2011

2011-05 Occupational Mobility, Occupation Distance and Specific Human Capital

Chris Robinson

Occupational Mobility, Occupation Distance and Specific Human Capital

by

Chris Robinson

Working Paper # 2011-5 **August 2011**

CIBC Working Paper Series

Department of Economics
Social Science Centre
The University of Western Ontario
London, Ontario, N6A 5C2
Canada

Occupational Mobility, Occupation Distance and Specific Human Capital *

Chris Robinson
Department of Economics, University of Western Ontario

August 2011

*The author wishes to thank participants at workshops and Yale and Michigan State Universities, Craig Riddell, discussant at the Canadian Economic Association Meetings in Quebec City, May 2010, and participants at the RCEA Conference in Rimini, Italy, June 2010 for comments on related preliminary working papers. Special thanks are also due to David Autor for making available a copy of the Treiman file with dual 1970 and 1980 occupation coding. This work is supported by the CIBC Human Capital and Productivity Centre at the University of Western Ontario.

Abstract

Measures of occupation distance based on underlying skill portfolios are constructed and used to contrast involuntary and total mobility. One component of total occupational mobility is voluntary mobility, including moves to higher job offers using the same skills, as well as promotions that may reflect augmented skills. These are not sources of specific human capital loss. By contrast, the involuntary mobility component due to plant closure involves a higher incidence of loss of specific capital. The evidence indicates that a decreasing fraction of occupation switches involve significant skill portfolio switches: the mean distance in involuntary occupational mobility declines significantly over time. Wage losses following displacement are strongly related to distance. This is reflected in a marked downward shift in the skill portfolio for involuntary occupation switchers. By contrast, the direction of the skill portfolio change in total mobility shows a neutral or modest upward pattern, suggesting limited specific human capital loss from voluntary occupational mobility.

JEL CODES J24,J31

1 Introduction

Occupational mobility has received increased attention in recent years. A major reason for this is that new research has linked human capital specificity with occupations. Some forms of job mobility, such as the movement up a career ladder, or the search theory job-to-job transitions following a better wage draw at a different firm, are typically considered to be a positive type of mobility, at least from the point of view of the worker. Several other forms of mobility come with costs. Losses associated with plant closings and more general worker displacement have received considerable attention, including the length of unemployment spells and the extent of wage losses.[1] Attention has recently focused on the loss of specific human capital associated with more general job mobility, especially occupational mobility, and its effect on wages.

Human capital specificity has been investigated in several recent papers. Neal (1995) and Parent (2000) both investigated evidence for industry specific human capital, and contrasted this with the original focus in the literature on firm specific capital. Kambourov and Manovskii (2009) argue that human capital is specific to three-digit occupation rather than industry. Poletaev and Robinson (2008) present evidence from the job based skill measures in the Dictionary of Occupational Titles (DOT) to support the hypothesis that many occupations are similar in their basic skills and that human capital is not narrowly specific to three digit occupation, but rather to these basic skills. Overall, the recent evidence places more importance on occupation or basic skill or task related human capital specificity, than on firm or industry specificity.[2] The extent of specific capital losses depends on both the type and magnitude of mobility. The magnitude has been studied in several recent papers, leading to a consensus that average annual rates of three digit occupational mobility in the United States over the decades of the seventies, eighties and nineties has been around 20%. To the extent that human capital is specific to three digit occupation, as argued in Kambourov and Manovskii (2008), this level of mobility suggests a potentially large annual loss of specific capital.

This paper examines the sources and potential magnitude of specific human capital losses from

[1]For recent examples, see Neal (1995), Farber (2005), Poletaev and Robinson (2008).

[2]The results in Poletaev and Robinson (2008), however, suggest some role for industry. In particular The evidence from the Displaced Worker Surveys suggests that while broader (*fluid*) specific capital may be transferred across a wide variety of industries and occupations, a more narrowly specific (*crystallized*) form of human capital may be lost when workers switch industries.

involuntary and total mobility using measures of occupation distance based on the skill vector characterization of occupations in Poletaev and Robinson (2008). The distance measures provide a means of assessing how much of the large amount of three-digit occupational mobility is mobility across jobs that use different basic skills (occupations that are far apart) and how much is mobility across quite similar jobs (occupations that are close.)[3] By the nature of occupation coding, occupational mobility is a discrete measure: all occupation moves are equal.[4] The skill vector characterization of occupations provides a means of assessing the direction of an occupation move. In general occupation codes contain no means of distinguishing, or examining the direction or magnitude of "upward" moves like promotions from "horizontal" or "downward" moves. In assessing potential loss of specific human capital from occupational mobility, it is important to be able to distinguish the direction of a move. Without this distinction, promotion to a different occupation code in which, for example, all the skills are used at a higher level, would entail a loss of specific capital no different from a change to different skills or a move to an occupation at lower levels of all the same skills.

The structure of the paper is as follows. Section 2 outlines the relation between occupation, skills and tasks in the literature and introduces the features of distance and direction in relation to occupational mobility which are studied in this paper. Section 3 details the construction of a vector characterization and distance measures at the level of 3-digit occupations in terms of a low dimension vector of factor scores based on the job characteristics in the DOT. The distribution of this distance measure is simulated for a random sample of workers randomly switched from an initial job into a new job, using U.S. occupation employment weights, providing a benchmark for comparison with the actual distribution of the distance measure for displaced workers from the Displaced Worker Surveys (DWS) and for total mobility from the annual March Current Population Surveys (MCPS). The distribution is simulated unconditionally, i.e. for all job moves, and conditional on an occupation switch. The distance measure is almost always positive for occupation switches, but by construction

[3]As noted by Moscarini and Vella (2003), a more accurate view of the level and trends in occupational mobility requires an occupation distance measure that can be used to weight moves across various occupations.

[4]The coding classifications do have some grouping levels, so that a distinction can be made between moves within groups and across groups. The previous literature has presented mobility estimates at 1-, 2- and 3-digit levels, focussing on the 3-digit level, but has not used this coding information as a distance measure to weight the 3-digit occupation moves. Moreover, the analysis in this paper using U.S. data, and in Robinson (2009), using data from the 2006 U.K. Skills Survey, shows that grouping occupations under the higher level occupational classifications is very far from the grouping that would be produced using a distance measure based on skill or task vectors.

the distance measure is always zero for occupation stayers. However, there is variation in the underlying DOT jobs within 3-digit occupation. To aid in the interpretation of how close or far apart particular occupation switches are, the distribution of the distance measure is also simulated at the level of the DOT job for random switches within 3-digit occupation.

In Section 4, data from the DWS 1984-2010 are used to study involuntary mobility. Distance distributions estimated from the DWS are compared with the benchmark distributions from random mobility. The sources of potential losses of specific human capital are decomposed into displacement rates, occupation switching rates conditional on displacement, and expected distance conditional on switching occupation. Displaced workers find their way back to much more similar jobs than random mobility, both unconditionally and conditional on switching occupation for both males and females. Moreover, there is a significant secular decline in distance, conditional on switching occupation, indicating that displaced workers are increasingly finding their way back to similar jobs. Section 5 compares distance distributions from involuntary mobility from the DWS with the distributions observed for total mobility as measured in the MCPS, and Section 6 extends the analysis to a comparison of the direction of the mobility in terms of the underlying skill levels. The results show a close similarity in distance, but a striking difference in direction: the direction for involuntary mobility is typically downward (and strongly associated with wage losses), while for total mobility it is neutral or upward. Section 7 provides some discussion and suggestions for further work.

2 Occupational Mobility: Distance and Direction

The literature on developing detailed measures of the "distance" between occupations, in the sense of how similar the skill sets used (or tasks performed) are, is relatively recent. This recent literature is related to the broader literature on human capital specificity and occupational mobility that goes back for many years. The empirical work in this broader literature is based on data sets that record standard occupation and industry codes and, particularly in panel data sets, firm tenure. Changing firm, industry or occupation meant some kind of change in the skill set used, or tasks performed, by the worker. However, the nature of this type of data makes it difficult to rank the changes. This literature, therefore, bypassed the issue at the level of skills or tasks. Instead, it focused on firm, industry and occupation tenure, as measured by the length of time a worker spent with a firm, industry or occupation. It assumes that firm, industry or occupation specific capital is an increasing function of tenure.[5] By contrast, the more recent literature focuses directly on skills and on ranking changes in these skills or tasks in a job in a systematic way, based on data that go beyond standard industry and occupation coding.[6]

2.1 Occupations, Skills and Tasks

Occupations are sometimes characterized by the skill sets or by the tasks performed, or both. The German Qualification and Career Survey (GQCS), used in Gathmann and Schonberg (2010), is an example of characterizing occupations by tasks, such as "cleaning" and "correct texts or data". The Dictionary of Occupational Titles (DOT) and its successor O*NET, used in most of the recent literature, is an example of a combinations of tasks, such as "dealing with people", and skills, such as "finger dexterity". There is no single interpretation of the relation between skills and tasks in the literature. Indeed, the terms skill and task are often used interchangeably. Heckman and Sedlacek (1985) specified a Roy model of comparative sector advantage in which a worker's skills generated certain sector specific task levels via sector specific task functions. Thus the basic concept is a vector of skills that are transferable in the sense that they can produce more than one sector specific task. Autor and Handel (2009) collect their own individual level task data and use a Roy model to analyze occupational choice based on occupation specific "returns to tasks" where workers skills are fixed, but tasks are chosen. Acemoglu and Autor (2011) emphasize the need for a clear distinction between skills and tasks in understanding the returns to skills and the evolution of earnings inequality.[7] "A *task* is a unit of work activity that produces output (goods and services). In contrast, a skill is a worker's endowment of capabilities for performing various tasks....The distinction ... becomes particularly relevant when workers of a given skill can perform a variety of tasks.."

[5] See, for example, Neal (1995), Parent (2000) and Kambourov and Manovskii (2009).

[6] This literature is rapidly expanding, covering a number of diverse topics including the distinction between tasks that are most likely affected by the rapid decline in computing costs (e.g. Autor, Levy and Murnane (2003)), or most vulnerable to off-shoring (e.g. Firpo, Fortin and Lemieux (2009)) as well as the evolution of tasks over a career (Yamaguchi (2010), Gathmann and Schonberg (2010))

[7] Acemoglu and Autor (2011) argue that what they call the canonical model for understanding skill prices has no meaningful role for "tasks", imposing a one-to-one mapping between skills and tasks which makes it tractable, but results in an inability of the model to explain recent empirical trends.

The analytic distinction made by Heckman and Sedlacek (1985) between skills and tasks is important for an occupation distance measure in the sense that it tries to locate the source of specificity. Suppose there is a small number of general skills, such as reasoning ability, mathematical skills, literary skills, finger dexterity, gross motor skills, etc. that can be used to carry out a much larger number of specific tasks that are actually carried out in the workplace in jobs that are located across many different coded occupations. If a worker with a given skill set can easily be switched between these specific tasks then the most useful occupational distance measure would focus on measuring the distance between the skill sets. However, if the skill set only provides the potential to carry out a wide range of tasks, and the specific tasks require what could be called a specific "crystallized" form of the skill set to carry out that particular specific task, then an appropriate occupation distance measure would also need to take into account this "crystallized" form of the skill set or task.

Poletaev and Robinson (2008) use a distinction between fluid and crystallized skills as a possible explanation for the role played by switching or staying in the same industry after displacement. The psychology literature on intelligence introduced a distinction between *fluid* and *crystallized* intelligence.[8] There is a very large literature that explores this distinction in various contexts. A recent application of this kind of distinction to life skills including workplace skills, in connection with the International Adult Literacy Survey, is Murray, Clermont and Binkley (2005). In terms of specific human capital, crystallized skills could be thought of as more narrowly specific human capital, skill or knowledge, compared to fluid skills, which would have more broader application. Consider, for example, a worker with a basic skill portfolio that is consistent with the skills necessary to be a good sales person. The basic skill portfolio associated with being good as a sales worker could be carried across industries. However, the worker may also have specific human capital in the form of crystallized skills or knowledge connected to the product, or to buyers of the product, that would be lost if they switched industries and were involved in selling a different product to different customers. In this case, if the worker switched basic skill - i.e. was no longer working in sales, it would not matter whether they also switched industry or not since the specific human capital in the form of

[8]See, for example, Cattell (1971) for early work on this distinction.

crystallized skills or knowledge connected to the product, or to buyers of the product, would be lost in any case since the worker was no longer in sales. On the other hand, if the worker remains in sales it would make a difference if the worker switched industry or not. In particular, the crystallized specific capital would be lost if the worker switched industry.

There may also be a similarly useful distinction between skills and tasks, related to the Heckman and Sedlacek task functions. Thus, the selling or teaching "tasks" may both use communication or literary skills. The skills may fluid, in the sense of being able to produce both tasks, but the tasks may be crystallized in the sense that to convert the skill from one task to the other may not be costless. There may be some specialized features of the task that are similar to industry specific knowledge in that the adaptation is not immediate.

The distance measure used in this paper is based on a large number of characteristics from the DOT that include a mixture of what might be considered tasks, task complexity and skills. The distance measure itself does not distinguish between these components, nor between fluid and crystalized skills. Thus, in terms of distance, it measures the similarity of occupations on this mixture. The vector characterization on which the measure is based does have a primarily skills interpretation, as it incorporates a large amount of "skill level" information. However, caution needs to be used in the interpretation of the distance measure as it relates, for example, to wage losses or mismatch since these can be affected by the distinction between skills and tasks.

2.2 Distance and Direction

The distance measures developed in the recent literature all have two basic components. First, a vector of tasks or skills is specified for each identified "occupation", i.e. occupation code or job code in a data set. This requires a source of data on skills or tasks that are associated with jobs held by workers that goes beyond the usual information gathered for standard occupational or industry coding. Second, a vector distance measure is chosen to measure the distance between occupations or jobs in terms of the distance between their underlying skill or task vectors. The alternative occupation distance measures in the current literature can be classified according to how they deal with each of these components.[9]

Information on underlying skills or tasks associated with occupations used to create measures of occupation distance can often be used to characterize occupational mobility in terms of direction as well as distance. The main requirement for this is that the underlying information have level rankings rather than simple incidence. The information on skills should have a ranking on the level of the skill, and information on tasks should have a ranking on the level at which the task is performed. A substantial part of the DOT has this feature, since it uses analysts to rate many tasks, such as dealing with people, according to levels of complexity, and skills, such as mathematical ability, according to levels of ability. By contrast, the employee survey based GQCS measures mainly incidence of tasks ("do you do this task in your job?"), rather than the level of the tasks, so that an increase in the task "score" for an occupation represents simply an increase in the fraction of employees in that occupation performing the task rather than the typical level at which the task is performed in that occupation.

Characterizing an occupation by a skill or task vector with this level information provides a basis for distinguishing between upwards, downwards or sideways movements in occupational mobility. A worker, following a plant closing, may move from an occupation where they performed managerial tasks at a high level, to one where they perform managerial tasks at a lower level (and, say, all other tasks or skills remained at similar levels to the previous occupation). This move may have the same occupational distance as that of another managerial worker who received a promotion involving the opposite switch in level of the managerial task. Obviously, the multidimensional nature of the skill or task vector complicates the directional characterization. This is discussed in more detail in Section 6.

3 Construction of the Occupation Distance Measure

Construction of the occupation distance measure has two steps. The first step is a vector characterization of each three digit occupation in terms of underlying information on tasks or skills for these occupations. The second step is the choice of a measure of distance between the vectors. The

[9]See Robinson (2009) for a more detailed review of this literature.

underlying information on tasks or skills for these occupations is taken from the revised 4th edition of the DOT. The raw data contain information on more than 50 characteristics of the jobs. Following Poletaev and Robinson (2008), the dimension is reduced by employing a factor analysis to the raw data. While the standard factor analysis does not produce unique factors, the euclidean distance between the vector of factor scores is invariant to "rotations" so that it is not necessary to take a stand on the precise nature of the "true" factors in an analysis of distance.[10]

3.1 Vector characterization of the occupation

The vector characterization constructed in this paper is similar to Poletaev and Robinson (2008).[11] The data source is the DOT. The DOT contains information on 12741 unique DOT occupations or jobs.[12] This master file contains two forms of data on the characteristics of the occupation. First is a measure of the complexity of the interaction with "data", "people" and "things". Second is a set of ratings on a large number of very detailed characteristics. One important set of characteristics come from the ratings on General Educational Development (GED) which is subdivided into three factors: reasoning development, mathematical development and language development. Each factor is then given a rating for each job, based on a detailed description of the rating.

GED attempts to measure the general education or life experience necessary to perform a given job in a satisfactory manner. By contrast, Specific Vocational Preparation (SVP) ratings measure the time required "to learn the techniques, develop the facility, and gain the knowledge for acceptable performance in a specific occupation." The remaining characteristics are divided into three groups: "Physical Demands and Environmental Conditions," "Temperaments," and "Aptitudes". The Physical Demands include a rating on the amount of strength needed, and an indicator of the presence of various requirements of the job, such as "climbing" or "stooping." Temperaments are defined as

[10]The literature has taken two approaches to constructing vectors of characteristics for each occupation. Gathmann and Schonberg (2010) take a direct approach, in the sense that the 19 task questions ("raw" characteristics) in their data set were used directly to obtain an occupation task vector by calculating the fraction of workers in each coded occupation that reported the task. The rest of the literature takes an indirect approach in which the raw characteristics obtained from the data were first processed via some form of factor analysis to obtain a set of estimated values for a reduced number of more basic characteristics.

[11]It is also similar to Ingram and Neumann (2006) who used factor analysis to extract a small number of basic factors ("skills") and examined the pricing over time of these "skills."

[12]This is version 4.3.

"personal traits" required by specific job-worker situations, such as "Performing effectively under stress." Finally, the Aptitudes factors use a 5 point scale to rate characteristics of the job such as "numerical ability", "form perception", "motor co-ordination", "finger dexterity", where the highest point on the scale is:

The top 10 percent of the population. This segment of the population possesses an extremely high degree of this aptitude

and the lowest point is:

The lowest 10 percent of the population. This segment of the population possesses a negligible degree of the aptitude

The DOT thus contains a very rich description of each DOT job in the form of ratings on a large number of characteristics of the job. A direct approach would simply characterize each DOT job by a vector of these characteristic ratings. However, this involves very large dimension "skill" vectors. The basic rationale for using a factor analysis is the assumption that jobs can in fact be distinguished on the basis of their requirements for (or use of) a relatively small number of skills or tasks ("factors"), such as fine motor skill, and that the relatively large number of characteristic ratings are reflections of these underlying skills. Formally, the model assumes that the L characteristic ratings for job j are generated by $K < L$ underlying skill factors according to the linear model:

$$\begin{aligned}
C_{1j} &= \mu_1 + \lambda_{11} f_{1j} + \lambda_{12} f_{2j} + ... + \lambda_{1K} f_{Kj} + \varepsilon_{1j} \\
C_{2j} &= \mu_2 + \lambda_{21} f_{1j} + \lambda_{22} f_{2j} + ... + \lambda_{1K} f_{Kj} + \varepsilon_{2j} \\
. &= \\
. &= \\
C_{Lj} &= \mu_L + \lambda_{L1} f_{1j} + \lambda_{L2} f_{2j} + ... + \lambda_{LK} f_{Kj} + \varepsilon_{Lj}
\end{aligned}$$

where C_{lj} is the rating for characteristic l on job j, f_{kj} is the amount of underlying skill k used in job

j and λ_{lk} is the factor (skill) loading of characteristic l on skill k. The scale of each factor, which is arbitrary, is usually set by imposing $\lambda_{11} = \lambda_{22} = ... = \lambda_{KK}$. The zero mean errors, ε_{lj}, are assumed to be uncorrelated with the factors, so that all the correlation among the characteristic ratings is explained by the common factors.

