Job Changes and Wage Changes: Estimation with Measurement Error in a Binary Variable*

Adele Bergin

This Draft: 22nd July 2013

Abstract

Many studies of labour market dynamics use survey data so it is valuable to know about the quality of the data collected. This paper investigates job transitions in Ireland over the period 1995 to 2001, using the Living in Ireland Survey, the Irish component of the European Community Household Panel. In applied work on job mobility, researchers often have to rely on self-reported accounts of tenure to determine whether or not a job change has taken place. There may be measurement error in these responses and consequently observations may be misclassified. The paper finds that there are substantial inconsistencies or measurement error in the responses used to determine job changes so there is a risk of misclassifying cases as being job changes when truly they are job stays and vice versa. The paper explores the impact of misclassification in a model of job change using an estimator developed by Hausman, Abrevaya and Scott-Morton (1998). It finds that ignoring misclassification may substantially underestimate the true number of job changes and it can lead to diminished covariate effects. The paper then investigates the relationship between job mobility and wage growth. Misclassification in a binary explanatory variable causes attenuation in OLS estimates. A two-step approach to controlling for misclassification in job changes is adopted to estimate the wage effects of job mobility. The paper finds that controlling for misclassification has a substantial impact on the estimated effect of changing jobs on wage growth.

JEL Classification: J62, C25

Keywords: Job Mobility; Binary variables; Misclassification

* Earlier versions of this paper were presented at National University of Ireland Maynooth in December 2008 and March 2011 and the 12th IZA European Summer School in Labor Economics, 27th April 27 to 3rd May, 2009. The author would like to thank Donal O'Neill, Aedin Doris, Olive Sweetman, Tom O'Connor and Arthur van Soest for helpful comments and suggestions.

1. Introduction

Many studies of labour market dynamics use survey data. Therefore it is important to know about the quality of the data collected. There may be ambiguity in a survey question, respondents may misunderstand the question, they may have an incentive to misreport, they may have poor recall or responses may be coded incorrectly. This paper investigates job mobility or, more specifically, employer changes over the period 1995 to 2001 using the Living in Ireland Survey (LIS), the Irish component of the European Community Household Panel (ECHP). As is common with many surveys, there is no direct question in the LIS about job mobility; instead it is inferred from the responses of individuals to a question about tenure. The paper describes a potentially serious measurement error problem in the responses used to determine job changes. As a result, there is a risk of misclassifying cases as being job changes when truly no change has taken place and vice versa.[1]

The effect of a potentially misclassified job change dummy variable is examined in two cases: (1) where it is the dependent variable in a model examining the determinants of job change and (2) in a wage model where it is a regressor. Ignoring misclassification in estimating the determinants of job mobility can lead to estimates that are biased and inconsistent. A modified probit estimator developed by Hausman, Abrevaya and Scott-Morton (1998) is used to control for misclassification in the dependent variable. The paper finds that ignoring misclassification leads to the true number of job changes being substantially underestimated and to diminished covariate effects.

Then the relationship between job mobility and wage growth in Ireland and how it is affected by measurement error is explored. Most applied research on job mobility and wages tends to focus on issues such as whether heterogeneity in some unobserved individual characteristic can account for the effect of mobility on wages and/or address the possible two-way causation between job mobility and wage growth. These issues are also addressed here. Generally, studies find a positive wage effect of around 10 per cent associated with changing jobs which serves as a useful reference point for this paper. Campbell (2001) finds that the wage gain connected to changing jobs over a three-year period is around 10 per cent in the UK. Topel and Ward (1992) report a 10 per cent return to mobility for young men in the US. Abbott and Beach (1994) find that the average wage gain for Canadian women who change jobs is around 8-9 per cent. OECD (2010) find an average of a 3 to 4 percentage point wage premium associated with changing jobs for a range of European countries. Their estimate for the Irish wage premium is higher at around 9 per cent.

The main contribution of the paper is to control for misclassification in job mobility status when estimating the impact of job mobility on wage growth. A two-step approach is adopted. The key result is that ignoring misclassification leads to a significant downwards bias in the wage impact of job mobility and the paper provides an estimate of the wage effect of job mobility, which corrects for measurement error.

The paper is organised as follows: Section 2 examines the reasons for and incidence of reporting errors in labour market survey data and, in particular, focuses on studies relevant to job mobility. Section 3 describes the dataset and explores the extent of measurement error in the LIS data. Section 4 outlines the econometric approach. Section 5 presents some descriptive statistics on job

[1] Misclassification is a special case of measurement error which arises when the variable of interest is binary.

mobility and wage growth. Section 6 presents estimation results and some sensitivity analysis and Section 7 concludes.

2. Reporting Errors in Labour Market Survey Data

Many empirical studies of job mobility use survey data and usually surveys do not contain a direct question asking if the respondent has changed jobs in the past year. Instead job changes are inferred from the length of time an employee reports to have been with their current employer. There are several reasons to suspect that responses to questions about tenure are measured with error. Respondents may find it difficult to remember when they started working in their current job. Bound *et al.* (2001) describe studies that categorise the question and answer process in a survey as a four-stage procedure. These stages include understanding the question, recovering the information from memory, considering whether the information matches what was requested and communicating the response. Much of the measurement error literature focuses on the stage where respondents retrieve the information from memory. A general principle from this literature is that the longer the length of the recall period the greater the expected bias due to respondent retrieval error. Therefore we might expect respondents with longer tenure to be most likely to misreport tenure. In one sense, this does not pose a serious problem for calculating job changes as they are associated with people who have short tenures; provided those with longer tenures who misreport do not significantly underestimate their tenure. Farber (1999) and Ureta (1992) find a heaping of tenure responses at round counts of years or round calendar years and this rounding indicates that individuals do not provide precise responses about tenure.

There may also be ambiguity in the wording of the question about tenure or there may be changes to the wording of the question in other waves of a survey. Farber (1999) points to how the mobility supplements to the Current Population Survey in the US from 1951 to 1981 asked workers what year they "…started working at their present job or business" while in later years the supplement asked workers how many years they have "…been working continuously for the present employer". The earlier question refers to time on the present job rather than time with the present employer. Workers may experience other types of mobility (e.g. promotion, reassignment) which means that their time on the job will be shorter than their time with the employer. The interviewer notes for the LIS provide clarity in distinguishing between employer changes and other types of internal mobility as they state that the question refers to when they started working with their present employer even if there have been position changes with that employer. In addition, there were no changes to the wording of the question about tenure in the LIS. The interviewer notes for the LIS do not provide guidance on how to handle interrupted employment spells (in particular when someone returns to a previous employer). Farber (1999) indicates that if no reference is made to the continuity of employment that the natural inclination of workers will be to ignore interruptions of "reasonable" length.

Brown and Light (1992) examine the extent of measurement error in tenure responses in the Panel Study of Income Dynamics (PSID). They find that tenure responses are frequently inconsistent with calendar time.[2] In addition, they perform a validation exercise to gauge the accuracy of their measure of job changes. They adopt a range of definitions of job mobility (based on tenure

[2] The level of inconsistencies in reported tenure in the PSID is described in Section 3 where comparisons are made to the LIS data.

responses) and use them to partition the data into distinct jobs. They assess the accuracy of the various measures by comparing the number of jobs and the number of times each job is observed with those identified by the National Longitudinal Survey (NLS). The NLS contains unique employer codes which can be compared across interviews and so provides a more accurate count of the 'true' number of jobs.

Brown and Light (1992) investigate which definition of a job change performs best when there is measurement error in tenure data. One definition they employ is to assume that a job change has taken place whenever reported tenure is less than the time elapsed since the previous interview. If tenure was never misreported and if respondents never returned to previous employers then this method would identify job changes without error. They also examine defining a job change to occur whenever the change in tenure between adjacent interviews varies "too much" or "too little" in either direction. In another definition, a job change is defined whenever the change in tenure is not exactly equal to the change in calendar time between interviews. This permits no inconsistency in tenure responses within jobs. They also employ more flexible measures that permit various amounts of inconsistency in reported tenure within jobs, where the change in tenure differs from the change in calendar time by more than 6, 12, 18 and 24 months in either direction. As these latter definitions identify job changes when tenure changes by "too much" as well as by "too little" they are more likely to separate continuing jobs but less likely to link jobs that are truly separate than when job changes are defined as occurring whenever reported tenure is less than the time elapsed since the previous interview. They find the definition of job mobility that is the most accurate when compared to the NLS data is that a job change has occurred whenever reported tenure is less than the time elapsed between interviews. This is the definition of job change used in this paper.

These types of validation studies are also useful because they provide evidence on the magnitude of measurement error in tenure data. Bound *et al.* (2001) point out that few studies have investigated the quality of tenure data. An example is Duncan and Hill (1985) who present results from a validation study of a large manufacturing company in which administrative records are used to validate survey responses from workers. Overall they find very little evidence of bias. They find that reported tenure is typically quite accurate; 45 per cent of the sample accurately reported the year they were hired and 90 per cent were able to report year of hire to within one year. However, the unit of analysis is in terms of years and these types of error margins in a dataset could be problematic if we were to use the measure of tenure to calculate job changes. As job changes are identified from those who report short tenures the under or over reporting of tenure by a year, in particular by those with short tenures, could lead us to misclassify job changes and vice versa.

Weiss *et al.* (1961) also compare reported starting dates to employer records and find that 71 per cent of jobs in the prior 5 years had reported starting dates within one month of company records. They also find that validity significantly declines as a function of the length of time between the job start date and the date of interview. To capture job mobility, tenure, at least for those who have not been in their jobs long, needs to be reported accurately. These validation studies suggest that the quality of tenure data may not be sufficient to do this.

3. Measurement Error

3.1 Dataset

The LIS is the Irish component of the European Community Household Panel (ECHP) which began in 1994 and ended in 2001.[3] It involved an annual survey of a representative sample of private households and individuals aged 16 years and over and households were followed over time. The survey collected a wide range of data such as labour force status, education level, income, job attributes and firm characteristics.[4]

The panel dimension of the LIS is exploited to identify job changes. A revolving balanced panel of people aged 20 to 60 years, roughly the prime working age, is selected for each country. This means that individuals are included in the sample in every year that they meet this age restriction. A revolving balanced panel is preferable over a balanced panel as a balanced panel prevents the entry of younger people into the sample and so, over time, as the fixed sample ages the proportion of younger people would decline. Essentially, a revolving balanced panel allows younger people into the sample in later years and allows older people to leave the sample. In addition, they must also have completed the interview in each year in question. Finally, around 120 cases are deleted from the sample each year; these cases refer to where the respondent is working but the start date with their employer is missing in any year.

There is no explicit question in the LIS about whether or not a person has changed jobs; instead job transitions are inferred from responses to the question about when they started working with their present employer. If a person is employed in two consecutive years and in the second year they report a starting date that falls between the two interview dates we conclude that this person has changed jobs during that period. Table 1 shows the number of observations in the revolving balanced panel, the number of workers employed in consecutive two-year periods and the number of job changes each year.