Each observation in the factor analysis is an $L \times 1$ vector of characteristic ratings for the observation, say job, j:

$$C_j = \mu + \Lambda f_j + \varepsilon_j$$

where C_j is an $L \times 1$ vector of the characteristics ratings for observation j, μ is an $L \times 1$ vector of means, f_j is a $K \times 1$ vector of unobserved skill levels (factor scores) for observation j, Λ is an $L \times K$ matrix of the factor loadings and ε_j is an $L \times 1$ vector of errors for observation j. Given that f and ε are uncorrelated:

$$cov(C) = E(C - \mu)(C - \mu)' = \Lambda \Sigma_f \Lambda' + \Sigma_\varepsilon$$

where Σ_f is the covariance matrix of the factors and Σ_ε is a diagonal matrix of the so called *uniqueness* variances.

In general, the separate elements of Λ and Σ_f are not identified. Further, the diagonal elements of Σ_ε and $\Lambda \Sigma_f \Lambda'$ are not separately identified. Identification is achieved in the standard factor analysis by normalizing the factors to be mean zero, with a standard deviation of one and by assuming that the factors are orthogonal, so that Σ_f is diagonal. Factors are estimated sequentially, according to how much of the observed covariance in the characteristic ratings can be explained by the factor. The "first" factor is estimated so that it explains the maximum amount of covariance in the characteristic ratings; the second factor is estimated so that it explains the maximum amount of residual covariance that was not explained by the first factor, and so on. The resulting estimated or predicted factors are linear functions of the characteristics.

A standard practice in factor analysis is to "rotate" the factors as a way to find more easily interpretable factors. However, if attention is focused on the distance between occupations, standard euclidean distance is invariant to rotation of the factors. There are alternative ways of identifying the factors; for example, restrictions are often placed on Λ. This is appropriate in cases where the nature of the underlying "true" factors is known and there is prior knowledge about which

DOT characteristics could be considered measures of which factors. This is the approach taken in Yamaguchi (2010), Ingram and Neumann (2006) and Autor, Levy and Murnane (2003).[13]

In principle, the vector characterization can be done at the level of the DOT job. However, with the exception of a special 1971 Current Population Survey (CPS) file, there are no data sets with employment at the level of the DOT job. In practice, therefore, it is necessary to obtain a vector characterization at the level of common occupation coding used in standard data sets. Poletaev and Robinson (2008) started with the DOT master file, averaged the values of each DOT characteristic over DOT jobs within a three digit occupation according to the master file crosswalk to obtain (mean) values of each characteristic for each three digit occupation and then ran a weighted factor analysis. The unit of observation for the factor analysis was a three digit occupation and the weights were employment weights of the occupations in the 1992 population of employees.[14] The simple averaging of values of each DOT characteristic over DOT jobs within a three digit occupation was done in the absence of any DOT level employment information in the DOT master file to calculate weighted means.

In general, employment at the DOT job level is unavailable since this is an analyst rather than employee survey based skill source. However, a special sample of the CPS from April 1971 was coded with both 1970 occupation codes and DOT characteristic values from the original 4th edition of the DOT. This provides the basis for alternative approaches to constructing the vector characterization that allows DOT level employment weights to be used in place of simple averaging over DOT jobs within a three digit occupation. In addition it permits separate skill vectors to be calculated for males and females for a given 3-digit occupation, reflecting potentially different employment distributions across DOT jobs for males and females within 3-digit occupations.

[13] Yamaguchi (2010) imposes a strong *a priori* structure by assuming first that occupations can be characterized by four underlying skills: cognitive skill, interpersonal skill, motor skill and physical skill and second that various pre-specified subsets of the 53 DOT characteristics are the relevant measures for each of these four characteristics. A principal component analysis, a form of factor analysis, was applied to produce the four skill indexes as the first principal components from each of the four subsets of the underlying DOT characteristics. Ingram and Neumann (2006) identify the subdivision of the last two of their four factors by imposing *a priori* restrictions on the factor loading matrix, especially as it relates to what they consider their "physical strength" factor. Autor, Levy and Murnane (2003) assume an *a priori* structure that groups occupations by four underlying task types and uses pre-specified subsets of the 53 DOT characteristics to identify the task types.

[14] For the purposes of interpretation of the factors, it is important to use employment weighting in the factor analysis so that the derived factors represent standard deviations of the factors for the employed population.

These alternative approaches were used in a modified version of the Poletaev and Robinson (2008) procedure to construct the vector of factor scores. The alternatives both make use of the special CPS April 1971 file with original 4th edition DOT level employment data. The preferred alternative ("weighted crosswalk approach") uses the employment weighting information at the level of the DOT job and applies this to the crosswalk files that assign census occupation codes to DOT jobs in order to calculated mean values of the DOT characteristic scores by each of the census occupation codes for the different coding schemes, 1970, 1980/1990 and 2000. The other alternative ("dual occupation coded files approach") uses the employment weighting information at the level of the DOT job in the special CPS April 1971 file, but also uses the 1970 occupation coding in the file to calculates the mean values of the DOT characteristic scores by the 1970 census occupation codes. In order to calculate means by later census occupation coding schemes (1980, 1990, 2000) it is necessary to use two dual occupation code files to go all the way from 1970 codes to 2000 codes. The first, the "Treiman file", is a special dual 1970/1980 census occupation coded file from a subsample of the 1970 census. The second is a dual 1990/2000 occupation coded file assembled from matching "basic" and "extract" monthly CPS files from 2000, 2001 and 2002.[15]

In addition, in order to deal with the both the change in characteristic scores for some DOT jobs, and the change in the actual DOT code for the same DOT "job", between the original 4th edition of the DOT of 1977 and the revised 4th edition of the DOT of 1991, the information in official hard copy documentation on changes between the original and revised 4th edition of the DOT had to be applied.[16] The weighted crosswalk approach is the preferred approach as it produces a high degree of consistency in the distribution of distances from random mobility across all the occupation coding schemes. By contrast, the dual occupation coded files approach produces clear breaks that appears to be due to averaging that occurs in using the dual occupation coded files that reduces the mean distance at each dual coding transition. Given the well known problems of error in occupation coding, especially independent occupation coding, a significant part of this averaging is probably

[15]Since the 1980 and 1990 coding schemes were almost identical, a modified 1990 coding scheme that combines a small number of codes has an exact counterpart for a modified 1980 coding scheme so a dual coded file to go across this coding "break" is not necessary.

[16]The dual occupation (1970 and 1980) coded census sample is known as the Treiman file. David Autor kindly supplied a copy of this data file for use in this analysis.

due to coding error. The full details of these approaches are relegated to the Appendix.

The factor that explains most of the variance (about 40 percent) appears to capture some kind of general intelligence; the second factor (about 20 percent) emphasizes fine motor skills, while a third factor (about 12 percent) is more related to physical strength. After these three, the remaining factors contribute relatively little. One additional factor appears to pick up visual skills - the factor loadings emphasize color discrimination, color vision, far acuity, field of vision.[17]. These factors are very similar to the original Poletaev and Robinson (2008) factors. The factor loadings from the factor analysis are almost the same.[18] However, the new factors, unlike those used in Poletaev and Robinson (2008), have different values for males and females within occupation. In principle, the factor analysis could be done separately for males and females and a potentially different set of factors for males and females derived. The assumption used in this paper is that males and females have the same underlying skills, but that the amounts may differ. This requires a single factor analysis to impose the same factors or skills, and allows different levels of the skills for males and females within the same occupation to be generated by different mean characteristic scores within 3-digit occupation due to different DOT level job weights for males and females within occupation.[19]

3.2 The Distance Measure for 3-Digit Occupations

The four-dimensional vector of factor scores constitutes a skill or task portfolio associated with each three-digit occupation. The distance measures used in this paper are alternative descriptions of the distance between these vectors. The starting point is standard euclidean distance. An important advantage of the standard euclidean distance measure, is the invariance to factor rotation, so that no *a priori* information need be imposed. In particular, it is not necessary to interpret the factors

[17]After four factors the eigen values fall below 2, a common cut off point for significant factors. The analysis was restricted to these four factors that explain 77 percent of the covariance in the characteristic ratings.

[18]They are also similar to those computed by Ingram and Neumann (2006)in their study of skill pricing.

[19]Since the occupational distributions for males and females have become more similar over time, with females entering previously exclusively male or male dominated occupations, the same may have happened at the DOT job level within occupation. Given that the different characteristic score means within occupations can only be calculated using DOT level employment information by sex from the special 1971 file, caution is needed in interpreting the female factor scores applied to more recent data since the true underlying characteristic score means for females may have moved closer to the score means for males.

as the "true" underlying skills or tasks.[20] The euclidean based distance measure to define distance between these occupations is defined as:

$$edist4w = \sqrt{\Delta_1^2 w_1 + \Delta_2^2 w_2 + \Delta_3^2 w_3 + \Delta_4^2 w_4}$$

where Δ_i is the difference for factor i and the w_i are weights that sum to one. The benchmark continuous measure used in this paper, *edist*4, is the equally weighted continuous measure, where the weights are all 0.25. In terms of the units of the underlying factors, this measure is equal to λ when there is a change of λ standard deviations in each of the factors. This measure is invariant to rotation. In Poletaev and Robinson (2008), a distinction was made between the "main" skill and less important skills in an occupation, based on the individual factor scores. To capture this idea with the euclidean distance measure the weights can be adjusted according to the importance of the factor in the occupation as reflected in the factor scores. However, this requires taking a stand on the nature of the "true" factors, and results in a distance measure that is no longer invariant to rotation, though for the range of weights used in this paper the measures are in fact quite insensitive to rotation.

In order to provide a benchmark for the distribution of distances observed in the empirical analysis of the DWS and MCPS data, the following "random move" experiment is performed. Assume a fixed distribution of occupied jobs (employees) by occupation at time t and $t+1$. Displace a person at random from a job in this distribution in t and assign the displaced person at random to another job in $t+1$. The expected distance for a displacement from occupation j in t is a weighted sum of the distance between occupation j and the occupation picked at random in $t+1$ where the weights are the employment shares of the occupations. The overall expected distance is a weighted sum of these expected distances from occupation j where where the weights are, again, the employment shares of the occupations.

Denote j as an occupation in t and j' in $t+1$. Define $d(j, j')$ as the distance between occupations j and j', and $\omega(j)$ as the fraction of jobs (employees) in occupation j. The overall expected distance

[20] By contrast, in analyzing the direction rather than simply the distance of the moves, the precise nature of the factors is important, especially as it relates to specific capital or wage losses.

is:

$$\sum_{j=1}^{J}[\sum_{j'=1}^{J} d(j,j')\omega(j')]\omega(j)$$

where J is the total number of occupations. The number of possible pairs, $d(j,j')$ is $J \times J$. In a population with equal employment in all occupations, the expected distance is the simple mean of the $J \times J$ distance terms $d(j,j')$. If employment is unequal across occupations, it is a weighted mean, where the weights are the product $\omega(j)\omega(j')$.

The primary analysis in this paper uses 493 modified 1990 occupation codes in total, which are consistent across all the DWS surveys from 1984-2002.[21] From the factor analysis, most of these (406) were assigned skill vectors separately for males and females so distances between the same pair of occupations may be different for males and females.[22] The frequency distribution of distances for a population of random movers of between each of the possible 493 occupations (including moving to the same occupation) using the actual employment weights from the 1992 sample used in the factor analysis are given in Figures 1a & 1b.[23] The frequency distribution of distances for males and females are very similar. Less than 1.5% of the random mobility results in staying in the same occupation and only 6.8% of the mobility for males and 7.3% for females involves distances less than 0.5. The expected values of the distance measure are 1.29 for males and 1.27 for females. The maximum distance is 3.46 for males and 3.40 for females.

The same distance measure was extended to the 2000 census occupation codes. However, due to the change in the occupations, there is a potential break in the series. The extension of the distance measure to 2000 codes was done using both the weighted crosswalk approach and dual occupation coded files from the CPS for 2000-2002. As noted above, the weighted crosswalk approach produces a distance distribution for random mobility that is constant across the different coding schemes. The shape of the distribution and the mean are both almost the same across the occupation coding break using the weighted crosswalk method.[24]

[21]See the Appendix for details.

[22]The remainder were either exclusively male (77) or exclusively female (10).

[23]The employment weights are for total employment in the occupation. Thus, the random experiment allows males and females to move with probabilities equal to total employment shares, not sex specific shares.

[24]By contrast, using the dual occupation coded files produces distributions of distance from random mobility that have very similar shapes to those for the 1990 occupations, but they are shifted to the left due to averaging that occurs in using the dual occupation coded files that reduces the mean distance at the dual coding transition. More details

3.3 Distance Within 3-Digit Occupation

The distance measure at the level of 3-digit occupations can be positive only for occupation switches. One benchmark in assessing the magnitude of the distance for these switches is the standard deviation units for the individual factors in the 1992 population. The measure *edist4* is constructed such that it is equal to λ when there is a change of λ standard deviations in each of the factors. An alternative is to compare the magnitude for a particular 3-digit occupation switch with the average distance of switches within 3-digit occupation at the DOT job level. Given 12741 DOT jobs in the revised 4th edition DOT, compared with about 500 3-digit occupations, there are about 25 DOT jobs per 3-digit occupation. Analysis of the crosswalk between these 12741 DOT jobs and their associated 1990 census occupation codes shows that not all DOT jobs within an occupation have the same values on the DOT characteristics. In fact, there is considerable variation.

The four factors estimated from the factor analysis and used to describe distance across occupations are simple linear functions of the DOT characteristics. The same linear functions may be used to estimate the factors at the level of the DOT job. The same distance measure may then be used to describe distance across DOT jobs. Table 1 describes the distribution of the distance measure *edist4* for random mobility across DOT jobs within 3-digit occupation. The revised 4th edition DOT has no employment data. Consequently, the results in Table 1 are for random mobility within 3-digit occupation unweighted by DOT employment within occupation. In addition, there is no distinction between males and females. The distribution was generated using occupation weights from the 1992 population.[25] The table also reports the results for distance conditional on switching occupation from the random mobility experiment across 3-digit occupations from the previous section for comparison. The distances for moves within occupations are, as expected, much smaller than those across occupations. Approximately 92% of random mobility within occupation has a distance below one compared to only 29% (31%) for random mobility for males (females) across occupations. However, as reported in Table 1, the mean (and median) distance is still quite large at about one half of that across occupations. Thus only occupation switches with a distance in excess of 0.7 are

are given in the Appendix.

[25]The results without occupation weights are very similar.

larger distance moves relative to a random within 3-digit occupation move.

4 Involuntary Mobility

Occupational mobility has a variety of different forms. A worker following a well defined job ladder within a firm may make a number of voluntary occupational switches in the form of promotions up the job ladder. Since they are progressions up a job ladder, the direction of these switches is, presumably, positive. The switches may vary in distance, with possibly large distances at early ages and shorter distances at later ages in a pattern that parallels concave age-earnings profiles. Alternatively, a worker may be involuntarily displaced from a job ladder by a plant closing. To the extent that workers of this type are unable to immediately find a similar job, they may experience an occupational switch which may be downward in direction and of varying distance. In contrast to the first case, since the direction is now typically downward, a larger distance implies a worse outcome. In this section data from the DWS are used for an analysis of involuntary mobility.

4.1 The Data for Involuntary Mobility

The DWS is a supplement questionnaire applied every two years since starting in 1984 to the monthly CPS. From 1984 to 1992 respondents are asked to provide information on job separations in the previous five years; from 1994 onward, the period is reduced to three years. The respondents are asked whether they had been displaced from a job, and why. The supplement is designed to focus on the loss of specific jobs that result from business decisions of firms unrelated to the performance of particular workers. Table 2 presents estimates of two types of "involuntary" job loss in the form of three year rates for displacement, separately for males and females. In Neal (1990) and Poletaev and Robinson (2008) attention was restricted to plant closures, which has always been the first listed reason for displacement. Table 2 presents estimates for this type of exogenous involuntary job loss, as well as for displacement according to the broader BLS definition, which includes "slack work" and "position or shift abolished". Figure 2 plots the three year displacement rates together with the unemployment rate.

There are several problems of comparability over time discussed in detail in Farber (2004). The respondents in the DWS are asked if they were displaced from a job in the last three years (surveys 1984 - 1992) or five years (surveys 1994-2000). One way of dealing with this is to restrict observations to displacements in the last three years for all survey years. However, as Farber (2004) notes, there remains a comparability issue due to the fact that respondents in the 1984-1992 surveys were asked to report the displacement from the longest job over the last five years. In the case, for example, where there was a displacement three years ago and another five years ago, dropping this observation would undercount displacements in the last three years for the earlier surveys. Based on evidence from the Panel Study of Income Dynamics, Farber (2004) suggests an upward adjustment of about 11% in the estimated rates for these surveys.[26]

The pattern in Figure 2 is the same for both types of displacement. It shows a decline from job losses in 1981-83 up to the 1990 survey representing displacements in 1987-1989. The job loss rate jumps up for the 1992 survey, representing displacements in 1989-91, and thus affected by the onset of the recession in 1990. The next survey covers the period 1991-93 and already reflects the recovery as job loss rates begin a long decline to the recession in the early 2000s. For the recession in the early 2000s, job loss rates rise for both the 1999-2001 and 2001-03 periods before a brief decline for the 2003-05 and 2005-07 periods.[27] This brief decline is followed by a large increase for the 2007-09 period reflecting the steep recession in 2008. Using the BLS measure, the rate reaches a historic high in 2007-09 for both males and females. By contrast, the rate of plant closings are similar to rates in earlier recessions and are lower in 2007-07 than in 1981-83.[28] Thus the time path of involuntary job loss has a cyclical pattern, similar to that noted in Farber (2005) for the period up to the 2004 survey, though no clear secular pattern.[29]

[26]See Farber (2004) for more details and a full discussion of the comparability issues.

[27]A minor exception is the BLS rate for the males which only declines in 2003-05.

[28]Farber (2004) uses a broader definition than the BLS, and reports the patterns for different age and education groups. The same basic pattern in Figure 1 is seen in Figures 1-3 in Farber (2005) across age and education groups. (There is a minor difference between the 1992 and 1996 surveys where the Figures in Farber (2005) show a small upward movement between 1994 and 1996.)

[29]Farber (2005) argued that there was no clear evidence of a secular pattern up to the 2004 survey, though noted that while job loss rates among older and more educated workers did decline after 1995, they remained higher up to the 2004 survey than they were at the peak of the 1980s expansion, possibly due to restructuring.

4.2 Distance Distributions for Involuntary Mobility

The results in Poletaev and Robinson (2008) suggest that the overall cost of mobility among occupation switchers following displacement is highly sensitive to whether or not the worker made a significant switch in their skill portfolio, defined using a discrete distance measure. Poletaev and Robinson (2008) constructed several discrete distance measures by defining a switch in the skill or task portfolio based on the difference between the vectors. The simplest uses only the order information: does the occupation switch involve a switch in the main skill as defined by the factor (skill) in the portfolio with the highest score? The other measures use both level and continuous distance information at the level of the individual factors to construct alternative switch definitions.[30] These definitions of skill portfolio switching provide a coarse ranking of the distance between occupations for any move. By construction, all three-digit occupation stayers are skill stayers. However, three digit occupation switchers may switch between occupations that are relatively close (no skill portfolio switch) or between those that are far apart (skill portfolio switch). However, these measures are not invariant to rotation.[31] The analysis in this paper uses a continuous euclidean distance measure that is both invariant to rotation, and permits a much more detailed description of the distribution of distances among displaced workers than the discrete skill portfolio switching measure.

The potential loss of specific capital from involuntary mobility depends on the actual distribution of the distances of the involuntary moves following displacement. The actual distribution of distances for mobility following displacement is first compared with the distribution of the distance measure from random mobility described in Section 3. Let d denote the distance of a move. Figure 3 compares the cumulative distribution of d from random mobility with the actual distribution of observed distances, d_i, for males age 20-64 displaced from the private sector by plant closing from the pooled DWS sample for 1984-2002. Both use the same modified 1990 occupation coding to construct the

[30] Skill Portfolio Change 1 (PC1) denoted a skill change when the order information indicates a switch in the skill portfolio and this is confirmed by the distance and level information: the change in the score of the original main skill cannot be too small and the level cannot be too low. Skill Portfolio Change 2 (PC2) allowed for the possibility that some of those classified as stayers by the order information should really be classified as switchers on distance grounds. Thus, the original classification to stayer on the basis of keeping the same main skill is reversed if there is a large change in the main skill factor score.

[31] While not invariant to rotation, the main results in Poletaev and Robinson (2008) were robust to a variety of common forms of rotation.

distance.[32] The observed distances, d_i, from the DWS are constructed using a comparison of the occupation codes in the pre- and post-displacement jobs. However, there is a well known problem of measurement error in occupation coding in most data sets, including CPS data sets such as the DWS. These potential coding errors have been extensively discussed in connection with accurate measures of occupational mobility, but they also have implications for distance measures.

The normal occupation coding procedure uses coders to assign particular occupation codes from information provided by survey respondents. The information usually consists of a job title and a brief description of the job. There is a great deal of evidence of substantial measurement error in assigning these codes. The problem is particularly severe in identifying true occupation switches when occupations held at two different points in time are coded independently. In this case, the coding for the occupation in the later period is done independently of any information regarding the occupation in the earlier period.[33] The occupation coding for the DWS is independent coding. There are two types of error from independent coding. One is the simple error of indicating a switch when this is not true. This converts a number of true $d_i = 0$ into $d_i \geq 0$. This could be simply random, or it could be that they tend to be "close" occupations and have values close to zero. The other is that a switch may be true, but there was error in assigning one or both the independent codes compared to the true, but different codes. Suppose we can write the coding error in terms of the underlying factor scores as:

$$\hat{f} = f + \epsilon$$

where the factor score assigned on the basis of the actually coded occupation, $\hat{f}$, is the sum of the factor score for the true occupation f, and a classical measurement error, ϵ, with mean zero and variance, σ_ϵ^2. The terms in the distance measure based on the coded occupations can then be written:

$$(\hat{f}_1 - \hat{f}_2)^2 = [(f_1 - f_2) + (\epsilon_1 - \epsilon_2)]^2$$

Under independent coding, the covariance in the errors is zero, and the calculated distance measure

[32] The details of the sample are described in the Appendix.

[33] See Kambourov and Manovskii (2008, 2010) and Moscarini and Thomsson (2007) for a full discussion, including the distinction between dependent and independent coding and further references.

will overestimate the true distance measure by an amount determined by the variance of the error terms. That is, the expected value of the terms in the distance measure are:

$$E(\hat{f}_1 - \hat{f}_2)^2 = E(f_1 - f_2)^2 + 2\sigma_\epsilon^2$$

If the coding error tends to involve "close" occupations, then the variance, σ_ϵ^2, will be relatively small.

Figure 3 shows that less than 1.5% stay in the same occupation under random mobility, but about 40% stay in the same occupation among males 20-64 displaced from the private sector by plant closing, based on the independent occupation coding in the DWS. Only 29% of the mobility from the random experiment benchmark have $d < 1$, compared to 68% of the mobility from plant displacement. Since the coding error tends to impart an upward bias to the estimated distance, this is a lower bound on the fraction in the DWS with true $d < 1$. A large part of the difference from random mobility is due to the difference in the fraction coded as switching occupation, i.e. with $d = 0$. However, in addition, conditional on switching occupation, the distances are smaller in the DWS data compared with random mobility as shown in Figure 4.

Figure 4 compares the cumulative distribution functions for distance conditional on switching occupation in the DWS sample and the random mobility benchmark for both males and females. Looking at the distance conditional on switching occupation, for the random mobility benchmark, for males only 29% have $d < 1$ while 47% of males displaced from a plant closure in the DWS have $d_i < 1$. For females the difference is even more pronounced: 31% compared to 59%. Again, the measurement error in the independent coding implies that the distance estimates for the DWS are upward biased, so displaced workers find their way back to much more similar jobs than random mobility, both unconditionally and conditional on switching occupation for both males and females. In terms of magnitudes, it is not possible to estimate the mean distance for DOT job switchers who stay in the same occupation in the DWS, however, the estimated mean for random DOT job switchers within occupation (Table 1) is 0.7. For displaced workers that switch occupations, 29% of males and one third of females have distances less than the within occupation mean distance of 0.7.

The amount of occupational switching following displacement is much lower than random mobility, but is nevertheless very high. Tables 3a & 3b show a large amount of occupational switching in the DWS 1984-2010 for displacement from the private sector by plant closings, or for the broader BLS definition. There is an occupation coding break after the 2002 survey, when the DWS switched to the 2000 census codes for occupation. There is only a small increase in the number of occupations from 489 modified 1990 codes to 502 codes in the 2000 census coding. A small increase in measured 3 digit occupational mobility would be expected if the 2000 codes were mainly adding codes by splitting, say, some single miscellaneous codes of the previous scheme as the components of the earlier miscellaneous group increased in employment in the form of growth in "new" occupations. However, the relationship between the coding schemes is more complicated, and it is possible that it is tending to group together what were frequent switches across what were previously considered different occupations, but what are now considered the same occupation. Other things the same, this would reduce measured switching under the new codes. In addition, the principles behind the aggregation into minor and major groups changed, making it especially difficult for consistent comparison across the coding break at this level. Tables 3a & 3b thus only report the 3-digit codes after the break.

Occupational mobility is around 60% at the 3 digit level. It is only a little lower at the more aggregate occupation groupings: even at the major group level with only 13 groups about 45% switch occupation. This appears to be a large amount of switching. However, this covers periods between pre- and post-displacement jobs as long as three years. The general consensus from the earlier literature is that annual 3-digit occupational mobility is around 20%, so that a relatively large figure may be expected for the longer interval. Moreover, the estimates in Tables 3a & 3b are subject to the upward bias from the independent coding. Previous studies suggest that the upward bias in the estimates of occupation switching can be large, though the bias is lower when attention is restricted to workers who switch employers, as is the case for the DWS. Kambourov and Manovskii (2008) compare original independent coding with retrospective dependent coding available in the Panel Study of Income Dynamics (PSID) for the period 1968-1980 to estimate adjustment factors for occupational mobility based on independent coding after 1980. The time path of the adjustment factors used to deal with this bias in Kambourov and Manovskii (2008) suggest that it is stable over time, so that the interpretation of trends in Tables 3a & 3b is subject to fewer problems than the interpretation of levels.

Figure 5 shows the change over time between the cumulative distribution of distance for displaced males who switched occupation in the 1984 survey compared to the 2000 survey. The cumulative distribution for 2000 is well to the left of the distribution for 1984, such that while 41% of the occupation switchers in 1984 have $d_i < 1$, this increases to 53% in 2000. Table 3a shows that a higher fraction of male workers displaced by a plant closure switched occupation (64% vs. 59%), which suggests increased "mobility", but this is more than offset in terms of the mean distance by the reduction in the distance of the conditional moves described in Figure 5. The overall effect can be seen in Figure 6 which compares the unconditional distributions for males for 1984 and 2000. There is a smaller mass at zero distance for 2000 due to the higher fraction of occupation switching; however, the lower conditional distance for 2000 results in the cumulative distribution functions crossing fairly rapidly; by the 55th percentile the 2000 cdf has moved to the left of the 1984 cdf. In other words, the increased occupation switching is offset by the smaller conditional distances, such that overall workers displaced by a plant closure are on average finding new jobs that are closer to their old jobs than they could before. The unconditional mean for 2000 is 0.646 compared to 0.687 in 1984.

4.3 Expected Distance for Displaced Workers: 1984-2002

Summary parameters of interest that relate to the potential loss of specific human capital through involuntary mobility are the expected distance between pre- and post-displacement jobs, and the fraction displaced. Given that the distance after displacement for a worker staying in the same occupation is zero, in terms of the continuous distance measures defined in Section 3, the expected distance for a displaced worker ($E(d_i|\delta_i = 1)$) is the product of the expected distance, conditional on switching occupation ($E[d_i|\alpha_i = 1]$) and the probability of switching occupation, conditional on being displaced $Pr(\alpha_i = 1|\delta_i = 1)$:

$$E(d_i|\delta_i = 1) = E[d_i|\alpha_i = 1]Pr(\alpha_i = 1|\delta_i = 1)$$

where $\delta_i = 1$ if worker i is displaced, and zero otherwise; and $\alpha_i = 1$ if worker i switches occupation, and zero otherwise.

Estimates of $E[d_i|\alpha_i = 1]$ based on workers displaced by plant closure in the DWS data for 1984-2002 are presented in Tables 4a & 4b. The dependent variable in Table 4a is the unweighted euclidean distance, *edist4*. This is replaced in Table 4b by the weighted euclidean distance, *edist4w*, with larger weight on the more important factors in the pre-displacement job.[34] The results are very similar. There is strong evidence of a decline over time in the expected distance following displacement, conditional on switching occupation. The first two columns report estimates for males and females together; the last four columns report the estimates separately for males and females. The mean distance over the period is 1.0290 using *edist4*, and 1.0403 using *edist4w*; The results in column 1 show a decline over time of 0.1317 for *edist4*, and 0.1434 for *edist4w*. The decline occurs mainly in the mid 1980s to mid 1990s. It is a little smaller for females compared to males (columns 3 & 5), but on average for females the average distance is significantly lower: 0.9405 for females vs. 1.1022 for males using *edist4* and 0.9355 vs. 1.1271 using *edist4w*, so that the percentage declines are very similar. The decline reflects the shift in the conditional cumulative distribution function in Figure 5.

The remaining columns in Tables 4a & 4b use a quadratic time trend in place of the time dummies and include the unemployment rate, age and education. There is little evidence of a cyclical effect on the conditional distance using the unemployment rate. In all specifications that include either a linear or quadratic time trend, the unemployment rate is insignificant. The results for females including a quadratic time trend and a quadratic in unemployment show insignificant results for both; dropping unemployment results in statistically significant results for the time trend, similar in magnitude to that for males. Column 2 shows the significantly lower distance for females. It also shows significantly lower distance for older workers and for male college graduates. There is an interesting difference in the results for college graduates between males and females. The very large and significant effect in the pooled results (-0.1832) is primarily due to the large and significant effect for males (-0.2815).

All the qualitative patterns are the same for the larger sample (20929 vs. 9016) of all workers displaced in the DWS 1984-2002 according to the BLS definition, except that there is now significant

[34] The equal weights of 0.25 are replaced by declining weights, 0.4, 0.3, 0.2 and 0.1. Full details are given in the Appendix.

evidence of cyclical effects. Table 5 shows the main results for the pooled sample for both un-weighted and weighted distance measures. The magnitudes are all similar to the results in Tables 4a & 4b. The unemployment rates ranged from 4.3% to 9.0% over this period. The estimated effect using the quadratic is positive up to 6.9%. The effect at the mean is around 0.013, so while the estimates are statistically significant, the magnitudes are relatively small.

The shift in the unconditional cumulative distribution functions in Figure 6 reflects the time path of the average distance for displaced male workers. Let $\alpha(t)$ denote the fraction of displaced workers that switch occupation in period t, and $D_c(t)$ be the average distance of the involuntary move of the displaced worker conditional on switching occupation. The average distance for a displaced worker is then simply $D(t) = \alpha(t)D_c(t)$. Finally, let $\delta(t)$ denote the fraction of workers displaced in period t. The overall cost, per worker, of involuntary mobility is then $\delta(t)D(t) = \delta(t)\alpha(t)D_c(t)$. Table 6 reports estimates of $\delta(t)D(t) = \delta(t)\alpha(t)D_c(t)$, putting together its components from Table 2 ($\delta(t)$, Tables 3a & 3b ($\alpha(t)$ and Tables 4a & 4b ($D_c(t)$)[35]

The results in Table 6 show a decline in the overall cost per worker, $\delta(t)D(t)$, for both males and females from all types of displacement as defined by the BLS, and from the subset of displacements following plant closure for the displacement periods from 1981-83 to 1999-2001. For males, at the end of the period the costs per worker from displacements from plant closings are only 74.9% of what they were at the beginning of the period and from all displacements they are only 77.5%. For females there is a similar reduction from plant closing (73.3%) but a smaller reduction from all displacements (86.0%). The components in the table show that in all cases there is no contribution from the time path of occupation switching. This is either the same at the beginning and end of the period, or is slightly increased. Both the decline in the rate of displacements and the decline in mean occupation distance are important sources of the overall decline in costs per worker. The decline in mean occupation distance is a substantial contributor in all cases. Moreover, even using the rate of major group occupation switching, there is no indication of this strong decline in mean distance for except for females from plant closings. For displacements for males, the rates of major group switching are almost unchanged at 44.6% at the end of the period compared to 45.5% at the

[35] The reported results in Tables 4a & 4b used pairs of adjacent surveys which generate five year displacement periods. The results in Table 6 report the individual survey years.

beginning from plant closings, and 43.6% compared to 44.0% from all BLS defined displacements. For females for plant closings there is a decline from 51.3% to 46.2%, but for all displacements there is a much smaller decline of 48.2% compared to 46.4%.

4.4 Effects of the 2008 Recession

Table 2 showed an effect of the recent 2008 recession, especially for males, in increasing the rate of displacement. There is a break in the occupation coding after the 2002 DWS that complicates the comparison over time for the expected distance for displaced workers, conditional on switching occupation. However, analysis of the distance distributions from random mobility using the modified 1990 occupation codes and the 2000 codes shows that when the weighted crosswalk method is used the distributions are almost identical. A detailed analysis is provided in the appendix. The expected distance for males is 1.29 for modified 1990 codes and 1.26 for 2000 codes; for females it is 1.27 for modified 1990 codes and 1.25 for 2000 codes. Thus, using the weighted crosswalk method the break in the distance measure is minimal. Figure 7a plots the mean distance for males and displaced by plant closing or using the BLS definition; Figure 7b repeats the plot for female displaced workers. There is evidence of a cyclical effect for the recent recession; relative to the trough (2004 survey for males and 2006-2008 surveys for females, the distance increases for 2010. However, at least up to the time of the 2010 survey, the effects are relatively modest. Apart from the spike in the estimate for females from plant closure, which is a relatively small group, the estimated distance for the 2010 survey remains low relative to the 1990s.

5 Comparison with Total Mobility

Potential loss of specific human capital is not confined to involuntary mobility as reflected in the DWS. In documenting an increase in occupational mobility in the United States, especially over the period 1969-1985, Kambourov and Manovskii (2008) remark: "This rise in mobility may imply a substantial increase in the destruction rate of human capital."[36] There is an emerging consensus on the time path of occupational mobility over the last four decades. Kambourov and Manovskii

[36]Kambourov and Manovskii (2008), p.56.

(2008) present evidence from the Panel Study of Income Dynamics (PSID) of increased occupational mobility from 1968-1997, with the main increase occurring in the 1970s. Moscarini and Vella (2003), using the March Current Population Survey (MCPS), and Moscarini and Thomsson (2007), using matched monthly basic Current Population Surveys (CPS) and sophisticated filtering to mitigate the effect of occupation coding error, document a pro-cyclical pattern of occupational mobility and a secular pattern similar to Kambourov and Manovskii (2008) for the overlapping period in the studies of 1979-1997. Moscarini and Thomsson (2007), estimate monthly occupational mobility of about 3.5% that they argue is consistent with the annual level of about 20% estimated in Kambourov and Manovskii (2008). In all the studies, the pattern of increase in mobility is much more muted after the 1980s, and Moscarini and Thomsson (2007) document a strongly declining trend from the mid 1990s, resulting in mobility rates falling below 1970s levels.[37].

In order to get a better understanding of the importance of these patterns in terms of specific human capital losses, it is necessary to distinguish between different forms of occupational mobility, both in terms of the distance between the occupations and the direction of the mobility. In this section, the characteristics of the distribution of distances for total mobility are contrasted with the patterns detailed in Section 4 for involuntary mobility.

5.1 The Data for Total Mobility

The data for total mobility, covering the same period as the DWS data Section 4 and using the same modified 1990 coding, come from the MCPS. The MCPS was used in a study of occupational mobility in Moscarini and Vella (2003), comparing the occupation code of the current job in the MCPS with the occupation code of the longest job held last year. As Moscarini and Vella (2003) note, the major advantage of these data are that they use a form of dependent coding of the occupation in the longest job last year and the current job, which they argue "eliminates virtually all of measurement error of one type (attributing different occupations to the same job)." Unlike the monthly CPS, which did not adopt dependent coding until 1994, the MCPS has used dependent coding since 1970. The MCPS used 1980 or 1990 coding throughout the survey years 1983-2002 (earnings years 1982-2001)

[37]The time path also shows evidence of pro-cyclicality. See Moscarini and Thomsson (2007)

so that the same modified 1990 coding as used for the DWS analysis in Section 4 may be applied to data for a similar period in the MCPS. The full details of the MCPS sample used in the analysis are given in the Appendix.

The dependent coding procedure in the MCPS in principle allows the identification of a change of "job", since in the supplement the interviewer checks whether the longest job last year is the same as the current job, and if it is, the interviewer marks a check item, and the longest job last year is given the same industry and occupation code as the current job. If not, the industry and occupation codes for the longest job last year are collected independently. The job change can involve a change within employer. Unfortunately, this check item is not available on the March tapes so a change of job is not directly identified.[38] The observation of the original 3-digit industry and occupation codes in the MCPS allows the partition of observations into: (A) occupation and industry switchers; (B) occupation switchers and industry stayers; (C) occupation stayers and industry switchers; and (K) occupation and industry stayers. This results in an incompletely observed partition into job changers: $A + B + C + \theta K$; and job stayers $(1 - \theta)K$, where θ is the (unknown) fraction of occupation and industry stayers that changed jobs. In order to estimate job changing rates or the unconditional (on switching occupation) distribution of distances for all job changers it is necessary to have an estimate of θ.