< Table 1 here >

However, in the absence of exogenous job change information we cannot be certain that the number of job changes reported in Table 1 is correct. Responses to questions about tenure are frequently inconsistent. For example, in one interview a person may report that their job spell started in January 1995 while the following year they may report that it started in January 1993. The concern in this paper is not necessarily that tenure is misreported but rather that if tenure is misreported there is a risk that cases may be misclassified as job changes and vice versa. For example, suppose a worker is interviewed in January 1995 and January 1996 and in January 1996 they report a start date in January 1993. Using the measure of job change defined above, we would conclude that no job change has taken place between the interviews in 1995 and 1996. However, suppose this person cannot accurately recall when the spell started and they misreport their starting date to be January 1995. Then we would erroneously conclude that this person has changed jobs between their interviews in 1995 and 1996. Now, suppose that their true starting date is January

[3] Additional details on the structure of the LIS Survey and the sample design are available at: http://issda.ucd.ie/documentation/esri/lii-overview.pdf

[4] There was some attrition in the sample in the earlier years, although the representativeness of the sample was improved in 2000 with the addition of new households. These new entrants to the LIS sample have been excluded from the analysis.

1995 so that they have truly changed jobs between interviews but they misreport their starting date to be January 1993. In this case we would erroneously conclude that no job change has taken place between interviews. In an attempt to ascertain how reliable the responses to the question about when a worker started working with their current employer are the next section examines the consistency of these responses over time.

3.2 Consistency of Starting Dates within Jobs

Given the possibility of measurement error we need to ascertain how reliable the information on starting dates is and therefore how useful it is for deducing job changes. If there were no measurement error then starting dates would be constant within jobs. By separating the dataset into distinct jobs and comparing starting dates across interviews we can investigate how consistent the data is.

Beginning with the 1995 data, there are 1,163 workers but as 76 workers changed jobs a total of 1,239 distinct jobs are observable (see Table 1). For this year, the previous jobs of those who changed jobs are excluded from the analysis as we only have one observation on their previous jobs (the starting date in 1994) so we cannot check the consistency of responses whereas we can track the new jobs across following interviews. We start with 1,163 distinct jobs in 1995. In each subsequent year one of four alternatives occurs:

1) A worker can stay in their job so the total number of jobs remains the same and the job survives an additional year.
2) A worker can drop out of the sample if they become unemployed, leave the labour force for more than a year or are over the age of 60.Here the total number of jobs remains the same and we no longer observe that particular job.
3) A worker can change jobs and accordingly the total number of jobs increases by one.
4) There can be a new entrant to the sample. This could be someone from the revolving balanced panel who is now 20 and so was excluded in earlier years or a worker who was unemployed or out of the labour force may come back into the analysis and this would increase the total number of jobs observed by one.

This results in 2,529 jobs observable for various durations over the period 1995 to 2001. Of these, there are 1,755 jobs observable for more than one year and this set of jobs is considered in the analysis in this section (so there are at least two starting dates to compare for each job). Table 2 shows how many jobs display consistency in starting dates. Of the 1,755 jobs considered, only 352 or 20 per cent have the same reported starting date each year the job is observed. If we adopt a less stringent definition of consistency such as all starting dates being within 3 months of each other then 37 per cent of jobs meet the criterion. If we relax the criterion further to where all starting dates are within 6 months of each other then 42 per cent of jobs display consistent responses. This leaves 1,014 or 58 per cent of jobs that survive for more than one year where all starting dates do not fall within 6 months of each other.

< Table 2 here >

This level of inconsistency in starting dates is quite alarming; however, it is in line with what has been found in other datasets. Brown and Light (1992) take a sample from the Panel Study of Income Dynamics (PSID) from 1976 to 1985 and partition the data into distinct jobs in an analogous fashion.

They find that only 7 per cent of jobs have identical starting dates in each year the job is observed, while 54 per cent of jobs have starting dates that do not all fall within 6 months of each other.

Brown and Light (1992) highlight another aspect of this definition of consistency that may be quite restrictive. Suppose a job is observed in every year of the survey and every reported starting date is equal with the exception of one which is different to the others by 7 months, then this job will not meet any of the measures of consistency define above. They argue most researchers would agree that this outlier could be 'fixed' to match up to the other observations for that job. Therefore, the measure of consistency used in Table 2 can be extended by requiring that only a majority of starting dates for jobs be in agreement.

Table 3 shows how many jobs have a majority of starting dates in agreement. A total of 654 jobs or 37 per cent have a majority of starting dates in agreement, while 84 per cent of all jobs identified have a majority of starting dates that are within 3 months of each other. The bottom panel of the table reports comparable statistics for the PSID taken from Brown and Light (1992). As before, the magnitudes of the consistency measures are broadly comparable with the Irish data. Given that both datasets display similar discrepancies, it is likely that any study using a similar question to deduce job changes contains measurement error.

<Table 3 here>

The method for dividing the dataset into separate jobs uses job changes to identify when one job ends and another one begins. The analysis in this section implies that the measure of job change may not accurately identify the true number of job changes i.e. there are probably cases identified as job changes when no change in jobs took place and vice versa. This means we may over or underestimate the true number of jobs and therefore the level of inconsistent starting dates within jobs.

Tables 2 and 3 focus on the extent and magnitude of inconsistencies evident in the data and it is clear that there is the possibility of substantial measurement error. In this paper, the main concern about measurement error is not directly that starting dates are misreported but rather that the misreporting of starting dates may cause cases to be misclassified as job changes and vice versa.

There are cases where it is very unlikely a job change has taken place, even though there are inconsistencies in starting dates, such as if the reported starting dates are sufficiently long ago. Of particular concern are inconsistencies in jobs where reported tenure is low. For example, suppose we observe a job every year between 1995 and 2001; it is more likely that this person has changed jobs at some point over this period and it has not been captured if the inconsistency in starting dates falls close to or within that period. However, if all inconsistencies in reported starting dates refer to a time period sufficiently far back then it is more likely that this person hasn't changed jobs recently and just cannot accurately recall when they started working in their current job.

Table 4 examines the timing of inconsistencies in reported starting dates within jobs. It reports how many of these jobs have the dates of all inconsistencies occurring at least three years prior to the date that we first start observing the job.[5] There are 722 jobs where all discrepancies fall reasonably

[5] For example, if we observe a job for the first time in 1995, this measure counts all jobs where each inconsistency refers to dates earlier than or in 1992.

far in the past so that these are probably truly continuing jobs. However, there are 681 jobs where the reported inconsistencies are more recent and it is more likely in these cases that we have linked jobs that are distinct or divided continuing jobs.

<Table 4 here>

This section focussed on examining discrepancies in reported starting dates within jobs. In the remainder of the paper, the focus is on how measurement error may lead us to misclassify a worker in a given year as being a job changer and vice versa so the unit of analysis switches from jobs to workers.

4. Econometric Approach

This section begins by discussing the Hausman *et al.* (1998) estimator to control for misclassification in the dependent variable in a discrete response model. A discrete choice model can be used to examine the decision to change jobs. Given the level of inconsistencies in the data, it is likely that incorrect inferences have been made about whether or not a worker has changed jobs so it is essential to control for misclassification. Measurement error in a binary variable results in misclassification i.e. some observations are misclassified as a zero when the variable is actually a one, and vice versa. In a linear regression model measurement error in the dependent variable only affects the precision of coefficient estimates; however the same problem leads to estimates that are biased and inconsistent in a nonlinear model.[6] Hausman, Abrevaya, and Scott-Morton (1998) use Monte Carlo simulations to demonstrate that even small amounts of misclassification can lead to substantially biased parameter estimates in a probit model. Then Section 4.2 investigates the relationship between job mobility and wage growth and how it is affected by misclassification so here the mismeasured binary variable is an explanatory variable. The econometric problems associated with estimating the impact of job mobility on wage growth, namely unobserved heterogeneity, possible two-way causation between job mobility and wage growth and measurement error in capturing job changes are also outlined. Section 4.3 describes the two-step empirical strategy used to deal with misclassification in a binary explanatory variable when estimating the impact of job mobility on wage growth.

4.1 Binary Choice Model with Misclassification

Hausman *et al.* (1988) develop a maximum likelihood estimator to control for misclassification in discrete dependent variables that consistently estimates the coefficients of a model and also the extent of misclassification. The decision to change jobs can be set in the usual latent-variable specification of the binary choice model.[7]

Let y_i^* be a continuous unobservable latent variable that represents the potential for a worker to change jobs:

$$y_i^* = x_i^{'}\beta + \varepsilon_i \qquad \text{where i=1, 2....n} \qquad (1)$$

and ε_i is an independently and identically distributed error term. We cannot observe y_i^* ; instead we observe whether a worker changes jobs or not so for each worker there is a threshold level, y_i^*, at or

[6] See Hausman (2001) for a discussion of the effects of measurement error in dependent variables.
[7] The details of the estimator come from Hausman *et al.* (1998).

above which they change jobs otherwise they stay in their job. The true response (or what we would observe if there was no measurement error), $\tilde{y}_i$, is given by:

$$\begin{aligned} \tilde{y}_i &= 1 \text{ if } y_i^* \geq 0 \\ &= 0 \text{ otherwise} \end{aligned} \tag{2}$$

However the classification of workers as having changed jobs or not is observed with error so let y_i denote observed job changes. Two types of misclassification can occur, so let the probability that a job stay is misclassified as a job change be given by α_0 and the probability that a job change is misclassified as a job stay be given by α_1. These probabilities depend on the true value, $\tilde{y}_i$, so the extent of misclassification depends on how good a proxy y_i is of $\tilde{y}_i$. The misclassification probabilities are assumed to be independent of the covariates, x_i, conditional on the true response, more formally: [8]

$$\alpha_0 = \Pr(y_i = 1|\tilde{y}_i = 0) = \Pr(y_i = 1|\tilde{y}_i = 0, x_i) \tag{3}$$
$$\alpha_1 = \Pr(y_i = 0|\tilde{y}_i = 1) = \Pr(y_i = 0|\tilde{y}_i = 1, x_i) \tag{4}$$

Let F(.) denote the cdf of ε_i. The probability that an observation is truly a job change is given by:

$$\Pr(\tilde{y}_i = 1|x_i) = \Pr(x_i'\beta + \varepsilon_i > 0) = F(x_i'\beta) \tag{5}$$

The probability that an observation is classified as being equal to one $\left(\Pr(y_i = 1|x_i)\right)$ is given by the probability that it has been correctly classified as being equal to one $(1 - \alpha_1)$ multiplied by the probability that it is truly equal to one $\left(F(x_i'\beta)\right)$ plus the probability that it has been incorrectly classified as being equal to one (α_0) multiplied by the probability that it truly is not equal to one $\left(1 - F(x_i'\beta)\right)$ as follows:

$$\Pr(y_{i=1}|x_i) = (1 - \alpha_1)F(x_i'\beta) + \alpha_0\left(1 - F(x_i'\beta)\right) = \alpha_0 + (1 - \alpha_0 - \alpha_1)F(x_i'\beta) \tag{6}$$

The expected value of the observed dependent variable y_i is given by:

$$E(y_i|x_i) = \Pr(y_i = 1|x_i) = \alpha_0 + (1 - \alpha_0 - \alpha_1)F(x_i'\beta) \tag{7}$$

When there is no misclassification ($\alpha_0 = \alpha_1 = 0$), this collapses to usual expression $F(x_i'\beta)$. Assuming ε_i are normally distributed, the log-likelihood function for the probit model with misclassification is:

$$\begin{gathered} \ln L = \sum_{i=1}^{n} \left\{ y_i \ln\left(\Pr(y_i = 1|x_i)\right) + (1 - y_i)\ln\left(\Pr(y_i = 0|x_i)\right) \right\} = \\ n^{-1} \sum_{i=1}^{n} \left\{ y_i \ln\left(\alpha_0 + (1 - \alpha_0 - \alpha_1)\Phi(x_i'\beta)\right) + (1 - y_i)\ln\left(1 - \alpha_0 - (1 - \alpha_0 - \alpha_1)\Phi(x_i'\beta)\right) \right\} \end{gathered} \tag{8}$$

where $\Phi(.)$ denotes the cdf of the normal distribution

Maximising the log-likelihood function given in (8) with respect to α_0 , α_1 and β yields consistent and efficient estimates of β as well as the probabilities of misclassification.