Kambourov and Manovskii (2010), based on evidence from the retrospective PSID indicate that less than one half of individuals switching employers also switch occupation. They also suggest that as many as 50% of occupation switchers based on raw coding $(A + B)$ may not be true switchers. Assuming 50% were not true switchers, and that 25% of job changers switch occupation, the implied value of θ from the observed values of $A + B$, C, and K in the MCPS for 1983-2002, is 0.0767. However, past research suggests that a majority of the false occupation switches occur when the individual does not switch employer; since all the measured occupation switches in (A+B) were assigned to independent coding, and since most of this occurs when there is an employer switch, the 50% number for observed occupation switchers that are actually stayers is probably too high. Assuming that 25% are actually stayers implies that θ equals 0.1683. The estimated percentage of

[38]See Stewart (2002).

workers switching jobs is about 18% using a value for θ equals 0.1 and about 25% using = 0.17. Stewart (2002), using more detailed information on the job history, reports job mobility rates around 23% for the same period, suggesting that θ closer to 0.17 may be preferred. In either case, the job mobility rates are stable or slightly declining over the period, which is also the case in Stewart (2002).

The dependent coding in the MCPS results in different measurement error issues for the distance measure in the MCPS compared to the DWS in which all workers change jobs and there is no dependent coding. The dependent coding implies that $(1-\theta)K$ are correctly assigned a zero distance, while $(\theta K + C)$ are assigned a zero distance on the basis of independent coding. Some of these could actually be switches that were, by coincidence, independently assigned the same code. This is presumably a small number. Finally, $(A + B)$ are assigned a positive distance. Some of these will not be true occupation switches even though they switched jobs, perhaps 25% using the earlier estimate; the remainder are true switches, but the estimated magnitude of the true distance will be subject to an upward bias due to measurement error.

5.2 Distance Distributions for Total Mobility

Figure 8 compares the cumulative distribution of d from random mobility with the distribution of d_i for males age 20-64 that switch jobs, assuming two alternative values of θ, 0.17 and 0.10. While less than 1.5% stay in the same occupation under random mobility, 68% of job changers stay in the same occupation in the MCPS, assuming $\theta = 0.17$ and 57% assuming $\theta = 0.10$. The results for females are very similar (66% stay in the same occupation for $\theta = 0.17$, and 54% for $\theta = 0.10$). As with the involuntary mobility in the DWS, while a large part of the difference from random mobility is due to the difference in the fraction coded as switching occupation ($d = 0$), there is also a significant difference due to smaller distances conditional on switching occupation. The estimated fraction of job switchers with $d = 0$, and hence the unconditional cumulative distribution of distances, is sensitive to the estimate of θ. However, the cumulative distribution of distance conditional on switching occupation does not require an estimate of θ. Figure 9 compares the conditional cumulative distribution functions in the MCPS and the random mobility benchmark for both males and females. For males under random mobility only 29% have $d < 1$ while 45% of male job switchers have $d_i < 1$. For females the difference is more pronounced: 30% compared to almost 58%. As with the DWS distance estimates based on independent coding, the distance estimates in the MCPS for job switchers are likely to be upward biased since occupations for job switchers are independently coded.

Table 7 reports estimates of the expected (conditional) distance, $E[d_i|\alpha_i = 1]$, for total mobility. The patterns for education and age, and the differences between males and females are similar to the patterns in involuntary mobility in Tables 3a, 3b & 5. The smaller distances for older college graduates are almost exclusively due to males. The specifications using time trends and unemployment rates produce mixed results. Unlike the results for involuntary mobility, there is no evidence of a strong downward trend. Total mobility combines involuntary with voluntary forms of mobility such as promotions, moves to employers with higher wages for the same job, job experimentation or learning, etc. The contrast with the DWS suggests that displaced workers are finding their way back to increasingly similar jobs, but promotions, job experimentation, etc. are occurring at similar rates across the period.

Table 8 reports basic summary results from the two types of mobility. Comparison of the magnitudes in the distance results for total and involuntary mobility is complicated because of the different data sources and time interval for the mobility.[39] The fraction of all job changers (voluntary and involuntary) that are observed in the same occupation in their current job and their longest job last year in the MCPS is smaller than the fraction of displaced workers that are observed in the same occupation in their pre- and post-displacement jobs in the DWS, especially so for females. The gap is sensitive to the estimated θ, but is generally large. The smallest is the increase about 40% in the fraction staying in the same occupation in the MCPS compared to the DWS for males using the estimate of $\theta = 0.1$; the largest is the 80% increase for females using the estimate of $\theta = 0.17$. Similarly, there is a larger fraction with an unconditional (on switching occupation) distance less than one in the MCPS (for example, 82% vs. 68% for males with $\theta = 0.17$), and a much smaller (unconditional) mean distance (for example, 36% vs. 65% for males with $\theta = 0.17$).

There are conflicting factors that can influence the difference in occupation switching rates be-

[39] Given the large sample sizes, the standard errors are very small so that the differences are almost all significant.

tween involuntary and total mobility. For the DWS, assuming this is random displacement, uncorrelated with the match, to the extent that the typical worker was in a good match prior to displacement, they will try and find a similar job after displacement, and hence observed occupation switching will be relatively low. By contrast, the mobility in the MCPS includes workers leaving a job voluntarily in search of a better job match, as well as promotions, both of which lead to occupation switching. However, it also includes workers in search of higher pay for the same job, which leads to relatively low occupation switching rates. The comparison is also complicated by measurement error issues. In the DWS, all cases of job changes (displacements) are independently coded. The evidence in the previous literature strongly suggests that this will bias upward the estimated fraction of occupation switchers, though since the cases in the DWS are known job changes, this literature also suggests that the bias will not be as large as in independent coding at two points in time where there may not have been a job or employer change.

In contrast to the unconditional distributions, the distance distributions conditional on switching occupation in the MCPS and DWS reported in the lower half of Table 8 are very similar. To the extent that independent coding tends to assign "close" occupations to what should have been the same occupation for the false occupation switches, the upward bias in occupation switching will be accompanied by a downward bias in the conditional (on an occupation switch) distance. This applies to both DWS and MCPS since all cases of different occupation codes for the current job and the longest job last year in the MCPS are cases to which independent coding must have been applied. However, while the conditional distance distributions are very similar, the direction of mobility, as measured by the direction of change in the individual factor scores in the skill vector, is very different.

6 The Direction of Mobility

The analysis in Poletaev and Robinson (2008) showed that among the displaced workers that switched three digit occupation, large wage losses were experienced only by those that significantly switched their skill portfolio. In fact, provided the worker did not make a significant switch in the skill portfolio, wage losses were relatively insensitive to switching occupation. Almost identical results are obtained using the new factors in this paper to construct the same definitions of significant portfolio changes. Further analysis (Poletaev and Robinson (2008), Table 7) suggested that the large wage losses associated with significant skill portfolio switches following displacement were due to the fact that these changes reflected not just skill switching, but also a typically downward direction move in the skill portfolio. In this section, direction changes in the individual factor scores in the skill portfolio for displaced workers that switch occupation are compared with the changes for occupational switchers in total mobility. A caveat is necessary for the analysis of direction. As noted in Section 3, the distance measure is invariant to rotation after the factor analysis and hence it is not necessary to take a stand on the definition of the underlying "true" factors. The interpretation at the individual factor level of changes in direction, however, is sensitive to rotation, or more generally, the precise definition of the factors. The analysis in this section assumes that the particular linear combinations of the DOT scores on the characteristics used in the factor analysis that define the four factors can be interpreted as reasonable approximations of four true factors.[40]

6.1 Wage Losses and Distance

The larger wage losses for skill portfolio switchers identified in Poletaev and Robinson (2008) are reflected in a strong negative relationship between wage changes following displacement and the continuous euclidean distance measures used in this paper. Tables 9 & 9a report regression results for log wage changes for males and females, respectively, following displacement using the equally weighted euclidean distance measure, *edist4* and the alternative measure *edist4w* that places greater weight on the more important skills in the pre-displacement occupation. To control for general year effects, including variation in unemployment rates, the specification reported includes a full set of survey year dummy variables; the results are almost identical if a quadratic in years and the unemployment rate are used instead. The remaining independent variables are the same age and education variables used in Section 4. The first row reports the results for the DWS 1984-2002 using the distance measure *edist4w*. This uses larger weights on the more important skills to capture some features of the discrete skill switch variables from Poletaev and Robinson (2008) that emphasize skill

[40]Regressing the log wage for the current job for the pooled DWS 1984-2002 sample yields highly significant "skill price" coefficients on the first three of these four factors with or without the inclusion of years of schooling and a quadratic in age. The addition of the factors substantially increases the explanatory power.

switching. It is sensitive to rotation since it uses the ranking information from the factors, though in practice the sensitivity is low to a number of conventional rotations. The second row reports the full specification using the benchmark distance measure *edist4* which is invariant to rotation. The coefficients on the age and educations variables, as well as the year or unemployment controls, are very similar with either distance measure. Moreover, the coefficients on the two types of distance measures are also very similar, especially for males.

The estimates are presented both for displacement from the private sector from plant closing and from displacements according to the broader BLS definition. The dependent variable is the difference in the log usual weekly earnings on the current job, from those of the pre-displacement job. The estimates are presented for all workers age 20-61, as well as for the more restricted sample of workers that were full time at both jobs. In addition, the relationship is estimated unconditionally and conditionally on switching 3-digit occupation. The final rows report the results for the same specifications using the more recent DWS 2004-2010. The results in Table 9 show generally highly significant negative coefficients on the distance measures. In terms of magnitude, these are very similar to the wage loss difference from the discrete skill portfolio switching reported in Poletaev and Robinson (2008).[41]

The results are quite similar for displacement according to the BLS definition, and for the more restrictive sample of displacements from plant closure. The effect from increasing the sample by removing the full time restriction on the pre- and post-displacement jobs is to increase the distance coefficient substantially. This suggests that in addition to finding a similar job after displacement, displaced workers may also initially have a problem in finding full time work in the jobs that are similar to their pre-displacement job. Thus the coefficient on the distance measure is reflecting both of these effects. For females, there are some insignificant effects of distance for the coefficients conditional on switching occupation when the full time restriction is imposed. Otherwise, the results for females are similar to those for males with generally large and significant distance coefficients. These results all control for education and age effects as well as the year (or unemployment rate)

[41] Applying the coefficients in Tables 9 & 9a to the difference in the mean difference for skill switchers and skill stayers, for example, using the PC1 definition, produces similar differences between the wage losses of skill switchers and skill stayers to those reported in Poletaev and Robinson (2008) Tables 4 & 5.

effects. Education itself, is usually insignificant, but there is a strong negative effect for older workers, reflecting a common result in the previous displacement literature.

6.2 Comparison of Direction in Total and Involuntary Mobility

The wage losses in Tables 9 & 9a reflect a downward direction in the skill portfolio for displaced workers. An important question for overall specific human capital loss is whether there is a similar downward direction in total mobility. A comparison of direction in the two forms of mobility is presented in Table 10. The four factors can be given the same shorthand names as in Poletaev and Robinson (2008), "general intelligence related" (f_1), "fine motor skills related" (f_2), "strength and gross motor skills related" (f_3) and "visual skills related" (f_4), based on inspection of the characteristics that loaded most heavily.[42] Results are reported for a restricted sample (Sample A) requiring full time work at both jobs and positive wage observations in both jobs in the DWS, and a larger sample (Sample B) relaxing these requirements. The sample sizes for the MCPS are large enough to yield precise estimates for either sample, while the smaller sample sizes for the DWS makes Sample B preferred in terms of precision. In either case, the main result is the same.

The top half of Table 10 reports the absolute changes (distance) in the individual factor scores, $|\Delta_i|$ for workers aged 20-61 switching occupation in the DWS and in the MCPS; the lower half reports the changes (direction), Δ_i. The lower half of Table 8 showed very similar conditional distance distributions overall for the DWS and the MCPS. The top half of Table 10 shows that this is reflected in the individual components of the distance measure which are quite similar in the DWS and the MCPS. The lower half of Table 10 shows that while the distance is similar, the direction is quite different. The direction for the MCPS is neutral or positive, while the direction for the DWS is typically downward. For males the strongest contrast is the significantly negative change for "intelligence" in the DWS (-0.1246 in Sample A and -0.1520 in Sample B) compared with the significantly positive change in the MCPS (0.0399 and 0.0469.) This associated with no significant changes in "fine motor" or "visual", but a shift towards the "strength" factor is apparent. For females there is also a significant contrast for "intelligence": the positive effect in the MCPS is similar in

[42] The pattern of factor loadings for the four factors are very similar to those in Poletaev and Robinson (2008), and again f_1 is the factor that explains most of the variance in the factor analysis.

magnitude to the males (0.0476 and 0.0545) and highly significant in both samples; the negative effect in the DWS is smaller than for the males (-.0230 and -.0463) and is statistically significant only in the larger Sample B. The negative effect for females in the DWS, however, is reflected in all the factor changes. For the smaller Sample A, only "strength" is statistically significant, but for the larger Sample B they are all statistically significant.

Table 10a repeats the analysis for older workers (40-61). The results confirm the downward direction for the DWS in contrast to the MCPS. As before, females show a drop in all factors in the DWS and a mildly positive change in the factors in the MCPS. For males, there is again a very large and significant drop in "intelligence", but this is now accompanied by a larger negative change in "fine motor", which is statistically significant in Sample B. These results are consistent with the total mobility in the MCPS reflecting lateral moves and promotions in addition to the typically negative changes from involuntary mobility reflected in the DWS evidence. The positive changes in the MCPS are muted in the older sample, especially for males, which may reflect a smaller fraction of promotions than in the full sample due to the removal of the age range in which the age earnings profile is steepest.

6.3 Interpretation of the Direction: Some Caveats

The interpretation of changes in direction at the individual factor level is subject to identification of the "true" factors. One approach to identification of the factors is to declare a specified number of true underlying factors (cognitive or non cognitive skills, fine motor skills, managerial skills, routine or non-routine tasks, etc.) and then, directly or indirectly, impose *a priori* identifying restrictions on the factor loading matrix in the factor analysis.[43] In the DOT case, this amounts to an *a priori* specification of which DOT characteristics measure which particular *a priori* factors. A common use of factor analysis is as an exploratory description of the data in the form of the factor loadings; subsequently the factor loadings, rotated or un-rotated, guide the search for "interpretable" factors. In this sense, pre-specified factors and standard factor analyses with a particular interpretation of

[43] For a general approach of this kind, see Carneiro, Hansen and Heckman (2003); for a more specific approach to defining true underlying factors from the DOT characteristics, see Yamaguchi (2011), Firpo, Fortin and Lemieux (2009).

the same number of factors based on a limited number of DOT characteristics loading heavily on each factor can be very similar.

In terms of the different direction in the MCPS and the DWS, the most sensitive issue occurs when not all the components move in the same direction. For example, a worker that previously used mainly gross motor skills that acquired fine motor skills or supervisory skills to get a promotion and higher earnings may have increased f_1 and f_2 at the expense of f_3. The relationship between the multidimensional skill vector and earnings is complicated, and a full analysis is beyond the scope of this paper. One approach that has been used in the literature is the simple linear skill pricing model.[44] A simple way of capturing the idea of the promotion and higher earnings example in the direction of the change in the multidimensional skill vector is to use the implied "skill prices" from a regression of earnings on the four skills as weights. In most cases in Tables 10 & 10a, the factor changes are significant in only one direction, so weights are not required to show a difference of direction in the two forms of mobility. However, for males 20-61 the pattern is for a large reduction in the "intelligence" factor f_1 combined with a smaller, but sizeable increase the "strength" factor f_3. A simple wage regression approach to obtaining skill prices for use as weights shows a much larger coefficient on f_1 than on f_3, for a variety of specifications. This is consistent with the interpretation of this particular combination of skill changes as a demotion.[45]

7 Conclusions

Recent research has linked occupational mobility and specific human capital loss. This paper develops an occupation distance measures that is used to provide a better understanding of occupational mobility, displacement and career mobility. Wage losses from involuntary mobility are strongly related to distance. The main results from the empirical analysis for involuntary mobility are as follows. There is considerable dispersion in the distance between pre- and post-displacement jobs, but over the period 1984-2002, displaced workers are increasingly finding new jobs that are close

[44] There is a general problem of whether the skills can be "unbundled" to allow this additive form, discussed in Heckman and Sheinkman (1987). The literature employing some form of skill pricing model includes Welch (1969), Ingram and Neumann (2006), Autor and Handel (2009) and Firpo, Fortin and Lemieux (2009).

[45] A regression of the log wage for the current job for the pooled DWS 1984-2002 sample yields highly significant "skill price" coefficients for f_1 of 0.152 for males and 0.195 for females, for f_2 of 0.079 for males and 0.073 for females, and for f_3 of 0.098 for males and 0.033; f_4 has a zero coefficient for males and a negative coefficient for females.

to their previous job. The rate of displacement also declined from 1984-2002, suggesting a decline in specific capital losses from involuntary mobility over this period. The major recession of 2008 is reflected in a significant increase in displacement and shift outward in the cumulative distribution of distance for displaced workers.

Displaced worker mobility is only one component of total mobility. Previous work has argued that specific human capital loss is also associated with other forms of mobility. A comparison of the distribution of distances for involuntary and total mobility shows quite similar distances. However, the two types of mobility differ markedly in their direction, as reflected in changes in the skill vectors associated with each occupation. While total mobility generally shows a neutral or mildly upward direction, involuntary mobility shows a significant strong downward direction. In addition, the downward direction is reflected in a strong negative association between wage change and distance following displacement. This suggests that the type of mobility is likely to have a significant effect on the amount of specific human capital loss.

The declining conditional distance for displaced workers suggests that the labour market may be more efficiently re-employing workers following displacement. The distance measure may therefore be useful in assessing how well the labour market matches workers and firms both over cycles as well as over time. The strong difference in direction of the total and involuntary mobility, however, indicates the need for future research to better understand the nature and measurement of the underlying skill vectors and their differential evolution in voluntary and involuntary mobility.

References

[1] Daron Acemoglu and David Autor. Skills, Tasks and technologies: Implications for Employment and Earnings. In Orley C. Ashenfelter and David Card, editors, *Handbook of Labor Economics*, volume 4B. Elsevier, 2011.

[2] David Autor and Michael Handel. Putting Tasks to the Test: Human Capital, Job Tasks and Wages. Working Paper, 2009.

[3] David H. Autor, Frank Levy, and Richard J. Murnane. The Skill Content of Recent Technological Change: An Empirical Exploration. *Quarterly Journal of Economics*, CXVIII:1279–1334, 2003.

[4] Pedro Carnero, Karsten T. Hansen, and J. Heckman, James. Estimating Distributions of Treatment Effects With an Application to Returns to Schooling and Measurement of the Effects of Uncertainty on College Choice. *International Economics Review*, 44:361–422, 2003.

[5] Raymond B. Cattell. *Abilities: Their Structure, Growth, and Action.* Houghton-Mifflin, Boston, 1971.

[6] Henry S. Farber. Job loss in the United States, 1981-2001. *Research in Labor Economics*, 23:69–117, 2004.

[7] Henry S. Farber. What do we know about job loss in the United States? Evidence from the Displaced Workers Survey, 1984-2004. *Economic Perspectives*, 2Q/2005:13–28, 2005. Federal Reserve Bank of Chicago.

[8] Sergio Firpo, Nicole Fortin, and Thomas Lemieux. Occupational Tasks and Changes in the Wage Structure. Working Paper, 2009.

[9] J.J.Heckman and G.Sedlacek. Heterogeneity, Aggregation, and Market Wage Functions: An Empirical Model of Self-Selection in the Labor Market. *Journal of Political Economy*, 93:1077–1125, 1985.

[10] Gueorgui Kambourov and Iourii Manovskii. Rising Occupational and Industry Mobility in the United States: 1968-1997. *International Economic Review*, 49:41–79, 2008.

[11] Gueorgui Kambourov and Iourii Manovskii. Occupational Specificity of Human Capital. *International Economic Review*, 50:63–115, 2009.

[12] Gueorgui Kambourov and Iourii Manovskii. A Cautionary Note on Using (March) CPS Data to Study Worker Mobility. Working Paper, 2010.

[13] Giuseppe Moscarini and Kaj Thomsson. Occupational and Job Mobility in the US. *Scandinavian Journal of Economics*, 109:807–836, 2007.

[14] Giuseppe Moscarini and Francis Vella. Aggregate Worker Reallocation and Occupational Mobility in the United States: 1976-2000. Working Paper, 2003.

[15] Giuseppe Moscarini and Francis Vella. Occupational Mobility and the Business Cycle. NBER Working Paper No. 13819, 2008.

[16] T. Scott Murray, Yvan Clermont, and Marilyn Binkley. *International Adult Literacy Survey. Measuring Adult Literacy and Life Skills: New Frameworks for Assessment*. Statistics Canada, Ottawa, 2005. Catalogue no. 89-552-MIE,no.13.

[17] Daniel Parent. Industry Specific Capital and the Wage Profile: Evidence from the National Longitudinal Survey of Youth and the Panel Study of Income Dynamics. *Journal of Labor Economics*, 18:306–321, 2000.

[18] Maxim Poletaev and Chris Robinson. Human Capital Specificity: Evidence from the Dictionary of Occupational Titles and Displaced Worker Surveys 1984-2000. *Journal of Labor Economics*, 26:387–420, 2008.

[19] Chris Robinson. Measures of Occupational Distance: Final Report. Report Prepared for Human Resources and Social Development Canada, 2009.

[20] A Spitz-Oner. Technical Change, Job Tasks, and Rising Educational Demands: Looking Outside the Wage Structure. *Journal of Labor Economics*, 24:235–270, 2006.

[21] Jay Stewart. Recent Trends in Job Stability and job Security: Evidence from the March CPS. BLS Working Paper 356, 2002.

[22] Shintaro Yamaguchi. Career Progression and Comparative Advantage. *Labour Economics*, 17:679–689, 2010.

[23] Shintaro Yamaguchi. Tasks and heterogeneous Human Capital. *Journal of Labor Economics*, 2011, forthcoming.

A Appendix

This appendix describes the construction of the skill vectors at the level of 3-digit census occupation codes. It also details the sample selection criteria used with the DWS and MCPS.

A.1 Construction of the Skill Vector by Modified 1990 Occupation Codes

The basic information source for the occupation skill vectors is the DOT. The final DOT master file (version 4.3) has 12741 unique DOT codes. Each DOT job has characteristics associated with it. Three characteristics come from the fourth, fifth and sixth digits of the DOT code itself: the complexity of the interaction with data, people and things. This provides a numerical ranking of the complexity. All other characteristics are provided in separate fields in the DOT master file. There are various ways of assigning values of the DOT characteristics to 3-digit occupations. Poletaev and Robinson (2008) used the final DOT master file (version 4.3) crosswalk between DOT coding and the 1990 census 3-digit occupation codes and averaged the DOT characteristic scores over the DOT jobs in each 1990 census 3-digit occupation code in the crosswalk. This implicitly assumes that employment within a three digit occupation is equally divided across the DOT jobs for that occupation, which is not in fact the case. An alternative approach is to use DOT level employment information made available in a special CPS monthly file for April 1971.

The special dual DOT and 3-digit occupation coded file is the DOT augmented CPS monthly file (April 1971), available at ICPSR as project 7845. The study description is as follows:

"The April 1971 CPS was coded routinely with 1970 occupation codes by census personnel. The occupational descriptions from this CPS were also coded with the 9 digit 3rd edition (1965) DOT codes and added to the tape by the staff of the Division of Occupational Analysis of the U.S. Employment Service. (Lloyd) Temme cleaned and edited the data. Using a map which relates third and fourth edition DOT codes (created by the Division of Occupational Analysis), he then added third edition GED, SVP and worker trait group scores as well as all fourth edition (1977) DOT occupational characteristics to the tape. In several instances there was no equivalence between the third and fourth edition codes, either because the title had been deleted between the two editions, or multiple fourth edition codes were created from a single third edition code. Temme reviewed the non matches on a case by case basis, assigning fourth edition codes by hand where possible. If no match could be made, 9's were assigned to the DOT codes. On the computer file there are 6984 persons with no information on fourth edition DOT codes."

The analysis in this paper covers the period 1982-2010. The occupation data for this period is coded in 1980, 1990 and 2000 codes. The special CPS April 1971 file contains 1970 occupation codes, and the DOT information in the file is for the original 4th edition (1977) DOT characteristics scores. Since many of the DOT jobs were given revised scores in the revised 4th edition (1991), these scores are most appropriate for the period of the analysis. There are two approaches to using the information in the special April 1971 CPS file. One is to use the file directly to obtain mean original (and revised) DOT characteristic scores by the 1970 occupation codes represented in the file and then to use dual occupation coded files to convert these to means by 1980, 1990 and 2000 codes ("dual occupation coded files approach"). The 1980 and 1990 coding schemes are almost the same. However, there is no simple relation between the 1970 and 1980 codes, or between the 1990 and 2000 codes, so there is no simple direct occupation crosswalk and special dual occupation codes files have to be used. The other approach is to use only the DOT level employment information from the special April 1971 CPS file to use as weights when calculating means by occupation for each of the relevant DOT crosswalk files ("weighted crosswalk approach").

A.1.1 Revised 4th Edition Characteristics

Characteristic scores have changed over time at the DOT job level. This is documented in the revised 4th edition (1991) of the DOT. The procedure for using the revised characteristic scores together with the original 4th edition codes available in the the April 1971 file is not straightforward because of a complicated relation between the original and revised 4th edition DOT jobs, especially the fact that the three middle digits of the job code are, in fact, ratings on the three characteristics "data", "people" or "things", so the same job gets a new code if the rating changes. The special file in fact has 3885 unique DOT (1977) codes assigned to the 53457 individuals in the file. This is a subset of the total number of unique original revised 4th edition DOT codes, which is over 12000. However, since the CPS special file is a large representative sample, the unused codes from the 12000 must have very low employment. These 3885 original (1977) DOT job codes have original characteristic scores assigned to them by those responsible for creating the special file. (These are in almost all cases the same as the values in the DOT 4th edition original (1977) crosswalk.) The task is to trace these 3885 1977 DOT jobs into the 1991 DOT edition which records the revised characteristic scores.

Unfortunately the jobs cannot be traced by the codes because the codes can change for two reasons: the middle three digits of the code have to change if there is an updating to the rating on any of the three characteristics, "data", "people" or "things"; and some codes were changed or merged or divided or dropped. Thus it is necessary to find the DOT jobs (codes) in the revised edition that correspond to the 3885 DOT jobs in the April 1971 file when the job is the same but the code may have changed. The revised 4th edition DOT master file containing the final updating for all the DOT characteristics contains the variable *dlu*, which indicates the last date on which the code or characteristics were updated. The file was examined for duplicates on the DOT code alone and on the DOT code and DOT title together, before construction of a file with unique codes with revised (1991) characteristic scores. Similarly, the April 1971 file was examined for duplicates on the DOT code alone and on the DOT code and DOT title together, before construction of a file with unique codes with original (1977) characteristic values that appear in the special 1971 CPS file.

Examination of *dlu* shows 10289 of the 12741 unique DOT jobs in the revised 4th edition had not been updated since 1977. These should therefore be codes that are the same in both the original and revised editions with characteristic values that are unchanged between the original and revised editions. The remaining 2452 DOT jobs in the revised 4th edition may be new or revised codes, or the same codes with trailer characteristics updated.[46] The next step was to identify which of the the original 3885 codes in the April 1971 file needed updating to the equivalent job code in the revised edition, and those that do not. There were 2715 codes that represented DOT jobs that were unchanged between 1977 and 1991. The characteristic values for these in the April 1971 file should be the same as in the 1991 master file.[47] The titles from both original and revised editions were compared for these 2715 codes. For 2563 of the the 2715, the titles are identical. Inspection of the

[46]The trailer characteristics are all the characteristics except for the three complexity scores for "data", "people" and "things".

[47]In fact, using the GED characteristics scores to test this, 98.9% are the same, but 30 cases are different by 1.

remaining 152 shows that almost all titles differ only in the use or non use of hyphens or roman numerals, confirming that they are the same job.

The 1170 codes that appear only in the April 1971 file and not in the master file have to be traced into the revised 1991 edition. Autor et al. (2003) used the DOT materials in the ICPSR study (6100) that describe the relation between the original DOT 4th edition and the revised 4th edition. The ICPSR project 6100 documents the changes between the original 4th edition and revised 4th edition. The relevant section is "Conversion Tables of Code Changes Fourth to Revised Fourth Edition Dictionary of Occupational Titles." This provides tables containing "occupations with a DOT code and/or title which changed or was deleted between 1977 and 1991 or appeared since the publication of the Fourth Edition in 1977." The occupational analysts revised 646, combined 136 and deleted 75 occupational codes and titles, so these 875 DOT codes and/or titles do not appear in the revised 4th edition, but are identified in these tables. In addition, between 1977 and 1991 analysts reviewed, updated and/or added 2452 occupational elements to the Dictionary. These were too extensive to tabulate, but can be identified by the "Date of Last Update" (*dlu*) variable in the DOT data files. For these cases *dlu* will have a value other than "77". (I follow Autor et al. (2003) in assuming that any job that was not revised experienced no task change.) While there were 2452 "reviewed, updated and/or added 2452 occupational elements", it was only necessary to trace the relevant 1170 1977 DOT job codes in the April 1971 file.

First, "Conversion Tables of Code Changes Fourth to Revised Fourth Edition Dictionary of Occupational Titles" was used to identify which of the relevant 1170 codes had a title or code change between the editions. Of the 1170 codes, 265 were identified in Table 1 "Occupational code and/or title changes," with 181 having code changes and 85 having only title changes. This left 905 codes still to be traced into the revised edition. Of the 905 remaining codes to be traced, 25 codes were identified from Table 2 "Occupations Deleted From or Combined with Another in the Revised Fourth Edition DOT" as deleted and 52 as combined. This left 829 original codes to be traced into the revised edition. The variable *dlu* identifies that they are codes that were reviewed and possibly updated after 1977. Most of these may have identical codes in the revised edition, where the "data", "people" and "things" characteristic scores stayed the same (otherwise the code would have changed) but some of the "trailer" values changed. To identify these cases the 829 codes were merged with a file of the 2452 unique revised 1991 DOT codes with *dlu* values indicating an update after 1977.

The merge shows 827 common codes. Of these, 784 (94.8%) also have the identical title. The remaining 43 do not have the same title, but examination shows that the only differences in the titles are in spelling, hyphens, etc. Thus all these original 4th edition codes are the same jobs as the revised 4th edition codes, where the code is the same because the revision did not affect the "data", "people" or "things" rating, and hence the 3 middle digits stayed the same. In summary, of the 3885 original codes from the special 1971 CPS file, almost all can be traced into the revised edition. 2715 are the same job with the same code and title and so are all directly traced. Of the remainder all but 25 (deleted) can be traced. Examination of the April 1971 file shows that these 25 codes make up less than 0.2% of the population.

A.1.2 Factor Analysis

The skill vector characterization at the 3-digit occupation is provided by the normalized factor scores from a weighted factor analysis where the unit of observation is a modified 1990 occupation code and the data for each observation is a (49×1) vector of revised DOT characteristic scores for the occupation and the weights are modified 1990 census occupation employment in 1992. Given the period of availability of data from the DWS, the analysis uses the revised DOT characteristic scores. The strong similarity of the 1980 and 1990 census occupation codes allowed the construction of a modified 1990 census occupation coding scheme that provides consistent occupation coding for the DWS for all the surveys from 1984 to 2002. The correspondence between the original 1980 and 1990 codes and the modified 1990 coding is given in Appendix Table A1.

In order to implement this factor analysis, it remains to construct revised DOT characteristic score means by modified 1990 occupation. The dual occupation coded files approach starts with the means by the 1970 occupation coding directly from the special April 1971 file. (The means were calculated separately for males and females where possible.) There is no simple correspondence between the 1970 and 1980 occupation codes, so a simple occupation code crosswalk is not available to convert these to means by 1980 codes. Instead the special sample from the 1970 census that was retrospectively dual coded with 1980 codes, known as the "Treiman File" was used. First, the predicted revised characteristic scores were assigned to the individuals in this file based on their sex and their 1970 occupation code. Next, modified 1990 codes were assigned to individuals in this file based on their 1980 code. Revised characteristic score means by modified 1990 code were then estimated by averaging (separately by sex) over individuals in each modified 1990 code. This produced characteristic score means for 489 modified 1990 occupations; for 387 of these, characteristic mean scores are available separately for males and females; for 96, characteristic means are available for males only and for 6 for females only, since these occupations were exclusively male or exclusively female in the Treiman file.

The alternative weighted crosswalk approach to constructing revised DOT characteristic score means by modified 1990 occupation was simply to use the DOT employment level weights (separately by sex where possible) from the April 1971 file to calculate weighted means by 1990 occupation of the revised DOT characteristic scores directly available in the final version (4.3) master file of the DOT. From the master file, a file was created with 12741 unique DOT codes, the 1990 census occupation code and the revised DOT characteristics scores. This provides a single 1990 census occupation code for each DOT code.[48] The total number of 1990 occupations for which means can be calculated using this approach is a little less than the total of 501. The missing occupations are primarily post-secondary teaching occupations by specialty that do not have DOT codes in the master file. Most of these, however, are available from the dual occupation codes files approach.

To create the data for the factor analysis, the characteristic mean scores are assigned to all workers in the January 1992 CPS file, based on their modified 3-digit 1990 occupation using the weighted crosswalk approach means where available, supplemented by the means from the dual occupation coded files approach. There are characteristic score means for 493 modified 1990 occupations; for 360 of these, characteristic score means are given separately for males and females; for 124, characteristic score means are given for males only and for 9 characteristic means are given for females only. In assigning estimated characteristic means to all individuals in the 1992 CPS employment sample, the

[48]The crosswalk assignment of the 1990 census codes is done by the crosswalk center on the basis of the common underlying SOC codes.

values are assigned based on sex as well as occupation code. Where an individual male is employed in 1992, but no male characteristic score mean is available, the female value is assigned; similarly, where an individual female is employed in 1992, but no female characteristic value is available, the male value is assigned. There are a few occupations, 499, 564, 655, 659, 681 and 740 that are exclusively male in the characteristics score means file that do not have any employment in 1992; they are assigned a minimum weight of 1 to include them in the factor analysis.[49]

Since the unit of observation for the factor analysis is the individual in the 1992 sample and both male and female characteristic means are used in the factor analysis, they both contribute to the factors that explain the variance across individuals (or, in effect, weighted occupation/sex combinations.) By construction, the factor means across (weighted) individuals are zero and the variance is one. Predicting out the factors at the individual level means that within an occupation, males and females can have different values of the factors. The occupation employment distributions by sex in the 1992 CPS file resulted in the 493 modified occupations having 406 occupations with separate scores for males and females, 10 exclusively female and 77 exclusively male. The factors can be calculated for any individuals in any data sets with occupation coding by assigning the characteristic score means and applying the factor scoring coefficients from the factor analysis.[50]

The factor analysis produced a 4 vector of factor scores that are normalized to a mean of zero and standard deviation of one in the 1992 population. The factors are all treated symmetrically in the equally weighted euclidean distance measure *edist4*; for the weighted measure *edist4w* the highest weight, 0.4, is placed on the "main" skill in the occupation, as defined by the skill or factor with the highest score for that occupation. The remaining weights (0.3, 0.2, 0.1) are assigned in descending order to the second, third and fourth most important skill as defined by the factor with the second, third and fourth highest score for that occupation.

[49] There is one occupation where DOT characteristic score means are not available that has employment in 1992: 106. This is a tiny fraction of the sample and is dropped.

[50] The factor scoring coefficients have to be applied to standardized characteristic score means, subtracting the same mean and dividing by the same standard deviation as was done for the characteristic scores used in the factor analysis.

A.2 Extension to Census 2000 Occupation Codes and Consistency in the Distance Measure

The change from 1990 to 2000 Census occupation coding was a major one, causing a break in 3-digit occupation series.[51] There is no simple crosswalk between the 1990 and 2000 occupation codes. The same two approaches were taken to construct characteristic score means and factors for the 2000 occupation codes. The extension to the 2000 codes using the dual occupation coded files approach was carried out using dual coding available from merging the original 2000-2002 monthly basic CPS files coded with 1990 codes with the special 2000-2002 CPS Extract Files coded with 2000 codes.[52] Given the rotation structure of the CPS, a large sample (approximately 300,000) of individuals (without repetition) was constructed by taking all observations with observed occupation codes for the month-in-sample 1-4 from the January, May and September surveys in each of the three years. These monthly files were matched to the original CPS Basic Monthly data files using QSTNUM, HRMONTH, HRYEAR4, and OCCURNUM using the file formats created by Jean Roth at NBER and merged to form a dual 1990/2000 occupation coded file. Individuals in the dual coded file were then assigned factor scores based on their modified 1990 occupation coding, and the means of these factor scores by 2000 census occupation codes were then calculated by sex. This produced a vector characterization of the same four factors for 502 occupations using 2000 coding allowing the same distance measures, *edist4* and *edist4w* to be created.

The extension to 2000 occupation codes using the weighted crosswalk approach starts with the DOT crosswalk file, *dotcen00*. A duplicates analysis shows there are some duplicate cases, i.e. DOT codes that have more than one 2000 census occupation code. Almost all DOT codes have a single occupation code; a few have two occupation codes and a few have four occupation codes. The employment weights were assigned to the DOT codes such that DOT level employment was shared equally across the two or four occupations in the duplicates cases, and DOT characteristic score means by 2000 occupation were calculated separately by sex using these weights. The occupation

[51]For details see the Census Technical Paper #65 "The Relationship Between the 1990 Census and Census 2000 Industry and Occupation Classification Systems" (October 2003)

[52]From 2003, the CPS adopted the 2002 Census Bureau industry and occupation classification systems, which are derived from the 2002 North American Industry Classification System and the 2000 Standard Occupation Classification System. For comparison over time, the BLS released the 2000-2002 Current Population Survey Extract files. The Extract files contain the 2000 occupation codes.

counts show 478 in total with 126 male only and 8 female only. This is not a full count of all 2000 occupations due to some being unrepresented in the crosswalk. It was therefore supplemented with 26 occupations from the dual occupation coded files approach. This results in 504 occupations with 126 exclusively male and 8 exclusively female.