[8] Hausman *et al.* (1998) show how the model can be extended to allow for covariate-dependent misclassification.

The conditions for identification of α_0 , α_1 and β are similar to those for the traditional binary choice model. One additional assumption is needed, namely that the misclassification probabilities are not very large, specifically, $\alpha_0 + \alpha_1 < 1$.[9] When this assumption is not imposed the estimator cannot distinguish between the parameter values $(\alpha_0, \alpha_1, \beta)$ and $(1 - \alpha_0, 1 - \alpha_1, -\beta)$. Imposing this assumption excludes this situation because $\alpha_0 + \alpha_1 < 1$ implies $(1 - \alpha_1) + (1 - \alpha_0) > 1$. This implies that if $\alpha_0 + \alpha_1 > 1$ but we impose $\alpha_0 + \alpha_1 < 1$ the estimates of β will have the wrong sign. The assumption guarantees that $\alpha_0 + (1 - \alpha_0 - \alpha_1)\Phi(x_i'\beta)$ is strictly increasing in $x_i'\beta$ as $\Phi(.)$ is strictly increasing.

The model parameters are identified from the nonlinearity of $\Phi(.)$. Estimating α_0 and α_1 is only possible because they enter (8) additively. This can be demonstrated by taking limits of $E(y_i|x_i)$ as $x_i'\beta$ tends to $-\infty$ and $+\infty$ in (7).

$$\lim_{x_i'\beta \to -\infty} E(y_i|x_i) = \alpha_0 \ \text{ and } \lim_{x_i'\beta \to +\infty} E(y_i|x_i) = 1 - \alpha_1 \tag{9}$$

The identification of the misclassification probabilities comes from cases where $\Pr(\tilde{y}_i = 1|x_i)$ is close to 0 and 1 i.e. where $x_i'\beta$ is big in magnitude. The misclassification rates are assumed to be constant and depend only on the true value, $\tilde{y}_i$, so the probability of misclassifying a job stay, α_0, is identified from cases that have extremely negative characteristics in terms of job mobility and so are very unlikely to truly be job changes. However, as some constant proportion, α_0, are misclassified as job changes, $\Pr(y_i = 1|x_i)$ does not drop below α_0 no matter how negative $x_i'\beta$ is. In a similar fashion, the probability of misclassifying a job change as a job stay, α_1, is estimated from cases where $x_i'\beta$ is very large and positive and so these cases have a very high probability of truly being job changes but some proportion are misclassified. Therefore $\Pr(y_i = 1|x_i)$ will not rise above $1 - \alpha_1$.

This estimator, or variants and extension of it, have been used in a wide range of empirical applications where researchers suspect there is measurement error in a discrete variable. For example, Artís *et al.* (2002) investigate insurance fraud, Brachet (2008) and Kenkel *et al.* examine smoking data, Caudill and Mixon (2005) are interested in undergraduate student cheating, Dustmann and Van Soest (2001, 2004) examine language indicators, Flathmann and Sheffrin (2003) investigate self-reported non-compliance in completing tax returns and Jensen *et al.* (2011) examine patent applications.

4.2 Econometric Issues

The standard model for estimating the impact of changing jobs on wage growth is:

$$\Delta \log(w_{it}) = \beta m_{it}^* + \delta x_{it} + \varepsilon_{it} \tag{10}$$

where w_{it} is the natural logarithm of the wage of individual i at time t, m_{it}^* is a dummy variable indicating whether a job change has taken place between t-1 and t, x_{it} is a vector of personal, job and firm characteristics and ε_{it} is a random component that is mean zero and is uncorrelated with m_{it}^* and x_{it}. The key parameter of interest, β, captures the average percent difference in wage growth between job changers and job stayers adjusted for worker and job characteristics. Pooled OLS estimation of (10) is likely to produce biased estimates because of: (1) unobserved

[9] Hausman *et al.* (1998) refer to this as the 'monotonicity condition'.

heterogeneity, (2) the endogeneity of job mobility and (3) measurement error in capturing job changes. The first two issues have been tackled in the literature and the main contribution of the paper is to control for measurement error in job mobility when estimating its impact on wage growth.

Unobserved Heterogeneity

There may be unobserved factors that affect both wage growth and the decision to change jobs that can cause bias in the estimates. Intuitively, we want to compare the wage growth of a job changer with what they would have received had they stayed in their job. The estimate of β from (10) is unlikely to provide an accurate measure of this effect, if the average wage growth of stayers does not reflect the average wage growth job changers would have received had they stayed in their jobs. For example, in the mover-stayer model of Blumen *et al.* (1955), stayers experience higher wage growth than changers because they have some underlying personal characteristic that makes them more likely to stay in their job which also makes them more productive, which leads to higher wage growth. Therefore, the mover-stayer model suggests the estimate of β from (10) may be biased downwards.

Several techniques are used in the empirical literature to overcome this problem. One approach suggested by Bartel and Borjas (1981) and developed by Mincer (1986) is to use a proxy for the wage growth changers would have obtained had they not changed jobs. Mincer uses wage growth of those who do not change jobs in the current period but who change jobs in the following period as the proxy. The returns to mobility are then measured as the difference between the wage growth of workers who change jobs in the current period and the wage growth of workers who do not change jobs in the current period but do change jobs in the following period. This approach has been used by Abbott and Beach (1994), Campbell (2001) and Keith and McWilliams (1999). The key assumption is that workers who stay in their job in the initial period and who change jobs in the subsequent period are more similar, in terms of unobservable characteristics, to those who change jobs in the initial period than workers who stay in their jobs in both periods.

More recently, because of the availability of panel data, the issue of unobserved heterogeneity has been dealt with in a fixed effects estimation framework:

$$\Delta\log(w_{it}) = \beta m_{it}^{*} + \delta x_{it} + \varepsilon_{it} \qquad (11)$$
$$\text{where } \varepsilon_{it} = \eta_i + \upsilon_{it}$$

The error component has two distinct parts: η_i captures unobservable individual-specific effects that can vary across individuals but are constant over time for each individual and υ_{it} is assumed to be uncorrelated with the observed and unobserved characteristics across individuals and time.[10,11]

Unobserved heterogeneity is usually handled using a fixed effects or random effects model. The fixed effects model allows η_i to be correlated with other regressors. The estimator transforms all variables to deviations from their sample means for all time periods, so η_i drops out of the equation because it is constant over time. This framework has been used by Davia (2005), Le Grand and Tahlin

[10] In contrast, the pooled OLS model assumes that the intercept is common across all individuals.
[11] This assumes the only source of unobserved heterogeneity is at the individual level. The error term could be expanded to include unobservable job-specific effects which might capture, say, the quality of the job match (see for example, Light and McGarry (1998)).

(2002), Light and McGarry (1998), Naticchiono and Panigo (2004), Munasinghe and Sigman (2004) and Pavlopoulos *et al.* (2007). The key difference between the fixed and random effects models is that the random effects model assumes that the η_i are uncorrelated with the other regressors in the model; a quite restrictive assumption. Section 6 reports results from tests for individual effects and tests that discriminate between fixed and random effects models.

Reverse Causality

One source of endogeneity in (10) is two-way causation; not only is wage growth affected by job mobility but job changes may occur in anticipation of higher wage growth. If this feedback occurs from wages to job mobility then m_{it}^* will be correlated with υ_{it} in (11) as m_{it}^* depends on $\Delta \log(w_{it})$ which directly depends on υ_{it}. While it is reasonable to assume that there is no contemporaneous correlation between job mobility and the error term, the error term and future job changes may be correlated if workers decide to change jobs in the future based on shocks to wage growth in the past. In this case, the assumption of strict exogeneity underlying the random effects and fixed effects estimators (conditional on the unobserved individual effect in the case of fixed effects) can fail. Generally, random effects and fixed effects will be inconsistent if job mobility in some time period is correlated with υ_{it}. (Note, pooled OLS does not require all explanatory variables to be strictly exogeneous.)

One approach in applied work to deal with this problem is to use an instrument for mobility status. Possible instruments suggested in the literature include housing tenure status, job satisfaction (in particular the components of job satisfaction that do not refer to satisfaction with wages) and dummies for the region in which a person lives. Davia (2005) uses the predicted probabilities from a probit model of job change as an instrument for job mobility.

Misclassification

Measurement Error in Binary Regressors

Another source of potential bias in (10) is misclassification of job changes which can result in attenuation bias. Define m_{it} to be a noisy indicator of the binary variable m_{it}^*. Then we can write the observed value, m_{it}, as the sum of the true value, m_{it}^*, plus a measurement error, u_{it}:

$$m_{it} = m_{it}^* + u_{it} \tag{12}$$

where u_{it} is mean zero. When $m_{it}^* = 1$, m_{it} can only take on two values; 1 if it is correctly classified so $u_{it} = 0$, or $m_{it} = 0$ so $u_{it} = -1$. When $m_{it}^* = 1$, m_{it} can never overestimate/over-report the true value. Likewise, when $m_{it}^* = 0$, m_{it} can never underestimate the true value; u_{it} is either 0 or +1. Therefore the measurement error is negatively correlated with the true variable so misclassification in a dummy variable leads to non-classical measurement error.

Aigner (1973) and others have shown that when a binary regressor is misclassified the least squares coefficient estimates are biased towards zero and that additional assumptions or knowledge about the extent of misclassification in the data is needed to correct the estimates. To illustrate this, consider the model given in (10) where m_{it}^* denotes true job changes. Suppose we do not observe m_{it}^* but we observe m_{it} (as defined in (12)), which misclassifies some of the observations. Similar to section 4.1, let α_0 denote the probability that a true job stay is misclassified as a job change i.e. $\alpha_0 = Pr(m_{it} = 1|m_{it}^* = 0) = Pr(u_{it} = 1|m_{it}^* = 0)$ and α_1 denote the probability that a job change is misclassified as a job stay i.e. $\alpha_1 = Pr(m_{it} = 0|m_{it}^* = 1) = Pr(u_{it} = -1|m_{it}^* = 1)$.[12] Let μ denote the mean of m_{it}^*. Since m_{it}^* is a binary variable, μ corresponds to the probability of truly changing jobs; the probability that m_{it}^* is equal to 1. It follows that $\Pr(m_{it} = 1) = (1 - \alpha_1)\mu + \alpha_0(1 - \mu)$. In what follows, let $\Pr(m_{it} = 1) = (1 - \alpha_1)\mu + \alpha_0(1 - \mu) = p$ for simplicity.

First, consider a model with a single binary regressor:

$$\Delta \log(w_{it}) = \beta m_{it}^* + \varepsilon_{it} \tag{13}$$

However, we cannot observe m_{it}^* only the mismeasured proxy m_{it}, from (12), so:

$$\begin{aligned} \Delta \log(w_{it}) &= \beta m_{it} - \beta u_{it} + \varepsilon_{it} \\ &= \beta m_{it} + (\varepsilon_{it} - \beta u_{it}) \end{aligned} \tag{14}$$

Using m_{it} as a proxy for m_{it}^* means the measurement error becomes part of the error term in (14) and therefore creates an endogeneity bias. Estimating (14) yields an OLS estimator for β with a probability limit:

$$plim\hat{\beta}_{OLS} = \beta\left(\frac{{}^{\mu}/_{p}\,(1 - \alpha_1 - p)}{(1 - p)}\right) = \beta\gamma^0, \quad \text{where } \gamma^0 = \frac{{}^{\mu}/_{p}\,(1 - \alpha_1 - p)}{(1 - p)} \tag{15}$$

where γ^0 is the attenuation coefficient in a model with a single misclassified regressor. As α_0, α_1, μ and p are all greater than zero but less than one and $\mu \approx p$, the attenuation coefficient γ^0 given in (15) is less than one which implies that the OLS estimate of β is biased towards zero. Without knowledge about the misclassification rates, α_0 and α_1, and the probability that an observation is truly a job change, μ, we cannot identify the true β from our data. Furthermore, for very high levels of misclassification the expression for γ^0 could be negative yielding an OLS estimate of the wrong sign (Kane, Rouse and Staiger (1999)).

Attenuation bias is typically exacerbated in multivariate regression (Angrist and Krueger (1999)). Card (1996) and others have shown that the attenuation factor in this case is given by:

$$\gamma = \frac{\gamma^0 - (R^2/(1 - \alpha_0 - \alpha_1))}{1 - R^2} \tag{16}$$

where γ^0 is the attenuation factor from the model with no other covariates, from (15), and R^2 is the theoretical R^2 from a regression of observed job changes on the other explanatory variables in the model.

Misclassification will cause both OLS and fixed effects estimates to be biased towards zero and inconsistent. However, the bias is likely to be amplified in the fixed-effects estimates (Bound *et al.* (2001)). Correctly measured explanatory variables tend to be correlated across time so there is typically much less within-group variation in these variables than in the measurement error (as this will tend to exhibit weak or no serial correlation). Therefore, misclassification in fixed effects models

[12] As before, this assumes that the misclassification rates are constant across individuals and time and that they only depend on the true value m_{it}^* and not on the other covariates in the model.

tends to reduce the variance in the signal relative to the variance in the noise so attenuation bias in this model can be more severe than both measurement error bias and heterogeneity bias in a pooled OLS model.

Implications for Instrumental Variable Estimation

With a misclassified binary regressor, conventional instrumental variable estimation does not yield a consistent estimate of β. A valid instrument must be correlated with the true value, m_{it}^* , and uncorrelated with the error term which is made up of the random error ε_{it} and the misclassification error u_{it}. As u_{it} is correlated with m_{it}^*, any variable (potential instrument) which is correlated with m_{it}^* will also generally be correlated with the measurement error. If an instrument is available, IV estimation will remove the correlation between m_{it}^* and ε_{it} but not between m_{it}^* and u_{it} and so the IV estimate of β will be biased.

In a case with a single binary (misclassified) explanatory variable the IV estimate of β is biased by a factor $1/(1-\alpha_0-\alpha_1)$ (see Angrist and Krueger (1999), Kane, Rouse and Staiger (1999)). As $0<\alpha_0+\alpha_1$ and generally $\alpha_0+\alpha_1<1$, the IV estimate will be biased upwards. In a bivariate regression with a mismeasured binary explanatory variable the OLS estimate is biased downwards and the IV estimate is biased upwards so these estimates can be used to bound the true coefficient.

Approaches in the Literature

There are several approaches to dealing with measurement error in binary regressors. One is to exploit external estimates of misclassification rates as might be available from validation surveys. For example, Freeman (1984) and Card (1996) examine the impact of union membership on wages using a validation survey that has both employer and worker reports of union status to estimate the misclassification rates in the reporting of union status. Kane, Rouse and Staiger (1999) adopt develop a generalised method of moments estimator to obtain consistent estimates when there are two noisy reports of the regressor. Another approach assumes additional sample information can be used. For example, Mahajan (2006) assumes that additional information, in the form of a second variable, is available that is correlated with the unobserved true variable but not related to the misclassification in the binary variable and shows how this can be used to identify and estimate a nonlinear model with a misclassified binary regressor.

Other authors derive bounds for the estimates. Bollinger (1996, 2001) establishes bounds for the true coefficients in a linear regression when a binary regressor is mismeasured. Bollinger also shows how these bounds can be made tighter if information about the misclassification rates is available. Frazis and Loewenstein (2003) extend the Hausman *et al.* (1998) estimator and compute bounds for the misclassification rates without making functional form assumptions. They combine these bounds with the OLS coefficient to bound the true effect of the mismeasured explanatory variable.

Card (1996) and Frazis and Loewenstein (2003) provide an expression for the inconsistency in OLS estimates due to misclassification, assuming the other explanatory variables in the model are correctly measured:

$$\beta=\left(plim\hat{\beta}_{OLS}\right)\left[\frac{p(1-p)(1-\alpha_0-\alpha_1)(1-R^2)}{(p-\alpha_0)(1-p-\alpha_1)-R^2p(1-p)}\right] \tag{17}$$

A sensitivity analysis is performed Section 6 that uses this expression to provide "corrected" OLS estimates of the impact of job mobility on wage growth.

4.3 Empirical Strategy

A two-step approach to controlling for misclassification in estimating the effect of job mobility on wage growth is adopted. It follows Brachet (2008) and is similar to Dustman and van Soest (2001). The first step uses the Hausman *et al.* (1998) estimator to generates consistent estimates of the coefficients from a model of job change including the misclassification probabilities and, most importantly for this paper, the probability of truly being a job changer, $\left(\Pr(m_{it}^* = 1)\right)$. This yields a proxy for m_{it}^* that removes the impact of misclassification. In the second step, (10) is estimated using pooled OLS substituting in for m_{it}^* using the fitted probabilities that an observation is truly a job change calculated in the first stage. The coefficient estimates will be consistent provided the functional form for F(.) in the first step has been correctly specified.[13]

The same approach is used to control for both measurement error in job changes and unobserved heterogeneity; the wage growth equation in the second step is estimated using a fixed effects or a random effects estimator. In addition, if we have an instrument for job mobility, it can be used in the first stage of the procedure to create a proxy for m_{it}^* that removes both the impact of misclassification and the correlation with the error term. It is hard to find good instruments for job mobility. The paper attempts to control for reverse causality using non-wage elements of job satisfaction as instruments for job mobility.

5. Descriptive Statistics

The starting point is the sample of workers described in Section 3.1.[14] Two additional restrictions are placed on the sample: (1) only income from paid employment is considered so self-employed workers and farmers are excluded and (2) workers are excluded in any year that they report they are working part-time (less than 30 hours).[15, 16] These restrictions ensure some degree of homogeneity in the sample. The dependent variable is the change in log real gross hourly wages between period *t-1* and *t*. In each year, there are around 90 cases where wage data are not available and these person-year observations are dropped. The final sample consists of 1,206 workers and 5,346 person-year observations, observable for various durations over the period 1995 to 2001.

Table 5 provides preliminary evidence of the relationship between wage growth and job mobility. It shows that annual average wage growth for all workers is 8.5 per cent. The next two rows of the

---

[13] See Brachet (2008) for proof.

[14] The focus of this paper is on job-to-job transitions. The sample restrictions and definition of job mobility used means that workers cannot be unemployed or leave the labour force for any considerable amount of time between jobs (specifically by more than the amount of time between interviews). Therefore, the sample is probably a length time biased sample of job changers; in the sense that it may over-represent those who experience a relatively short period of unemployment between jobs or who leave the labour force for a relatively short period between jobs and under-represent those who are unemployed or leave the labour force for longer durations between jobs.

[15] This means that part-time workers are included in the sample in other waves if they are working full-time; however the results presented in the next section are similar to those when part-time workers are deleted entirely from the sample.

[16] Another reason for focussing on full-time workers is the possibility of measurement error in reported usual hours worked. Baum-Snow and Neal (2009) show that there is substantial measurement error in hourly wages for part-time workers.

table divides workers into job ‘movers’ and ‘stayers’, where job movers are those who change jobs *at some point* over the period 1995 to 2001 and stayers are those who are observed in the same job over the entire period. Job movers experience higher but more variable wage growth than those who stay in their jobs. An examination of real wage growth at different points in the distributions for job movers and stayers reveal that they are closest at the 25th percentile; however at the median and 75th percentiles wage growth of job changers is over 1.5 times that of job movers.

It may also be important to distinguish between different types of mobility when looking at wage effects associated with changing jobs.[17] Voluntary movers experience higher wage growth than involuntary movers and stayers, as expected. However involuntary movers record higher wage growth than job stayers which is surprising. The previous empirical literature has shown that involuntary job movers can experience wage losses, not just at the time of job change but that these losses can be permanent, especially if there is a period of unemployment between jobs (e.g. Garcia Perez and Rebollo Sanz (2005)). This effect is not evident in Table 5. This may be due to the fact that the sample is one where workers have a very high attachment to the labour force; workers need to be employed in consecutive two-year periods to be included in the sample. In addition, the time period under consideration is one of exceptional economic and employment growth in Ireland so it is possible that any reputation effects related to involuntary mobility may be reduced and/or job search costs may be lower as jobs are more plentiful.