The weighted crosswalk method produces distance distributions from random mobility that are almost the same across 1970, modified 1990 and 2000 occupation coding. By contrast, the dual occupation coded files approach produces distributions that are very similar in shape, but that shift to the left after each application of the dual occupation coded files. Table A2 reports comparison statistics from the two approaches. Using the dual occupation coded files approach, as coding shifts from 1970 to modified 1990 to 2000, the mean distance for males shows the decreasing shift: $1.31 - - > 1.25 - - > 1.10$; and females $1.27 - - > 1.21 - - > 1.04$, which represent substantial shifts. A probably explanation for this is averaging caused by random occupation coding error in the dual occupation coded files which results in the measured mean distance declining with each application of a dual coded file to make the transition. By contrast, using the weighted crosswalk approach the sequences are $1.28 - - > 1.29 - - > 1.26$ for males and $1.26 - - > 1.27 - - > 1.25$ for females, which show essentially no break. The distributions from the dual occupation coded files approach are plotted for males and females in Figures A1a & A1b. The distributions using the weighted crosswalk method are given in Figures A2a & A2b.

A.3 Sample Restrictions for the DWS and MCPS

The distance distributions for the DWS are estimated for workers displaced from the private sector. The only other sample restrictions were the requirement for valid occupation codes for pre- and post-displacement jobs and displacements restricted to the last 3 years. The DWS analysis of direction at the individual factor level uses two samples. Sample A approximates the sample in Poletaev and Robinson (2008). It restricts the sample to full time private sector workers with positive current and past weekly earnings observations, and excludes displacements from construction and agriculture. It also drops workers 62 and older. The larger sample B drops the requirement for weekly earnings observations and full time status.

The data for the analysis of this section come from MCPS. This records a three-digit occupation code for employed workers in their current job in (the third week of) March of the survey year and a three-digit occupation code for the longest job held in the previous year ("earnings year".) The main sample restrictions follow Moscarini and Vella (2003). This restricts the sample to past and current employees and excludes observations after the 1988 survey in which all of the supplement responses were allocated, including occupation code. The allocation flag for current occupation shows a fraction around 1% or less up to 1988. The 1989 survey year was the first with a new system of treating allocated values, especially for the annual supplement, with the introduction of the variable *suprec* indicating allocated supplement responses. From this point the fraction allocated for current occupation increases, ranging between $3-6\%$ up to earnings year 1997. For the earnings years 1998 to 2001 the fraction increases to over 11%. The allocation flag for occupation in the longest job last year is not available until the survey year 1989. After that, the allocation fraction is extremely low, but this is because this only refers to allocated values in supplementary records where the whole record was not allocated. The *suprec* variable shows that starting in 1989 nearly 10% of supplementary records were allocated, including occupation. The only additional restrictions are the age range 20-64, and that the longest job last year was in the private sector, for comparison with the DWS. For the analysis of direction, the restricted sample A imposes the full time requirement and excludes those who were incorporated self-employed from the "employees" in the longest job last year. The larger sample A removes these restrictions.

Figure 1a

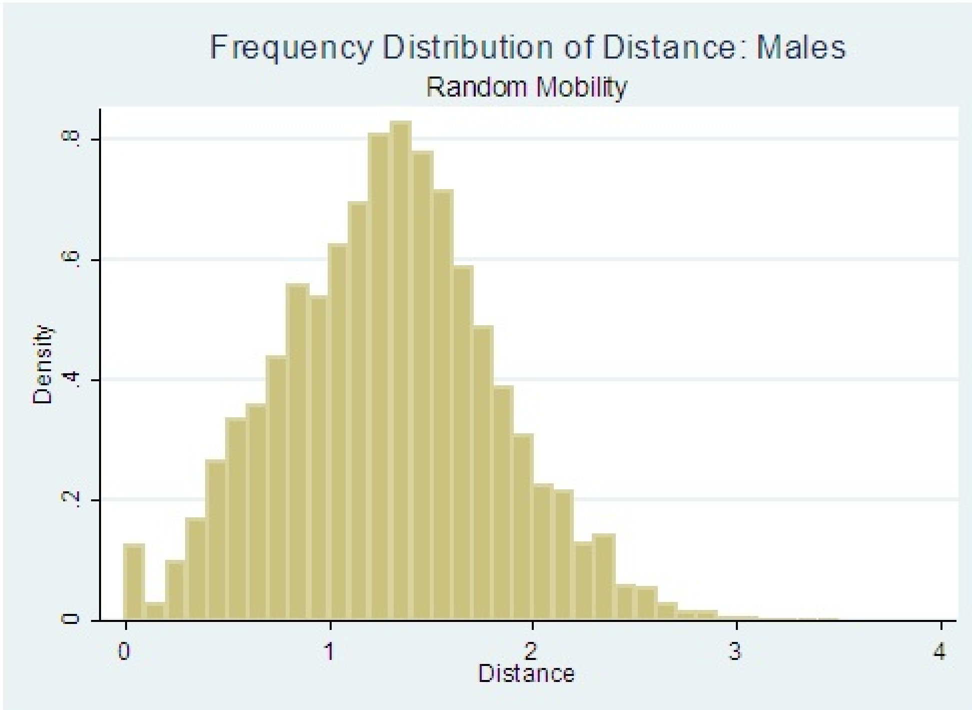


Figure 1b

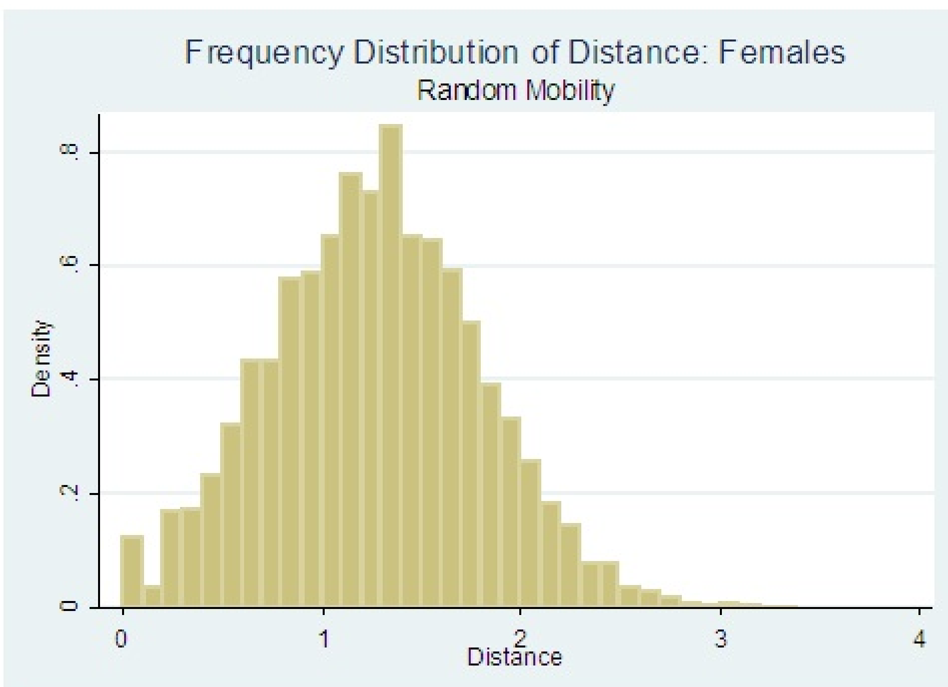


Figure 2

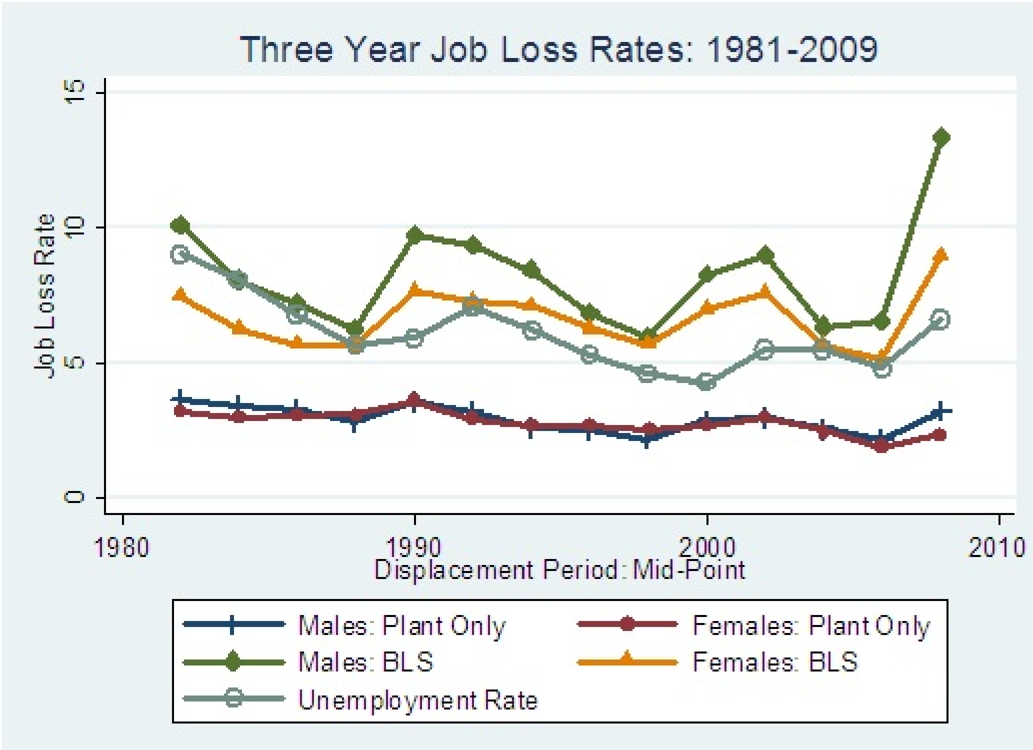


Figure 3

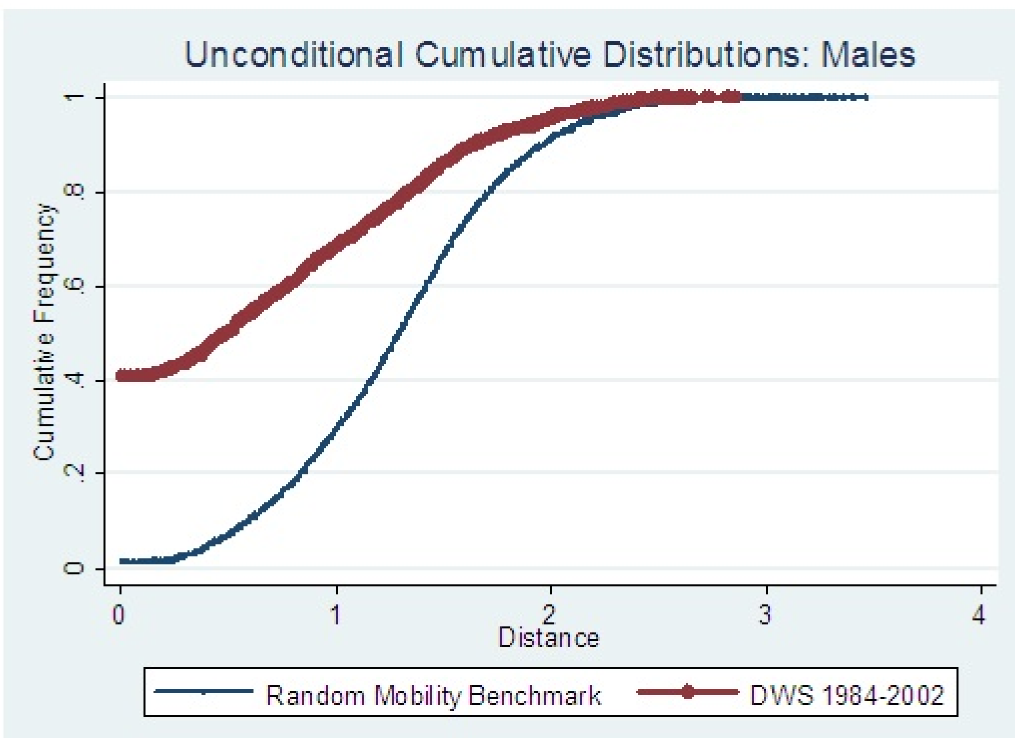


Figure 4

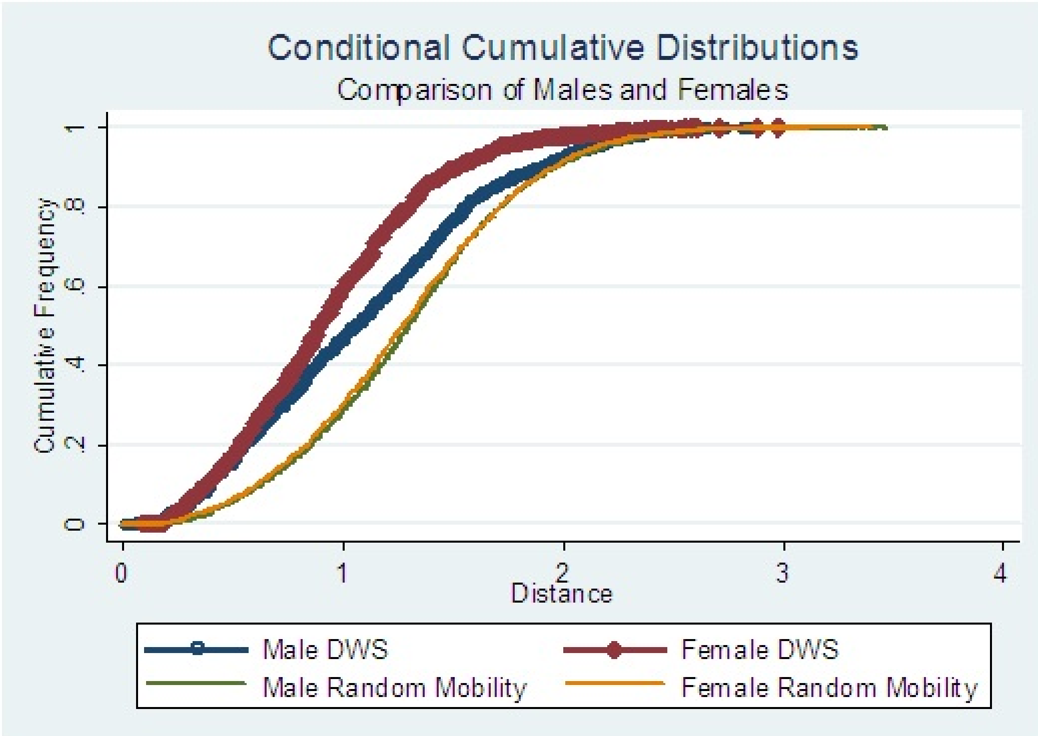


Figure 5

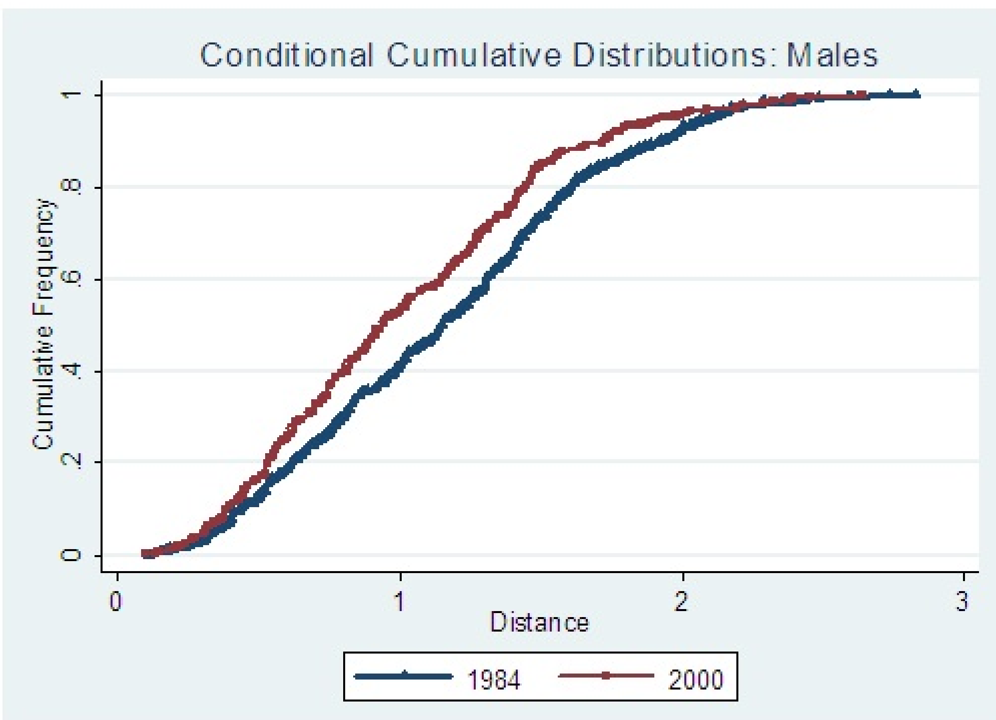


Figure 6

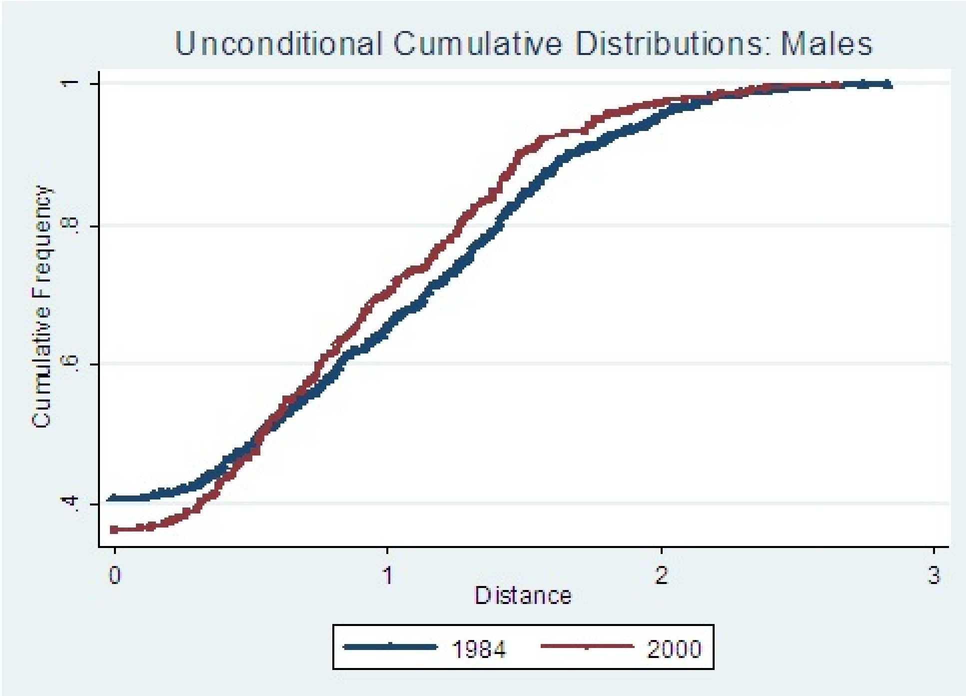


Figure 7a

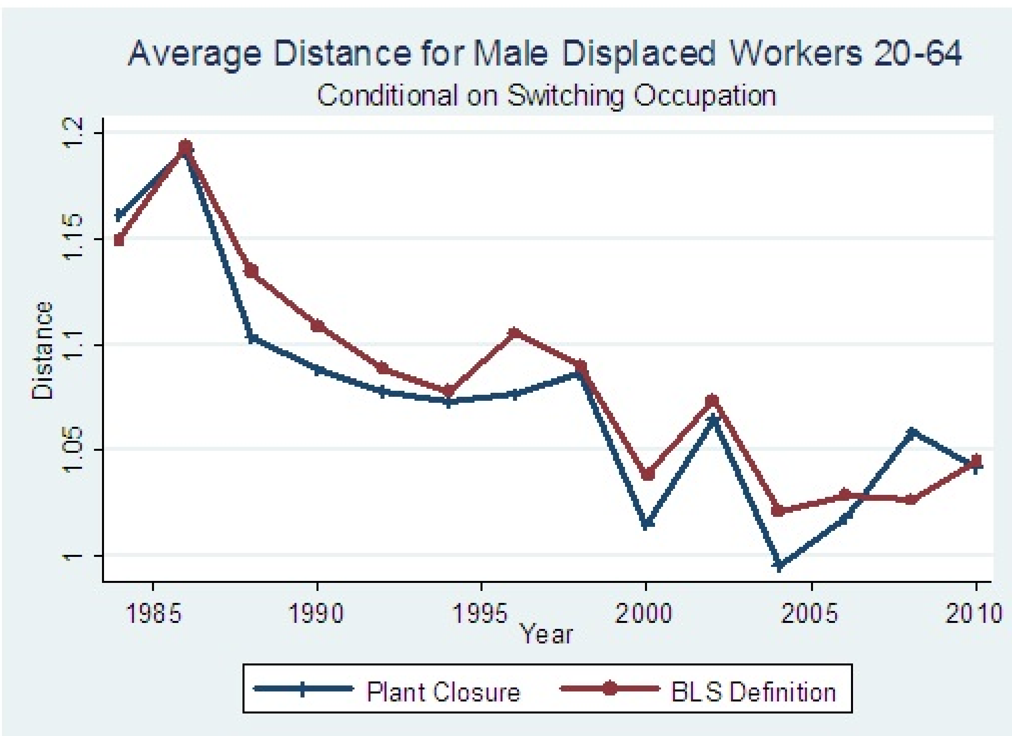


Figure 7b

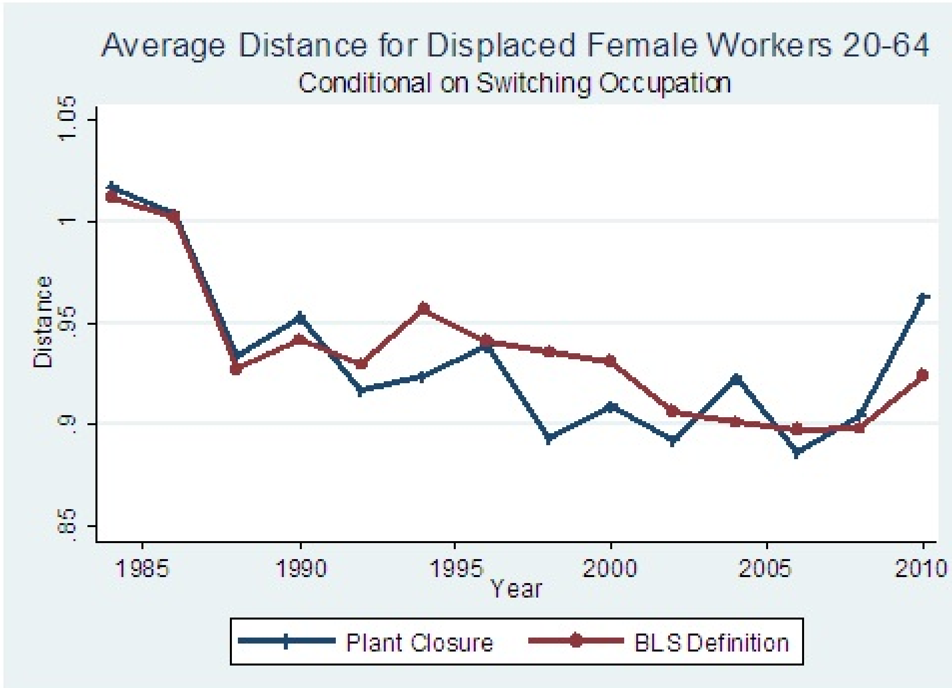


Figure 8

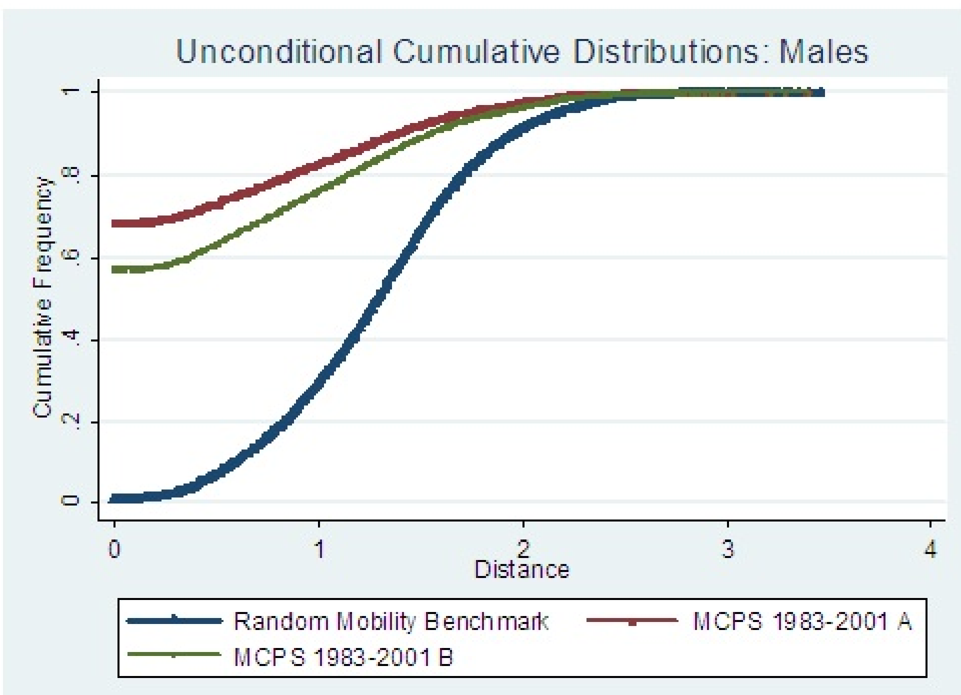
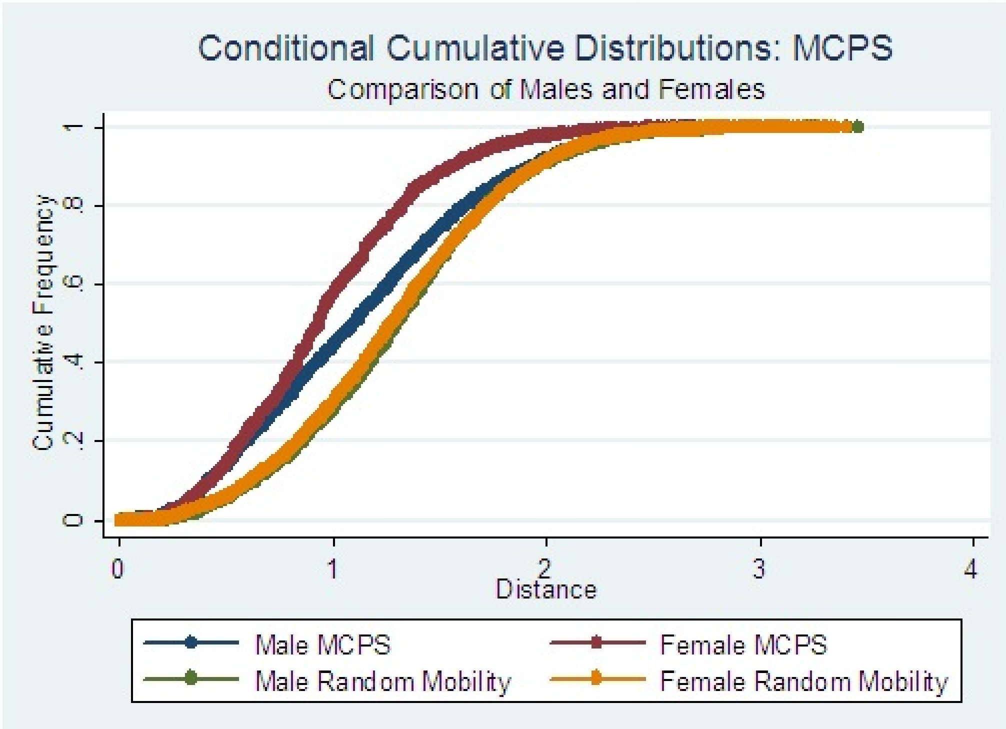


Figure 9



Conditional Cumulative Distributions: MCPS
Comparison of Males and Females
Cumulative Frequency
0
.2
.4
.6
.8
1
0
1
2
3
4
Distance
Male MCPS
Female MCPS
Male Random Mobility
Female Random Mobility

Table 1

Comparison of the Distance Distributions from Random Mobility Within and Across 3-Digit Occupations

	Within Occupation	Across Occupations	
		Males	Females
Mean	.694	1.288	1.269
Median	.709	1.296	1.269
Standard Deviation	.254	.523	.527
90^{th} Percentile	.982	1.96	1.956
Fraction d<1	.915	.290	.308

Notes: The distributions for moves across occupations are the distance distributions for males and females from random mobility conditional on switching occupation using the 1992 occupation weights. Given the very high fraction of occupation switching in random mobility, this is similar to the unconditional distance distributions in Figures 1a & 1b. The distribution for distance within occupation cannot use weighting at the DOT job level, but does use 1992 occupation weights. The results without weights are very similar.

Table 2
Three Year Displacement Rates for Workers 20-64 by Survey Year

	Males		Females	
	Plant Closing Only	BLS Definition	Plant Closing Only	BLS Definition
1984	.0366	.1010	.0322	.0746
1986	.0344	.0804	.0299	.0622
1988	.0330	.0722	.0306	.0567
1990	.0281	.0624	.0309	.0560
1992	.0361	.0968	.0362	.0763
1994	.0324	.0930	.0296	.0728
1996	.0263	.0844	.0273	.0713
1998	.0248	.0683	.0268	.0626
2000	.0224	.0592	.0256	.0567
2002	.0289	.0826	.0270	.0700
2004	.0300	.0899	.0301	.0758
2006	.0259	.0633	.0249	.0564
2008	.0219	.0652	.0194	.0516
2010	.0325	.1337	.0237	.0897

Table 3a
Occupation Switching for Workers 20-64 Displaced by Plant Closing by Survey Year

	Males			Females		
	3-digit	2-digit	Major Group	3-digit	2-digit	Major Group
1984	.5914	.5311	.4549	.6866	.6045	.5134
1986	.5932	.5273	.4623	.6784	.6009	.4985
1988	.5671	.5081	.4339	.5916	.5224	.4003
1990	.5648	.5062	.4277	.6067	.5457	.4340
1992	.5281	.4522	.3754	.5606	.4814	.3639
1994	.6113	.5271	.4573	.6511	.5530	.4533
1996	.6300	.5428	.4677	.6506	.5830	.4749
1998	.6393	.5863	.4974	.6870	.5689	.4626
2000	.6362	.5594	.4790	.6335	.5239	.4263
2002	.6115	.5338	.4461	.6857	.5857	.4619
Coding Break						
2004	.6566			.6765		
2006	.6226			.6829		
2008	.6355			.7005		
2010	.6734			.6448		

Notes: Displacements from the private sector.

Table 3b
Occupation Switching for Displaced Workers 20-64 (BLS Definition) by Survey Year

	Males			Females		
	3-digit	2-digit	Major Group	3-digit	2-digit	Major Group
1984	.5753	.5128	.4403	.6531	.5876	.4820
1986	.5651	.5080	.4377	.6561	.5885	.4827
1988	.5453	.4968	.4270	.5897	.5198	.4020
1990	.5202	.4636	.3915	.5709	.5166	.4097
1992	.4944	.4251	.3565	.5180	.4459	.3413
1994	.5948	.5165	.4445	.6446	.5535	.4399
1996	.6225	.5488	.4751	.6559	.5766	.4588
1998	.6436	.5748	.4870	.6647	.5623	.4531
2000	.6048	.5359	.4353	.6634	.5672	.4586
2002	.5845	.5107	.4355	.6683	.5717	.4638
Coding Break						
2004	.6410			.6809		
2006	.6117			.6604		
2008	.6227			.6889		
2010	.6149			.6323		

Notes: Displacements from the private sector.

Table 4a
Expected Distance for Workers Displaced by Plant Closing Conditional on Switching Occupation

	Pooled Males & Females		Males		Females	
Displacement Years						
1985-89	-.0808		-.0790		-.0670	
	(5.01)		(3.40)		(3.19)	
1989-93	-.1024		-.1003		-.0902	
	(6.40)		(4.35)		(4.31)	
1993-97	-.1056		-.0936		-.0950	
	(6.06)		(3.65)		(4.26)	
1997-2001	-.1317		-.1318		-.1111	
	(7.81)		(5.36)		(5.11)	
survey year		-.0239		-.0396		-.0040
		(2.32)		(2.63)		(0.29)
$(\text{survey year})^2$		.0011		.0020		.0000
		(1.92)		(2.34)		(0.02)
unemployment rate		.1344		.2867		-.0571
		(1.25)		(1.82)		(0.40)
$(\text{unemployment rate})^2$		-.0095		-.0211		.0051
		(1.18)		(1.80)		(0.48)
female		-.1618				
		(15.15)				
some college		-.0255		-.0468		-.0011
		(1.74)		(2.13)		(0.06)
BA+		-.1832		-.2815		-.0501
		(10.55)		(11.11)		(2.17)
Age 30-39		-.0417		-.0460		-.0320
		(3.14)		(2.38)		(1.82)
Age 40-64		-.0681		-.0686		-.0524
		(5.27)		(3.57)		(3.12)

Notes: Absolute value of t-statistics in parentheses. Omitted categories are high school graduates age 20-29. The dropout category was included but was insignificant.

Table 4b
Expected Weighted Distance for Workers Displaced by Plant Closing Conditional on Switching Occupation

	Pooled Males & Females		Males		Females	
Displacement Years						
1985-89	-.0942		-.0933		-.0765	
	(5.44)		(3.70)		(3.43)	
1989-93	-.1094		-.1108		-.0901	
	(6.36)		(4.46)		(4.05)	
1993-97	-.1141		-.0972		-.1040	
	(6.08)		(3.52)		(4.39)	
1997-2001	-.1434		-.1429		-.1194	
	(7.90)		(5.38)		(5.16)	
survey year		-.0259		-.0438		-.0031
		(2.34)		(2.71)		(0.22)
(survey year)^2		.0012		.0022		.0001
		(1.95)		(2.44)		(0.07)
unemployment rate		.1456		.3124		-.0640
		(1.26)		(1.85)		(0.42)
(unemployment rate)^2		-.0103		-.0230		.0057
		(1.18)		(1.81)		(0.51)
female		-.1918				
		(16.75)				
some college		-.0326		-.0555		-.0060
		(2.08)		(2.35)		(0.30)
BA+		-.2123		-.3227		-.0620
		(11.41)		(11.83)		(2.53)
Age 30-39		-.0370		-.0419		-.0266
		(2.60)		(2.01)		(1.43)
Age 40-64		-.0634		-.0680		-.0416
		(4.58)		(3.29)		(2.33)

Notes: See notes to Table 4a.

Table 5 (R6m)

Expected Distance for Displaced Workers (BLS Definition) Conditional on Switching Occupation

	edist4		*edist4w*	
1985-89	-.0636		-.0686	
	(5.61)		(5.62)	
1989-93	-.0811		-.0823	
	(7.67)		(7.24)	
1993-97	-.0795		-.0830	
	(7.12)		(6.91)	
1997-2001	-.1138		-.1210	
	(10.37)		(10.25)	
survey year		-.0243		-.0255
		(3.45)		(3.38)
(survey year)^2		.0012		.0013
		(3.09)		(2.98)
unemployment rate		.1992		.2036
		(2.69)		(2.56)
(unemployment rate)^2		-.0144		-.0148
		(2.59)		(2.49)
female		-.1590		-.1924
		(22.25)		(25.13)
some college		-..0330		-.0417
		(3.43)		(4.04)
BA+		-.1916		-.2122
		(17.45)		(18.03)
Age 30-39		-.0257		-.0245
		(2.94)		(2.62)
Age 40-64		-.0629		-.0633
		(7.34)		(6.89)

Notes: See notes to Table 4a.

Table 6
Decomposition of Unconditional Expected Distance from Involuntary Mobility

	Displacements from Plant Closing				All Displacement (BLS)			
	$\delta(t)$	$\alpha(t)$	$D_c(t)$	$\delta(t)D(t)$	$\delta(t)$	$\alpha(t)$	$D_c(t)$	$\delta(t)D(t)$
Males								
1981-83	.0366	.5914	1.1615	.0251	.1010	.5753	1.1496	.0668
1983-85	.0344	.5932	1.1915	.0243	.0804	.5651	1.1942	.0543
1985-87	.033	.5671	1.1038	.0206	.0722	.5453	1.1347	.0447
1987-89	.0281	.5648	1.0885	.0173	.0624	.5202	1.1091	.0360
1989-91	.0361	.5281	1.0778	.0206	.0968	.4944	1.0888	.0521
1991-93	.0324	.6113	1.0734	.0213	.0930	.5948	1.0779	.0597
1993-95	.0263	.6300	1.0772	.0179	.0844	.6225	1.1053	.0581
1995-97	.0248	.6393	1.0873	.0172	.0683	.6435	1.0899	.0479
1997-99	.0224	.6362	1.0149	.0144	.0592	.6047	1.0384	.0372
1999-01	.0289	.6115	1.0650	.0188	.0826	.5845	1.0740	.0518
Females								
1981-83	.0322	.6866	1.0168	.0225	.0746	.6531	1.0123	.0493
1983-85	.0299	.6784	1.0036	.0204	.0622	.6561	1.0025	.0409
1985-87	.0306	.5916	.9339	.0169	.0567	.5897	.9274	.0310
1987-89	.0309	.6067	.9527	.0179	.0560	.5709	.9416	.0301
1989-91	.0362	.5606	.9167	.0186	.0763	.5180	.9298	.0367
1991-93	.0296	.6511	.9235	.0178	.0728	.6446	.9567	.0449
1993-95	.0273	.6506	.9383	.0166	.0713	.6559	.9410	.0440
1995-97	.0268	.6870	.8927	.0164	.0626	.6647	.9355	.0389
1997-99	.0256	.6335	.9084	.0147	.0567	.6634	.9314	.0350
1999-01	.0270	.6857	.8921	.0165	.0700	.6683	.9063	.0424

Table 7
Expected Distance for Workers Switching Occupation: MCPS 1983-2002

	Males & Females		Males		Females	
1988-1992	-.0118		-.0110		-.0089	
	(2.27)		(1.40)		(1.37)	
1993-1997	-.0225		-.0212		-.0152	
	(4.16)		(2.58)		(2.27)	
1998-2002	-.0239		-.0308		-.0048	
	(4.43)		(3.72)		(0.72)	
survey year		-.0026		-.0017		-.0043
		(1.66)		(0.70)		(2.15)
$(\text{survey year})^2$		.0001		.0001		.0002
		(1.91)		(0.70)		(2.19)
unemp rate		.0313		.0591		.0031
		(2.64)		(3.24)		(0.21)
$(\text{unemp rate})^2$		-.0023		-.0042		-.0003
		(2.51)		(3.05)		(0.27)
female		-.1618				
		(43.07)				
dropout		-.0020		-.0175		.0068
		(0.32)		(1.95)		(0.79)
some college		-.0286		-.0403		-.0113
		(6.27)		(5.56)		(2.01)
BA+		-.1302		-.2245		-.0232
		(24.85)		(27.91)		(3.50)
Age 30-39		-.0391		-.0553		-.0128
		(8.73)		(8.09)		(2.27)
Age 40-64		-.0481		-.0774		-.0049
		(10.07)		(10.37)		(0.83)

Notes: Dependent variable is *edist4*. Absolute value of t-statistics in parentheses

Table 8
Comparison of Total and Involuntary Mobility

	DWS	MCPS	
		A: θ = .17	B: θ = .10
	Unconditional Distribution		
Occupation Stayers (%)			
Males	40.86 (0.54)	68.29 (0.14)	57.13 (0.17)
Females	36.05 (0.60)	65.49 (0.15)	54.41 (0.18)
d < 1 (%)			
Males	68.27 (0.51)	82.28 (0.11)	76.03 (0.15)
Females	73.73 (0.55)	85.27 (0.11)	80.55 (0.14)
Expected Distance			
Males	.6518 (.0076)	.3597 (.0018)	.4864 (.0023)
Females	.6015 (.0072)	.3351 (.0017)	.4426 (.0020)
	Conditional Distribution		
d < 1 (%)			
Males	46.35 (0.71)	44.10 (0.26)	44.10 (0.26)
Females	58.92 (0.77)	57.33 (0.26)	57.33 (0.26)
Expected Distance			
Males	1.1022 (.0079)	1.1346 (.0029)	1.1346 (.0029)
Females	.9405 (.0070)	.9709 (.0023)	.9709 (.0023)

Notes: Standard errors in parentheses.