The table also shows average wage growth for workers that move once and for workers that move more than once. Here, we do not distinguish between the types of move a worker may make, rather the number of moves. Workers who change jobs more than once experience higher wage growth than those who only move once. There do not appear to be any reputation effects associated with repeated mobiliy.[18]

<Table 5 here>

Controlling for the timing of job changes helps to disentangle whether the higher wage growth of job movers described in Table 5 is attributable to a discrete jump in wages at the time of starting a new job or if changing jobs shifts a worker onto a higher wage growth profile. Table 6 shows the annual average wage growth for job ‘moves’ and job ‘stays’. The unit of analysis has shifted from people in Table 5 to person-year observations. There are very large and variable wage gains related to job moves; a job move is associated with an average wage increase of around 17 per cent, compared to an average wage increase of around 6 per cent for a job stay. Comparing these figures with those from Table 5 implies that the bulk of the wage increase associated with job mobility happens at the time of changing jobs. The table also shows that wage growth is greatest for voluntary moves and that involuntary moves are associated with wage gains higher than those of job stays. Wage growth does not differ much depending on the whether it is the first move that we observe a worker making over the period or their second or third move etc over the period.

<Table 6 here>

[17] Other movers are those who do not state a reason for their job separation or who experience different types of mobility (e.g. they experience both a voluntary and an involuntary quit).
[18] Of course, we observe people at different stages in their working lives and the analysis cannot control for previous mobility history.

6. Results

This section presents econometric estimates of the effect of changing jobs on wage growth. It begins with pooled OLS results, which give an idea of the initial correlation between job mobility and wage growth. It then controls for unobserved heterogeneity and investigates whether there are differential wage impacts depending on the type of mobility. Then the results for controlling for misclassification in job changes are presented. A sensitivity analysis illustrates the effect misclassification has on the estimated wage effect associated with job mobility. Finally, an attempt is made to control for the bias due to the reverse causality between job mobility and wages.

Table 7 shows the pooled OLS estimates. The first specification contains no additional regressors (other than a constant term). The coefficient estimate on the job change dummy implies that the average increase in wage growth associated with changing jobs is around 10½ per cent and this effect is highly significant.

The second model includes the standard set of control variables that determine wage growth.[19] It includes traditional human capital variables such as age, experience and level of education. In addition, some job characteristics are controlled for, such as whether the job is in the public or private sector, firm size etc. Year dummies are included to control for changes in the macroeconomic environment. The estimate on the job change dummy variable is around 8 per cent indicating that some (around 2½ percentage points) of the higher wage growth associated with changing jobs is attributable to differences in observed characteristics.

The results show that wage growth declines with age and experience. This may reflect the fact that investment in human capital declines over the life-cycle. As expected, wage growth is higher for those who have at least a third level degree. The results reveal no significant difference between male and female, and public and private sector wage growth once differences in observable characteristics are controlled for. Workers in larger firms are expected to have higher wage growth, as larger firms are more likely to have internal labour markets etc, but the estimated effect is negative. Workers in sectors that are more exposed to market forces and where competitiveness is more important for growth, such as the manufacturing sector, have lower estimated wage changes. Working in the construction sector has a positive effect on wage growth, probably reflecting the fact that the sector was booming during the period. However, none of the sectoral wage effects are significant.

Next, unmeasured individual characteristics are controlled for using fixed effects and random effects models.[20] The key point to note about the estimation results is that mobility has a strong, positive and significant effect on wage growth even after controlling for unobserved heterogeneity. Overall,

[19] The explanatory variables are lagged by one year, so for job changers they refer to their characteristics in their previous jobs.

[20] The fixed effects model excludes time dummies as variables like age, and to some extent experience, change within individuals in the same way over time so the effect of a variable like age in a fixed effects model is interpretable more as a linear time trend. In addition, the estimate of job change on wage growth when age and experience are excluded and time dummies are included is practically identical. The education variables are also excluded as they have little within-person variation and reported changes in education level may reflect measurement error.

the fixed effects estimates are broadly comparable to the OLS estimates.[21] The estimates indicate that the impact of changing jobs on wage growth is around 11 per cent and the estimate is highly significant. This compares with the 8 per cent pooled OLS estimate and is consistent with the unobservable characteristic being negatively correlated with job mobility (and so the OLS estimate may be biased downwards). However, an F-test for the individual effects does not reject the null hypothesis that the individual effects are not jointly significantly different from zero.[22] This goes against the prediction of the mover-stayer model. The effects of the other variables included in the model are broadly comparable to the estimates from the pooled OLS model.

The random effects estimates are similar to the pooled OLS ones. A Breusch-Pagan Lagrange Multiplier test helps to discriminate between a random effects and pooled OLS model. The test indicates that we reject the null hypothesis that the variances across individuals are zero and this indicates that the random effects model is the preferred model. We might expect unobserved effects to be correlated with the explanatory variables i.e. that a fixed effects model is appropriate. A Hausman test can help decide between a fixed effects and random effects model. The Hausman test follows a chi-squared distribution and is equal to 22.47 with a corresponding p-value of 0.0962. This suggests we cannot reject the random effects model at the 5 per cent level of significance, but it can be rejected at the 10 per cent level of significance.[23]

<Table 7 here>

We expect to see differential wage impacts depending on the reason for job separation. Table 8 reports the random effects estimates of different types of mobility on wage growth.[24, 25] The first specification does not distinguish between different types of mobility. Model 2 distinguishes between voluntary, involuntary and other types of job changes.[26] Voluntary moves have the highest effect on wage growth, as expected; they are associated with a 14 per cent increase in short-term wage growth and this effect is significant at the 1 per cent level. Involuntary moves do not have a negative impact on wage growth; in fact the estimated effect is positive, although it is insignificant and much smaller than for voluntary moves. Although the sign of estimate is not as expected it is not significant and, as discussed before, may be attributable to the construction of the sample. In addition, it could reflect the tightness in the Irish labour market over the period under consideration

---

[21] As with the pooled OLS model the standard errors in the fixed effects model are clustered at the individual level. Fixed effects account for the time-constant part of the unobservable differences across people. However, it may be the case that unobserved random shocks that influence an individual at time *t* may also affect their behaviour at time *t+1* therefore leading to correlated errors within people.

[22] The F-test for the individual effects is calculated from a regression that does not use clustered standard errors because the test is based on the assumption of serially uncorrelated errors.

[23] The Hausman test is essentially testing whether the coefficient estimates from the fixed effects model are equal to those from the random effects model. As the fixed effects estimator only uses a small part of the information in the sample it usually has a large standard error. In practice, the Hausman test can very often accept the null hypothesis. Accepting the null hypothesis implies that either the two sets of coefficient estimates are reasonably close or it could indicate that the fixed effect estimates have very large standard errors and so we fail to reject the null hypothesis (Wooldridge (2002)).

[24] The same tests were conducted to help choose between the pooled OLS, fixed effects and random effects specifications. The random effects model is the preferred specification although the estimates from the three models are comparable.

[25] The models only include the relevant job change variable(s) and a constant term.

[26] Other types of job changes are those where the reason for changing jobs is not reported or the respondent chooses the 'other' category from a list of possible reasons for changing jobs.

where workers had many alternative employment opportunities. The estimated effect of 'other' types of mobility on wage growth is between the estimates of voluntary and involuntary mobility. Model 3 distinguishes between whether this is the first move a worker makes or whether they are observed changing jobs more than once during the observation window. The estimate connected with the job change being the first move observed is above that of a second or higher move but there is no evidence of wage penalties associated with repeated mobility. However, as mentioned before, it is important to note that in many cases we do not observe a workers' entire prior mobility history.

<Table 8 here>

Next we formally examine the impact of misclassification in job changes on the wage effects associated with job mobility using the two-step procedure described in Section 4.3. The first step uses the Hausman *et al.* estimator to control for misclassification in a model of job change.[27] Table 9 shows the marginal effects from a standard probit regression of the probability of job change and the Hausman *et al.* estimates that control for misclassification, with the later estimates used in the first stage of estimating the impact of job change on wage growth.[28] The table also shows the estimates of the misclassification probabilities. The estimated probability of misclassification for job stays, α_0, is very small at a ¼ of one per cent and the estimated probability of misclassification for job changes, α_1, is high at 51 per cent. Significance tests on α_0 and α_1 can be used as tests of misclassification. Workers who have truly changed jobs are more likely to be misclassified, as α_1 exceeds α_0. This means that the measure of job change is likely to undercount the true number of job changes. To put this estimate α_1 in context, the average mobility rate in the sample used is around 8 per cent and this estimate for α_1 implies that the true mobility rate is around 12 per cent. Hausman *et al.* also apply the estimator to a model of job change using US data from the January 1987 Current Population Survey from the Census Bureau. Their study provides external estimates of the misclassification probabilities. They estimate α_0 to be 6.1 per cent and α_1 to be 30.9 per cent.

Looking across the estimates from the misclassification and probit models, it is evident that when we allow for misclassification, the estimated coefficients have higher standard errors implying that errors in responses lead to a loss in estimation efficiency. The results also indicate that ignoring misclassification leads to diminished covariate effects.

Although both models indicate that the same factors determine job mobility the effect of misclassification in the dependent variable on the marginal effects of the various explanatory variables is sizeable. In the theoretical literature on job mobility, years of labour market experience is a key determinant of job change. Workers with less labour market experience are more likely to changes jobs as they have less knowledge of the labour market and their own preferences and abilities for different jobs. Both models provide findings that are consistent with this. However, in the probit model, an additional year of experience reduces the probability of changing jobs by 0.9 percentage points, while the marginal effect in the misclassification model is almost four times larger.

[27] The analysis only examines controlling for misclassification in the overall job change dummy variable.
[28] A series of models were run where the misclassification rates were allowed to depend separately on each of the covariates but either none of the additional probabilities estimated were significant or the models did not converge.

The models contain a range of individual controls that include household structure and personal characteristics. We may expect women to be more likely to change jobs as they have a weaker attachment to the labour force but the results do not indicate any significant gender difference in the probability of changing jobs. The marginal effect of having children is small and insignificant implying that the presence of children does not affect the probability of changing jobs. Workers may be less likely to change jobs if they are more constrained by non-market variables, such as being married or living in a couple but no significant effect is found.

The results also indicate that the negative effect of human capital on the probability of changing jobs is more marked in the misclassification model. For example, general human capital is proxied by education level and in the model incorporating misclassification the marginal effect of third level education is more than two times higher than in the probit model, although the effect is not significant. In addition, the marginal effects of higher levels of occupational attainment relative to those in elementary occupations are higher in the misclassification model. For example, the marginal effect indicates that those in a managerial occupation are almost 10 per cent less likely to change jobs than those who are in elementary occupations versus 4 per cent in the probit model. The results also show that workers who have undergone recent training are more likely to change jobs. This may reflect the fact that, typically, training is undertaken at the beginning of a job and there is a high hazard of new jobs ending early.

The job mobility models also contain variables that try to capture some job and firm characteristics. A variable to capture overskilling is included as it may signify a poor job match. A positive relationship between being overskilled and job mobility is found and the effect from the misclassification model of being overskilled is twice the impact from the probit model. A firm size effect is included to capture the fact that those working in a large firm may have more alternative employment opportunities within the firm and so are less likely to change jobs. The results indicate that workers in firms with more than 50 employees have a lower probability of changing jobs and, as before, the impact is more marked in the misclassification model.

Working in the public sector is found to exert a negative effect on the probability of changing jobs and the marginal impact from the misclassification model is more than twice the impact than in the probit model. The effect of the sector a worker was in the previous year (or for job changers the sector they previously worked in) is similar in both models but again the marginal effects are higher in the misclassification model. The results also show that workers in the construction and market services sector are more likely to change jobs than those in the nonmarket services sector.

The unemployment rate is included to control for factors such as access to alternative jobs and local labour market conditions. We expect the impact on the unemployment rate to be negative as a lower unemployment rate may signal to workers that jobs are more abundant and that job search is likely to result in them finding an alternative job. The impact of the unemployment rate is negative, as expected, but only insignificant in the probit model. It is likely that the unemployment rate variable is correlated with the time dummy variables. In fact, when the year dummies are dropped the effect of the unemployment rate is highly significant.

One way to demonstrate the differences between the two models is to graph the marginal effects of the variables. Figure 1 plots the marginal effect of experience from both models. The curves slope down as the probability of job change decreases as years of experience increases. The slopes of the curves are steep at lower values of experience and then flatten out at higher years of experience indicating that an additional year of experience reduces the probability of changing jobs but at a declining. Overall, the graph shows that the effect of ignoring misclassification error is large, especially at low values of experience.

<Figure 1 here>

The estimates from the misclassification model are used to construct the predicted probabilities that an observation is truly a job change. In the second step, this generated regressor is included instead of the job change dummy variable in the wage growth regression.[29,30]

The results from the second step indicate that the impact of changing jobs on wage growth is closer to 14 per cent when we control for misclassification (Table 9).[31] The comparable estimate from the model that ignores misclassification is around 8 per cent (see Pooled OLS model in Table 7). The results indicate that failing to control for misclassification leads us to seriously underestimate the wage effects of job mobility.

Sensitivity Analysis: Effect of Different Rates of Measurement Error on Estimates

This section illustrates the effect different rates of misclassification have on the estimates of job mobility in the wage growth regressions. It applies the formula for attenuation bias described in (17) and uses a range of misclassification rates to generate corrected OLS estimates. These adjusted estimates can be compared to the pooled estimate of 0.0794 from Model 2 in Table 7. Table 10 reports adjusted OLS estimates for different rates of misclassification. Using the first stage estimates of α_0 and α_1 from the previous section generates an adjusted OLS estimate of around 0.10, around 30 per cent above the pooled OLS estimate in Table 7.

The table also shows comparable corrected OLS estimates when α_0 is assumed to be equal to zero and α_1 varies between 1 per cent and 80 per cent. The corrected estimates indicate that when α_1 is low that the adjusted estimates are quite close to the pooled OLS one. However, as α_1 increases the adjusted estimate moves increasingly further away from pooled OLS estimate. In addition, the table reports corrected OLS estimates when α_0 is 1 per cent and 5 per cent and α_1 is assumed to be equal to zero. Even for these relatively low rates of misclassification, the adjusted OLS estimates are quite far away from the pooled OLS estimate. This stronger impact from misclassifying job stays arises from the fact that the proportion of observed job changes in the sample is around 8 per cent, so the proportion of job stays is 92 per cent and therefore the misclassification rate applies to a much higher number of cases.

<Table 10 here>

---

[29] The identification of the model comes from the fact that certain variables, such as whether a person reports if they are overeducated, have children, have undergone recent training and the national unemployment rate, are included in the model in the first stage but not in the second stage and also that the predicted probabilities are non-linear functions of the explanatory variables.

[30] The standard errors are adjusted in the second stage to take account of the additional variance due to the inclusion of generated regressor as described by Newey and McFadden (1994) and Murphy and Topel (1985).

[31] The results in Table 9 use a pooled OLS model in the second stage. The comparable estimates using a random effects model and a fixed effects model in the second stage are 13.77 per cent and 14.84 per cent respectively.

An additional difficulty with investigating the effect of job mobility on wage growth is the possibility of reverse causality. This can be addressed using an instrumental variable approach. We need instruments that are highly correlated with job mobility and that are uncorrelated with wage growth, so they have no independent effect on wage growth other than through job mobility. Here thenon-wage aspects of job satisfaction are used as instruments for job mobility.

Kristensen and Westergård-Nielsen (2004) and Gielen (2008) argue that job satisfaction may be a proxy for the worker's assessment of the quality of the match. Job satisfaction may capture unobserved aspects of work, such as the organisation of work, harsh working conditions etc. As such, job satisfaction should be a strong predictor of job mobility. However, the difficulty with using a measure of overall job satisfaction as an instrument for job mobility is that we expect a worker's satisfaction with earnings to dominate such as measure. Therefore it is likely that overall job satisfaction is highly correlated with wage growth. However, the LIS asks workers how satisfied they are with different aspects of their job, where satisfaction with earnings is only one component. Nonetheless, it may still be the case that satisfaction with earnings influences a worker's assessment of their satisfaction with other aspects of the job and this should be borne in mind when interpreting the results.

Table 11 shows the percentage of workers who are satisfied with various aspects of their jobs.[32] From the table, dissatisfaction with earnings is the most common source of dissatisfaction with the job.

<Table 11 here>

To assess whether satisfaction with wages affects satisfaction with other aspects of the job, Table 12 reports the percentage of workers satisfied with other areas of their jobs of those who are not satisfied with their earnings. The table shows that high proportions of workers are satisfied with other areas of their jobs even though they are unhappy with their earnings. This indicates that (dis)satisfaction with earnings may not influence satisfaction with other areas of the job. Therefore, the non-wage aspects of job satisfaction may be appropriate instruments for job mobility.

<Table 12 here>

The quality of the instruments can be checked by testing their significance in the first stage of the two-step approach. The results from the Hausman *et al.* estimator includes all the exogenous variables and all the instruments show that satisfaction with distance to job and working conditions are not significant and the coefficient on satisfaction with working hours has the incorrect sign and is only significant at the 10 per cent level. Consequently, these three variables are dropped from the analysis. Satisfaction with job security, type of work and working times are used as instruments for job mobility.[33] As before, the probability of truly being a job changer is calculated, and included in the wage growth equation in the second step. This controls for both misclassification and endogeneity.

[32] Workers are asked to indicate their degree of satisfaction with each area of their work on a scale of 1 to 6, where 1 indicates that they are not satisfied at all and 6 indicates that they are fully satisfied. In the table, satisfied corresponds to workers who report a level or 4 or above and not satisfied refers to those who report a satisfaction level of 3 or below.
[33] The results from the first-stage are available from the author on request.

As discussed earlier, we expect the IV estimates that don't control for misclassification to be biased upwards. Table 13 shows the results from the two-step procedure using IV but where misclassification is ignored i.e. the predicted probabilities in the first stage come from a standard probit model. The estimate on the job change dummy variable is around 26 per cent. This compares to the pooled OLS estimate of 8 per cent (see Table 7). As expected, the IV estimate is above the OLS one, but it is dramatically higher and arguably implausibly large.[34]

<Table 13 here>

Table 14 shows the results from the two-step approach controlling for both endogeneity and misclassification. The estimates of the misclassification probabilities from the first stage are practically identical to the estimates in Table 9. The second stage IV estimate implies that the impact of changing jobs on wage growth is around 13 per cent, when we control for misclassification. This is around half the IV estimate that ignores misclassification, implying that ignoring misclassification leads to a significant upwards bias in the IV estimate. In addition, the estimate is around 1.6 times the size of the pooled OLS estimate but quite similar to the estimate that controls for misclassification but ignores the possible reverse causality of job mobility. The results in tables 13 and 14 use pooled OLS in the second stage; very similar results are obtained when the random effects and IV approach are combined.

<Table 14 here>

Table 15 provides a summary of the various estimates of job mobility on wage growth.

<Table 15 here>

7. Conclusions

This paper investigates job mobility in Ireland over the period 1995 to 2001. It finds that there are substantial inconsistencies in responses to a question about tenure in the LIS. The extent of the measurement error is similar to what has been found in other studies. Survey data on tenure are very often used to deduce job changes and given the extent of response error evident in the data it is likely that cases are misclassified as job changes when they truly no job change has taken place and vice versa.

The decision to change jobs can be set in a binary choice framework. Misclassification in a binary dependent variable can lead to estimates that are biased and inconsistent so it is important to control for misclassification. An estimator developed by Hausman *et al.* is used to control for misclassification. The results indicate that, by ignoring misclassification, the true number of job changes is underestimated by around 50 per cent. The average mobility rate is calculated at around 8 per cent and the estimate for α_1, the misclassification rate for job changes, implies that the true mobility rate is around 12 per cent. In addition, the paper finds that ignoring misclassification leads to diminished covariate effects.

[34] This type of estimate is consistent with what Davia (2005) finds when she controls for endogeneity in job mobility using ECHP data. For most of the countries in her analysis, the estimates that control for endogeneity are multiples of the pooled OLS estimates (see Davia (2005), Table 2, page 24).

This paper also adds to the literature on the effect of job mobility on wage growth. It finds an OLS estimate of the effect of job mobility on wage growth of around 8 per cent. The wage effects differ depending on the reason for job separation. Voluntary job changes are associated with a 14 per cent increase in wage growth. However, there is no evidence of wage penalties associated with involuntary mobility. This may be attributable to the fact that the sample considered is one where workers have a very high attachment to the labour force or it may be due to the very high growth rates and tightness in the labour market over the time period under consideration.

The OLS estimate of the effect of changing jobs on wage growth may be biased due to unobserved heterogeneity, reverse causality and also because of measurement error. The paper finds that the effect of job mobility on wage growth persists even after controlling for unobserved individual heterogeneity. The magnitude of the estimates obtained from OLS regressions and regressions that control for unobserved heterogeneity are broadly in line with the existing empirical literature. For example, OECD (2010) finds a wage premium associated with changing jobs of around 9 per cent for Ireland which is very similar to what is found in this paper. However, these estimates ignore measurement error in job changes.

A two-step approach is adopted to control for misclassification in a binary explanatory variable. The paper finds that controlling for misclassification has a substantial effect on the estimated impact changing jobs has on wage growth; the effect is estimated to be closer to 14 per cent. Finally, controlling for reverse causality using an instrumental variables approach and ignoring misclassification produces an estimate that seems questionably high; however a more plausible estimate is obtained when the IV strategy is combined with the measurement error approach.

References

Abbott, M. and C. Beach (1994), "Wage Changes and Job Changes of Canadian Women", *Journal of Human Resources*, Vol. 29, No. 2, pp. 429-460.

Aigner, D.J. (1973), "Regression with a Binary Independent Variable Subject to Errors of Observation", *Journal of Econometrics,* Vol. 1, pp. 49-60.

Angrist, J.D. and A. B. Krueger (1999), "Empirical Strategies in Labor Economics", in *Handbook of Labor Economics*, Volume 3A, O. Ashenfelter and D. Card (eds), Amsterdam and New York: Elsevier Science.