Table 9
Expected Log Wage Changes for Males following Displacement: DWS 1984-2010

	Unconditional				Conditional on Switching Occupation			
	Plant		BLS		Plant		BLS	
	All	FT	All	FT	All	FT	All	FT
	Displaced Worker Surveys 1984-2002							
edist4w	-.0691	-.0444	-.0815	-.0500	-.0819	-.0398	-.0529	-.0214
	(5.63)	(3.96)	(9.52)	(6.28)	(4.19)	(2.22)	(3.91)	(1.71)
edist4	-.0685	-.0451	-.0822	-.0499	-.0802	-.0392	-.0476	-.0149
	(5.30)	(3.82)	(9.12)	(5.96)	(3.78)	(2.02)	(3.24)	(1.10)
dropout	-.0467	-.0289	-.0234	-.0196	-.1040	-.0704	-.0450	.0117
	(1.44)	(0.96)	(1.01)	(0.89)	(2.48)	(1.76)	(1.52)	(0.42)
some college	.0076	-.0045	-.0204	-.0140	-.0314	-.0299	-.0343	-.0139
	(0.31)	(0.20)	(1.18)	(0.88)	(0.96)	(0.99)	(1.53)	(0.67)
BA+	.0486	.0334	.0183	.0195	.0337	.0202	.0247	.0254
	(1.75)	(1.36)	(0.98)	(1.15)	(0.88)	(0.59)	(0.99)	(1.11)
Age 30-39	-.0947	-.0872	-.0908	-.0859	-.0961	-.0913	-.0993	-.0898
	(4.41)	(4.43)	(6.11)	(6.23)	(3.50)	(3.60)	(5.21)	(5.10)
Age 40-61	-.1981	-.1534	-.2121	-.1747	-.2323	-.1822	-.2395	-.1905
	(9.28)	(7.82)	(14.22)	(12.58)	(8.33)	(7.04)	(12.27)	(10.50)
	Displaced Worker Surveys 2004-2010							
edist4w	-.0994	-.0610	-.1011	-.0589	-.0585	-.0156	-.0750	-.0220
	(3.12)	(1.94)	(5.52)	(3.09)	(1.17)	(0.30)	(2.64)	(0.73)
edist4	-.1024	-.0638	-.1038	-.0640	-.0570	-.0130	-.0742	-.0267
	(3.07)	(1.93)	(5.40)	(3.22)	(1.06)	(0.23)	(2.43)	(0.83)

Notes: All specifications include survey year dummies. Absolute values of t-statistics in parentheses.

Table 9a
Expected Log Wage Changes for Females following Displacement: DWS 1984-2010

	Unconditional				Conditional on Switching Occupation			
	Plant		BLS		Plant		BLS	
	All	FT	All	FT	All	FT	All	FT
	Displaced Worker Surveys 1984-2002							
edist4w	-.0694	-.0468	-.0731	-.0338	-.0823	-.0169	-.0717	-.0065
	(3.75)	(2.87)	(5.77)	(2.94)	(2.84)	(0.64)	(3.70)	(0.36)
edist4	-.0508	-.0437	-.0596	-.0319	-.0408	-.0058	-.0419	.0010
	(2.66)	(2.63)	(4.57)	(2.72)	(1.32)	(0.21)	(2.01)	(0.05)
dropout	.0005	.0092	-.0717	-.0049	.0020	.0430	-.0599	.0281
	(0.01)	(0.26)	(2.47)	(0.18)	(0.04)	(0.89)	(1.63)	(0.78)
some college	.0332	.0276	.0176	.0154	.0450	.0463	.0392	.0381
	(1.15)	(1.08)	(0.89)	(0.85)	(1.25)	(1.41)	(1.59)	(1.67)
BA+	.0830	.0138	.0671	.0350	.0838	.0139	.0880	.0573
	(2.40)	(0.47)	(2.96)	(1.76)	(1.85)	(0.35)	(3.05)	(2.20)
Age 30-39	-.0750	-.0393	-.0957	-.0764	-.0780	-.0404	-.0978	-.0847
	(2.85)	(1.68)	(5.28)	(4.64)	(2.44)	(1.33)	(4.38)	(4.05)
Age 40-61	-.1295	-.1057	-.1422	-.1231	-.1519	-.1270	-.1620	-.1483
	(5.09)	(4.72)	(8.08)	(7.68)	(4.72)	(4.29)	(7.40)	(7.21)
	Displaced Worker Surveys 2004-2010							
edist4w	-.1071	-.0344	-.1150	-.0745	-.0444	.0339	-.0672	-.0529
	(2.66)	(0.69)	(4.11)	(2.54)	(0.77)	(0.54)	(1.66)	(1.41)
edist4	-.1003	-.0340	-.0999	-.0688	-.0207	-.0423	-.0282	-.0395
	(3.43)	(0.68)	(3.49)	(2.32)	(0.34)	(0.64)	(0.65)	(1.01)

Notes: All specifications include survey year dummies. Almost identical results occur using a quadratic in years and the unemployment rate. Absolute values of t-statistics in parentheses.

Table 10
Change in the Skill Portfolio of Occupation Switchers: Workers 20-61

	Sample A				Sample B			
	f_1	f_2	f_3	f_4	f_1	f_2	f_3	f_4
Distance of Change in Factor Score ($\lvert \Delta_i \rvert$)								
Males								
DWS	.7918	.9222	.7900	.9780	.8089	.9517	.8392	1.0122
MCPS	.7289	.8919	.7934	.9325	.7606	.9142	.8338	.9697
Females								
DWS	.6821	.8589	.9352	.6413	.7049	.8460	.9253	.6739
MCPS	.6365	.8407	.8307	.6496	.6706	.8644	.8700	.6986
Direction of Change in Factor Score (Δ_i)								
Males								
DWS	-.1246	.0045	.1048	-.0198	-.1520	-.0005	.1420	-.0310
	(.0230)	(.0266)	(.0225)	(.0312)	(.0174)	(.0206)	(.0177)	(.0241)
MCPS	.0399	-.0022	-.0054	.0328	.0469	-.0021	.0007	.0267
	(.0052)	(.0063)	(.0055)	(.0072)	(.0045)	(.0054)	(.0048)	(.0062)
Females								
DWS	-.0230	-.0360	-.1028	-.0128	-.0463	-.0493	-.0936	-.0319
	(.0227)	(.0287)	(.0305)	(.0220)	(.0154)	(.0184)	(.0199)	(.0151)
MCPS	.0476	.0192	.0213	.0101	.0545	.0239	.0266	.0234
	(.0052)	(.0069)	(.0068)	(.0050)	(.0040)	(.0051)	(.0051)	(.0043)

Notes: Sample B requires only valid occupation codes; sample A imposes restrictions on hours and earnings similar to Poletaev and Robinson (2008). Details are given in the Appendix.

Table 10a
Change in the Skill Portfolio of Occupation Switchers: Workers 40-61

	Sample A				Sample B			
	f_1	f_2	f_3	f_4	f_1	f_2	f_3	f_4
Distance of Change in Factor Score ($\lvert \Delta_i \rvert$)								
Males								
DWS	.7751	.8494	.7425	.9588	.7845	.8678	.7715	1.0050
MCPS	.6592	.8033	.6936	.8291	.6822	.8277	.7155	.8779
Females								
DWS	.6369	.8440	.9099	.6125	.6618	.8281	.9031	.6367
MCPS	.6048	.8120	.7765	.6458	.6441	.8308	.8061	.6927
Direction of Change in Factor Score (Δ_i)								
Males								
DWS	-.1760	-.0781	.0263	-.0020	-.2142	-.0839	.0797	.0282
	(.0388)	(.0428)	(.0371)	(.0522)	(.0294)	(.0328)	(.0288)	(.0419)
MCPS	.0071	.0042	-.0168	.0075	.0057	-.0075	-.0216	.0187
	(.0091)	(.0109)	(.0092)	(.0123)	(.0082)	(.0097)	(.0083)	(.0113)
Females								
DWS	-.0516	-.0939	-.0642	-.0428	-.0584	-.1053	-.0756	-.0646
	(.0354)	(.0461)	(.0493)	(.0355)	(.0243)	(.0300)	(.0328)	(.0242)
MCPS	.0343	.0228	.0187	.0542	.0366	.0208	.0136	.0626
	(.0091)	(.0121)	(.0117)	(.0100)	(.0072)	(.0093)	(.0091)	(.0081)

Notes: Sample B requires only valid occupation codes; sample A imposes restrictions on hours and earnings similar to Poletaev and Robinson (2008). Details are given in the Appendix.

Table A1
Modified 1990 Occupation Codes

			occc90 Pre 1996		*occ90* Post 1995
occ80	*occ80b*	*occ90b*			
463	463	461	461		461
464	464	462		etc.	
465	465	463			
466	466	464			
467	467	465			
633	633	628			
863	863	864			
864	864	865			
865	865	866			
866	866	867			
867	867	868		etc.	
873	873	874	874		874
{3, 16, 17, 18, 19}	19	22	{3,16,17,18,19,21,22}		{17,18,19,21,22}
{178, 179}	178	178	{178, 179}		178
{349, 353}	353	353	353		353
{368, 369}	368	368		etc.	
{436, 437}	436	436			
{673, 674}	674	674			
{794, 795}	795	795		etc.	
{804, 805}	804	804	804		804
468	468	468	{466, 467, 468}		{466, 467, 468}
{all other codes}	{identical codes}	{identical codes}			

Notes: The original 1980 and 1990 codes are *occ80* and *occ90*, respectively; the modified codes that are consistent over time are *occ80b* and *occ90b*.

Table A2
Consistency in the Distance Distribution Across Occupation Coding Changes

	Dual Occupation Coded Files		*Weighted Crosswalk*	
	Max	Mean	Max	Mean
	1970 Coding			
Males	3.56	1.31	3.59	1.28
Females	3.79	1.27	5.25	1.26
	1990 Coding			
Males	3.47	1.25	3.46	1.29
Females	3.55	1.21	3.40	1.27
	2000 Coding			
Males	3.28	1.10	3.32	1.26
Females	3.42	1.04	3.43	1.25

Figure A1a

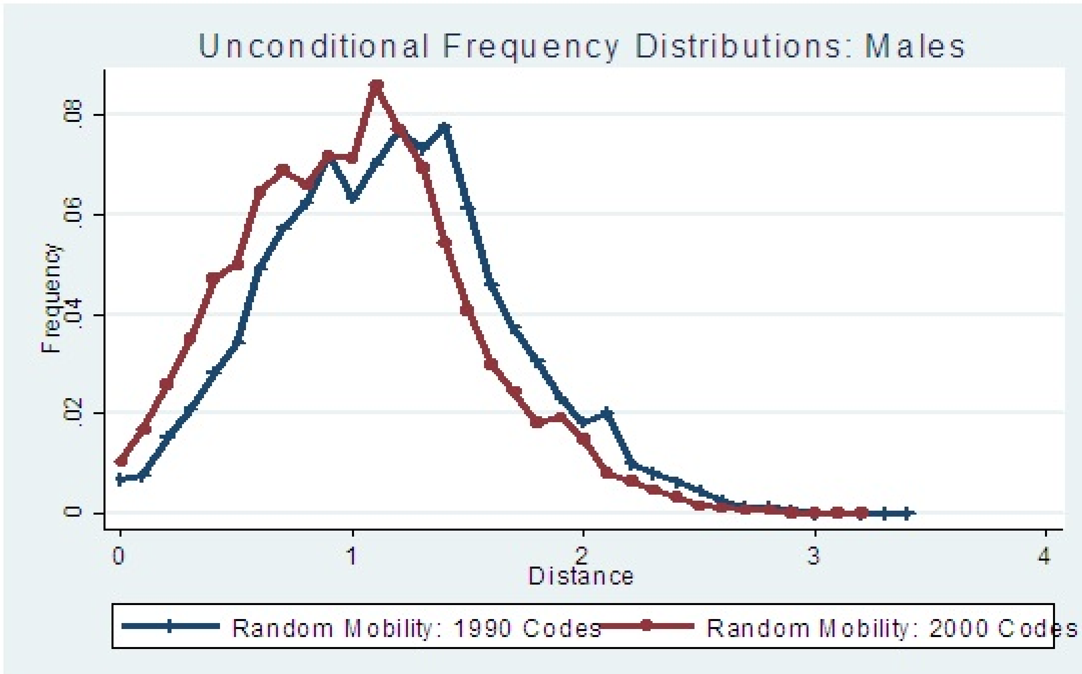


Figure A1b

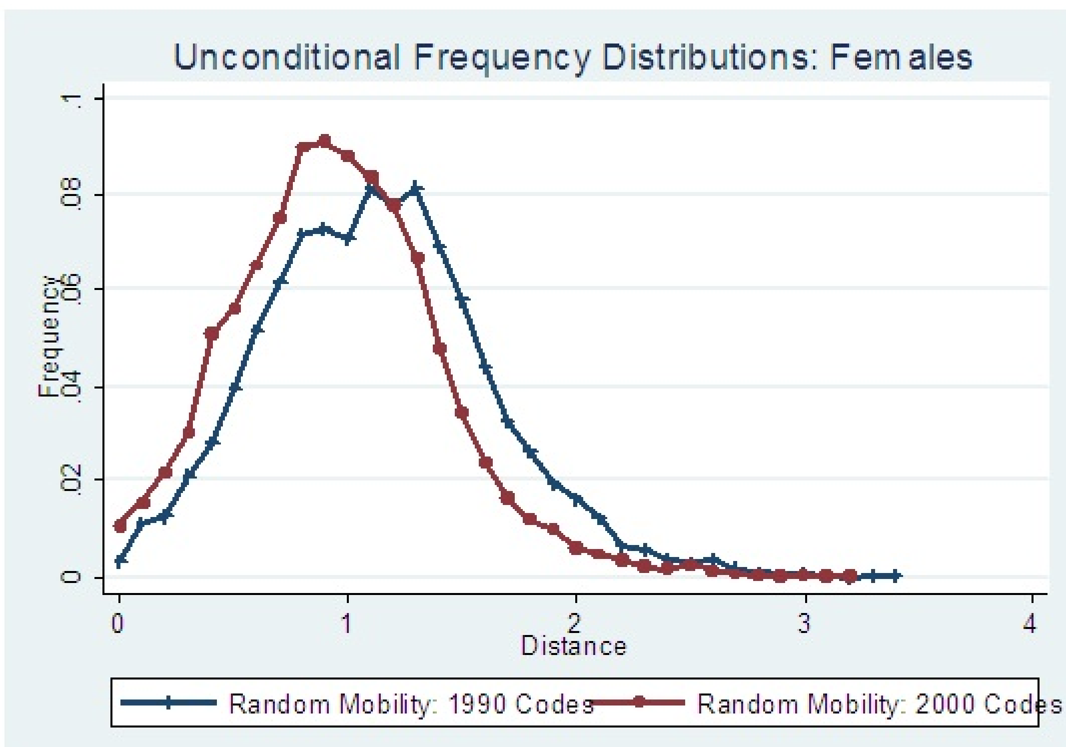


Figure A2a

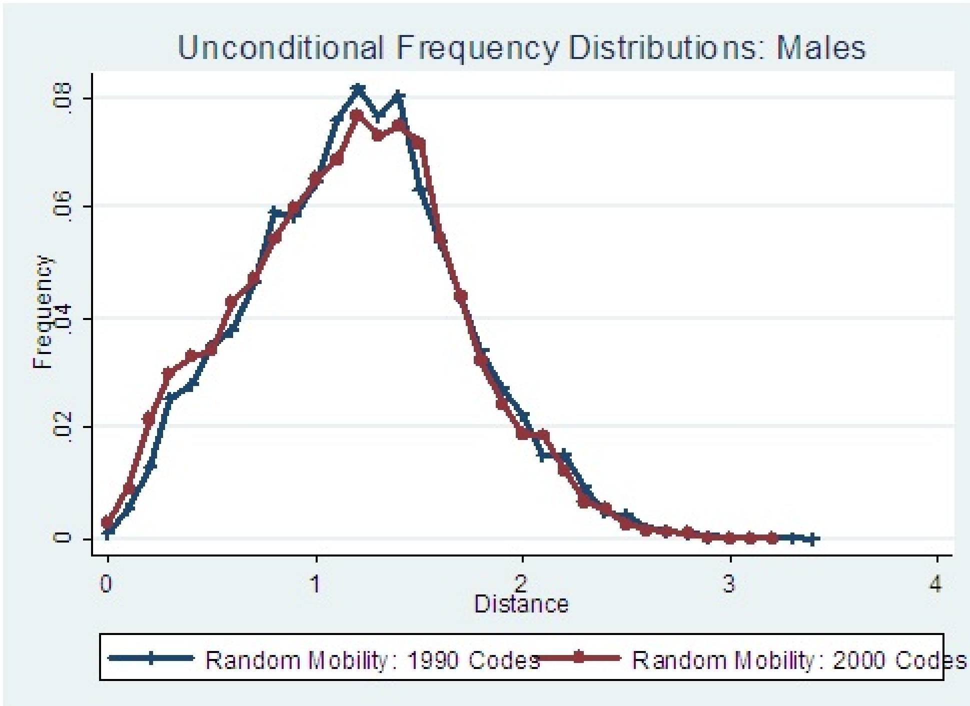


Figure A2b