Artís, M., Ayuso, M. and M. Guillén (2002), "Detection of Automobile Insurance Fraud with Discrete Choice Models and Misclassified Claims", *The Journal of Risk and Insurance,* Vol. 69, No. 3, pp. 325-340.

Bartel, A. and G. Borjas (1981), "Wage Growth and Job Turnover: An Empirical Analysis", in *Studies in Labor Markets*, Vol. 3, Rosen S. (ed.), pp. 65-84. University of Chicago Press.

Baum-Snow, N., and D. Neal (2009), "Mismeasurement of usual hours worked in the Census and ACS", *Economic Letters,* Vol. 102, Issue 1, (Jan., 2009), pp. 39-41.

Blumen, I. Kogen, M. and P. McCarthy (1955), "*The Industrial Mobility of Labor as a Probability Process"*, Ithaca, New York: Cornell University Press.

Bollinger, C.R. (1996), "Bounding Mean Regressions When a Binary Regressor is Mismeasured" *Journal of Econometrics*, Vol. 73, pp. 387-399.

Bollinger, C.R. (2001), "Response Error and the Union Wage Differential", *Southern Economic Journal*, Vol. 68, No. 1, pp. 60-76.

Bound, J., Brown, C. and N. Mathiowetz (2001), "Measurement Error in Survey Data", in J. Heckman and E. Leamer, eds, *Handbook of Econometrics*, Vol. 5. Elsevier, North Holland.

Brachet, T. (2008), "Maternal Smoking, Misclassification and Infant Health", MPRA Working Paper No. 21466. Available at: http://mpra.ub.uni-muenchen.de/21466/

Brown, J. and A. Light (1992), "Interpreting Panel Data on Job Tenure", *Journal of Labor Economics*, Vol. 10, No. 3 (Jul., 1992), pp. 219-257.

Campbell, D. (2001), "Estimating the Wage Effects of Job Mobility in Britain" Studies in Economics, No. 0117, Department of Economics, University of Kent.

Card, D. (1996), "The Effect of Unions on the Structure of Wages: A Longitudinal Analysis", *Econometrica*, Vol. 64, No. 4, pp. 957-979.

Caudill, A. and F. Mixon (2005), "Analysing Misleading Discrete Responses: A Logit Model Based on Misclassified Data", *Oxford Bulletin of Economics and Statistics,* Vol. 67, No. 1, pp. 105-113.

Davia, M. (2005), "Job Mobility and wage mobility at the beginning of the working career: a comparative view across Europe", ISER Working Paper No. 2005-03, Institute for Social and Economic Research.

Duncan, G.J. and D.H. Hill (1985), "An Investigation of the Extent and Consequences of Measurement Error in Labor-Economic Survey Data", *Journal of Labor Economics*, Vol. 3, No. 4, (Oct., 1985), pp. 508-532.

Dustmann, C. and A. van Soest (2004), "An Analysis of Speaking Fluency of Immigrants Using Ordered Response Models with Classification Errors", *Journal of Business and Economic Statistics,* Vol. 22, No. 3, pp. 312-321.

Dustmann, C. and A. van Soest (2001), "Language Fluency and Earnings: Estimation with Misclassified Language Indicators", *Review of Economics and Statistics*, Vol. 83, No. 4, pp. 663-674.

Dye, R. and D. McMillen (2007), "Teardowns and land values in the Chicago metropolitan area", *Journal of Urban Economics,* Vol. 61, pp. 45-63.

Farber, H. (1999), "Mobility and Stability: The Dynamics of Job Change in Labor Markets", in O. Ashenfelter and D. Card, eds, *Handbook of Labor Economics, Vol.3.* Elsevier, North Holland.

Flathmann, J. and S. Sheffrin (2003), *"Are Surveys of Taxpayers' Honesty Honest?"*, IIPF Congress Proceedings, Prague, August 2003.

Frazis, H. and M.A. Loewenstein (2003), "Estimating Linear Regressions with Mismeasured, Possibly Endogenous, Binary Explanatory Variables", *Journal of Econometrics*, Vol. 117, No. 1, pp. 151-178.

Freeman, R.B. (1984), "Longitudinal Analysis of the Effects of Trade Unions", *Journal of Labor Economics*, Vol. 2, No. 1, pp. 1-26.

Garcia Perez, J.I. and Y. Rebollo Sanz (2005), "Wage changes through job mobility in Europe: A multinomial endogenous switching approach", *Labour Economics,* Vol. 12, pp. 531-555.

Gielen, A.C. (2008), "Repeated Job Quits: Stepping Stones or Learning about Quality?", IZA Discussion Paper No. 3838.

Hausman, J., Abrevaya, J. and F.M. Scott-Morton (1998), "Misclassification of the dependent variable in a discrete-response setting", *Journal of Econometrics*, Vol. 87, pp. 239-269.

Hausman, J. (2001), "Mismeasured Variables in Econometric Analysis: Problems from the Right and Problems from the Left", *The Journal of Economic Perspectives,* Vol. 15, No. 4 (Aut., 2001), pp. 57-67.

Jensen, P., Palangkaraya, A. and E. Webster (2008), "*Misclassification between Patent Offices: Evidence from a Matched Sample of Patent Applications"*, *The Review of Economics and Statistics,* Vol. 93, No. 3, pp. 1063-1075. Kane, T., C.E. Rouse and D.O. Staiger (1999), "Estimating Returns to Schooling when Schooling is Misreported", National Bureau of Economic Research Working Paper No. 7235.

Keith, K. and A. McWilliams (1999), "The Returns to Mobility and Job Search by Gender", *Industrial and Labor Relations Review*, Vol. 52, No. 3, (Apr., 1999), pp. 460-477.

Kenkel, D., Lillard, D. and A. Mathios (2004), "Accounting for misclassification error in retrospective smoking data", *Health Economics*, Vol. 13, pp. 1031-1044.

Kristensen, N. and N. Westergård-Nielsen (2004), "Does Low Job Satisfaction Lead to Job Mobility?, IZA Discussion Paper No. 1026.

Le Grand, C. and M Tahlin (2002), "Job mobility and earnings growth", *European Sociological Review*, Vol.18, No. 4, pp. 381-400.

Light, A. and K. McGarry (1998), "Job Change Patterns and the Wages of Young Men", *The Review of Economics and Statistics,* Vol. 80, No. 2, pp. 276-286.

Mahajan, A. (2006), "Identification and Estimation of Regression Models with Misclassification", *Econometrica*, Vol. 74, No. 3, (May, 2006), pp. 631-665.

Mincer, J. (1986), "Wage Changes in Job Changes", in *Research in Labor Markets,* Rosen S. (ed.), NBER and University of Chicago Press, pp. 21-63.

Munasinghe, L. and K. Sigman (2004), "A Hobo Syndrome? Mobility, Wages and Job Turnover", *Labour Economics,* Vol. 11, pp. 191-218.

Murphy, K. and R. Topel (1985), "Estimation and Inference in Two-Step Econometric Models", *Journal of Business and Economic Statistics,* Vol. 3, No. 4, (Oct., 1985), pp. 370-379.

Newey, W. K., and D. McFadden (1994), "Large Sample Estimation and Hypothesis Testing", in R. F. Engle, and D. L. McFadden, eds, *Handbook of Econometrics*, Vol. 4. Elsevier, North Holland.

OECD (2010), "Institutional and Policy Determinants of Labour Market Flows", Chapter 3 in 2010 OECD Employment Outlook: Moving Beyond the Jobs Crisis.

Pavlopoulos, D., Fouarge, D., Muffels., R. and J. Vermunt (2007), "Who benefits from a job change: The dwarfs or the giants?", IRISS Working Papers, No. 2007-16, December 2007.

Topel, R. and M. Ward (1992), "Job Mobility and the Careers of Young Men", *The Quarterly Journal of Economics*, Vol. 107 (May, 1992), pp. 430-479.

Ureta, M. (1992), "The Importance of Lifetime Jobs in the U.S. Economy, Revisited", *American Economic Review*, Vol. 82, No. 1, pp. 322-335.

Weiss, D., Dawis, R., England, G. and L. Lofquist (1961), "Validity of Work Histories Obtained by Interview", Industrial Relations Center, University of Minnesota.

Wooldridge, J. (2002), Econometric Analysis of Cross Section and Panel Data, MIT Press, Cambridge, Massachusetts.

Tables & Figures

Table 1: Number of Workers and Job Changes

	1995	1996	1997	1998	1999	2000	2001
Revolving Balanced Panel	2,292	2,247	2,217	2,179	2,179	2,195	2,239
Number of workers	1,163	1,175	1,211	1,276	1,341	1,376	1,434
Number of Job Changes	76	85	102	139	146	184	156
Job Mobility Rate	6.5%	7.2%	8.4%	10.9%	10.9%	13.4%	10.9%

Table 2: Consistency of Starting Dates within Jobs

		Jobs with a Majority of Starting Dates:			
	No. of Jobs	Equal	Within 3 months	Within 6 months	Remaining jobs
Number of jobs	1,755	352	649	741	1,014
% of Jobs		20%	37%	42%	58%
Source: Living in Ireland Survey					
Number of jobs	3,318	246	1,170	1,514	1,804
% of Jobs		7%	35%	46%	54%
Source: Panel Study of Income Dynamics, taken from Brown and Light (1992)					

Table 3: Consistency of the Majority of Starting Dates within Jobs

		Jobs with a Majority of Starting Dates:			
	No. of Jobs	Equal	Within 3 months	Within 6 months	Remaining jobs
Number of jobs	1,755	654	1,471	1,513	242
% of Jobs		37%	84%	86%	14%
Source: Living in Ireland Survey					
Number of jobs	3,318	676	2,116	2,471	847
% of Jobs		20%	64%	74%	26%
Source: Panel Study of Income Dynamics, taken from Brown and Light (1992)					

Table 4: Timing of Inconsistencies within Jobs

	No. of Jobs	Equal Starting Dates	All inconsistencies at least 3 years prior to date job is first observed	Remaining jobs
Number of jobs	1,755	352	722	681
% of Jobs		20%	41%	39%

Table 5: Average Within-Person Wage Growth

	No. of People	*Mean*	*Standard Error*	*25th Percentile*	*50th Percentile*	*75th Percentile*
All Workers	1,206	0.085	0.005	0.01	0.06	0.13
Job Stayer	766	0.072	0.006	0.01	0.05	0.11
Job Mover	440	0.109	0.011	0.02	0.08	0.17
Voluntary Job Mover	223	0.118	0.014	0.03	0.10	0.17
Involuntary Job Mover	78	0.092	0.030	0.01	0.05	0.23
Other Movers	139	0.105	0.018	0.03	0.08	0.17
Move Once	256	0.101	0.015	0.01	0.07	0.16
Move more than Once	184	0.120	0.014	0.03	0.11	0.18

Table 6: Average Wage Growth for Job Stays and Job Moves

	No. of Person-Year Observations	*Mean*	*Standard Error*	*25th Percentile*	*50th Percentile*	*75th Percentile*
All Observations	5,346	0.070	0.004	-0.051	0.045	0.190
Job Stays	4,897	0.061	0.004	-0.051	0.041	0.172
Job Moves	449	0.167	0.019	-0.048	0.134	0.383
Voluntary Job Moves	282	0.200	0.025	-0.022	0.190	0.422
Involuntary Job Moves	103	0.094	0.040	-0.113	0.076	0.328
Other Moves	64	0.140	0.046	-0.054	0.073	0.334
First Move	232	0.177	0.026	-0.042	0.155	0.395
2nd + Move	217	0.157	0.029	-0.050	0.110	0.355

Table 7: Wage Growth Models

	No Additional Regressors		***Pooled OLS***		***Fixed Effects***		***Random Effects***	
	Coef	*Std Error*	*Coef*	*Std Error*	*Coef*	*Std Error*	*Coef*	*Std Error*
Job Change	**0.1064*** **	**0.0178**	**0.0794*** **	**0.0182**	**0.1122*** **	**0.0237**	**0.0805*** **	**0.0183**
Age			-0.0110**	0.0043	-0.0211	0.0267	-0.0113**	0.0044
Age Squared			0.0001***	0.0001	0.0003	0.0002	0.0002***	0.0001
Experience			-0.0019	0.0022	0.0062	0.0210	-0.0017	0.0022
Experience Squared			0.0000	0.0000	-0.0002	0.0002	0.0000	0.0000
Female			0.0014	0.0065			0.0016	0.0066
Education: (ref: Low Education)								
Education - Medium			-0.0087	0.0073			-0.0086	0.0074
Education - High			0.0078	0.0110			0.0080	0.0111
Public Sector			0.0070	0.0098	-0.0028	0.0282	0.0073	0.0100
No. Employees > 50			-0.0244***	0.0067	-0.0497***	0.0147	-0.0251***	0.0068
Occupation: (ref: Elementary Occupations)								
Manager			-0.0016	0.0111	-0.0076	0.0278	-0.0013	0.0113
Professional			0.0210**	0.0099	-0.0015	0.0268	0.0212**	0.0101
Clerk			0.0186**	0.0092	0.0106	0.0240	0.0190**	0.0093
Skilled			0.0125	0.0114	0.0385	0.0329	0.0128	0.0117
Sector of Origin: (ref: Non-Market Services)								
Ag., Mining & Utilities			0.0073	0.0184	0.1026*	0.0606	0.0080	0.0189
Manufacturing			-0.0108	0.0130	0.0269	0.0383	-0.0107	0.0133
Construction			0.0133	0.0181	0.0799**	0.0389	0.0137	0.0184
Market Services			0.0037	0.0097	0.0501*	0.0298	0.0039	0.0099
Year Dummies:								
1996			-0.0168	0.0167			-0.0166	0.0167
1997			0.0139	0.0155			0.0140	0.0155
1998			-0.0199	0.0146			-0.0198	0.0145
1999			-0.0069	0.0144			-0.0070	0.0144
2000			0.0120	0.0140			0.0120	0.0140
2001			-0.0072	0.0149			-0.0072	0.0149
Constant	0.0608***	0.0028	0.2987***	0.0707	0.3859	0.4649	0.3021***	0.0716
Obs.	5,346		5,320		5,320		5,320	
Number of People					1,205		1,205	
R-squared within					0.0157		0.0144	
R-squared between					0.0281		0.0715	
R-squared overall					0.0156		0.0294	
R-squared	0.0112		0.0294					
Prob > F	0.0000		0.0000		0.0000			
Prob > chi squared							0.0000	
F test that all $\eta_i = 0$					F(1,204, 4,100)=0.68			
$H_0: \sigma^2_{\eta_i} = 0$							chi-squared(1)=162.76	

*Notes: Standard errors are clustered by person. In the table * corresponds to 10%, ** to 5% and *** to 1% level of significance.*

Table 8: Random Effects Wage Growth Models, Controlling for Type of Job Mobility

	Model 1		***Model 2***		***Model 3***	
	Coef	*Std Error*	*Coef*	*Std Error*	*Coef*	*Std Error*
Job Change	0.1068***	0.0182				
Voluntary Job Change			0.1397***	0.0242		
Involuntary Job Change			0.0335	0.0380		
Other type of Job Change			0.0791*	0.0479		
First Job Change					0.1153***	0.0264
Second plus Job Change					0.0973***	0.0241
Obs.	5,346		5,346		5,346	
Number of People	1,206		1,206		1,206	
R-squared within	0.0095		0.0115		0.0093	
R-squared between	0.0142		0.0154		0.0156	
R-squared overall	0.0112		0.0134		0.0113	
Prob > chi squared	0.0000		0.0000		0.0000	

*Notes: * corresponds to 10%, ** to 5% and *** to 1% level of significance. Standard errors are clustered by person.*

Table 9: Effect of Job Mobility on Wage Growth Controlling for Misclassification

	First Stage Estimates			
	Misclassification Model		***Reference: Probit Estimates***	
	Marginal Effects	*P>|Z|*	*Marginal Effects*	*P>|Z|*
$\hat{\alpha}_0$	**0.0025**	**0.77**		
$\hat{\alpha}_1$	**0.5113**	**0.03**		
Experience	-0.0329	0.02	-0.0086	0.00
Experience squared	0.0005	0.05	0.0001	0.00
Female	-0.0289	0.16	-0.0112	0.17
Child	-0.0040	0.85	-0.0014	0.88
Living in a Couple	-0.0111	0.65	-0.0042	0.68
(Ref: Education – low)				
Education- medium	-0.0385	0.10	-0.0186	0.02
Education- high	-0.0505	0.21	-0.0214	0.07
Recent Training	0.1088	0.07	0.0389	0.00
Public Sector	-0.0692	0.02	-0.0312	0.00
Number of Employees > 50	-0.0457	0.06	-0.0188	0.01
Overskilled	0.0662	0.00	0.0297	0.00
Occupation of Origin:				
(Ref: Elementary Occ's)				
Manager	-0.0975	0.03	-0.0432	0.00
Professional	-0.0791	0.00	-0.0362	0.00
Clerk	-0.0807	0.05	-0.0337	0.00
Skilled	-0.0794	0.00	-0.0373	0.00
Sector of Origin:				
(Ref: Non Market Services)				
Agric. & Mining & Utilities	-0.0791	0.08	-0.0351	0.01
Manufacturing	-0.0559	0.14	-0.0234	0.09
Building	0.1536	0.09	0.0572	0.00
Market Services	0.0443	0.22	0.0188	0.14
Year Dummies:				
(Ref: 1995)				
1996	-0.0068	0.81	-0.0022	0.85
1997	0.0032	0.91	0.0028	0.83
1998	0.0458	0.34	0.0180	0.22
1999	0.0343	0.44	0.0166	0.31
2000	0.0587	0.30	0.0261	0.15
2001	0.0244	0.63	0.0100	0.59
Unemployment Rate	-0.0226	0.10	-0.0058	0.07

	Second Stage Estimates	
	Coeff	*Std Error*
Job Change	**0.1372**	**0.0532**
Number of Observations	5,217	
R-squared	0.0238	
Prob > F	0.0000	

Notes: Standard errors are clustered by person. The standard errors in the second stage are adjusted to take account of the fact that a generated regressor is included. The second stage regression includes the predicted probabilities from the first stage and the other variables from the pooled OLS model in Table 7.

Table 10: Sensitivity Analysis: Adjusted Pooled OLS Estimates for a Range of Misclassification Rates

	Estimate	*Standard Error*
Reference:		
Pooled OLS estimate (Table 7)	0.0794***	0.0182
Corrected OLS Estimates:		
Using α_0 =0.0025 and α_1 =0.5113 (from Table 9)	0.1008	
Varying α_1 (assume α_0 =0):		
α_1 =0.01	0.0795	
α_1 =0.05	0.0802	
α_1 =0.10	0.0811	
α_1 =0.20	0.0832	
α_1 =0.30	0.0860	
α_1 =0.40	0.0902	
α_1 =0.50	0.0968	
α_1 =0.60	0.1086	
α_1 =0.70	0.1364	
α_1 =0.80	0.2798	
Varying α_0 (assume α_1 =0):		
α_0 =0.01	0.0903	
α_0 =0.05	0.2131	

Table 11: Satisfaction with Various Aspects of Job

	% Satisfied	*% Not Satisfied*
Satisfied with:		
Earnings	67%	33%
Job Security	82%	18%
Type of Work	90%	10%
Number of Hours	83%	17%
Distance to Job/Commuting	88%	12%
Working Times (i.e. daytime, night-time, shifts etc.)	88%	12%
Working conditions/environment in place of work	88%	12%

Table 12: Satisfaction with Various Aspects of Job for those not Satisfied with Earnings

	% Satisfied	*% Not Satisfied*
Satisfied with other aspects of job, if not satisfied with earnings:		
Job Security	66%	34%
Type of Work	80%	20%
Number of Hours	70%	30%
Distance to Job/Commuting	83%	17%
Working Times (i.e. daytime, night-time, shifts etc.)	81%	19%
Working conditions/environment in place of work	78%	22%

Table 13: Second Stage IV Estimates of Job Mobility on Wage Growth^

	Estimate	*Standard Error*
Job Change	0.2590***	0.0706
Number of Observations	4,428	
R-squared	0.0275	
Prob > F	0.0000	

^ *Notes: * corresponds to 10%, ** to 5% and *** to 1% level of significance. Standard errors are clustered by person. The standard errors in the second stage are also adjusted to take account of the fact that a generated regressor is included in the model. The second stage model includes the same controls as the pooled OLS model in Table 7. The first stage model includes the same controls as the first stage model in Table 9 as well as the three instruments.*

Table 14: IV Estimates of Job Mobility on Wage Growth, Controlling for Misclassification^

	Estimate	*Standard Error*
First Stage Estimates		
$\hat{\alpha}_0$	0.0072**	0.0030
$\hat{\alpha}_1$	0.5107***	0.1337
Second Stage Estimates		
Job Change	0.1242***	0.0348
Number of Observations	4,428	
R-squared	0.0274	
Prob > F	0.0000	

^ *Notes: * corresponds to 10%, ** to 5% and *** to 1% level of significance. Standard errors are clustered by person. The standard errors in the second stage are also adjusted to take account of the fact that a generated regressor is included in the model. The second stage model includes the same controls as the pooled OLS model in Table 7. The first stage model includes the same controls as the first stage model in Table 9 as well as the three instruments.*

Table 15: Summary of Estimates of Job Mobility on Wage Growth

	Estimate	*Standard Error*
Pooled OLS (Table 7)	0.0794***	0.0182
Random Effects (Table 7)	0.0805***	0.0183
Controlling for Misclassification (Table 9)	0.1372**	0.0532
IV (Table 13)	0.2590***	0.0706
IV & Controlling for Misclassification (Table 14)	0.1242***	0.0348

Figure 1: Marginal Effect of Experience in Models of Job Mobility