November 2007

Validating U.S. Earnings Mobility Measures

Lisa M. Dragoset
Cornell University, LMD49@CORNELL.EDU

Gary S. Fields
Cornell University, gsf2@cornell.edu

Validating U.S. Earnings Mobility Measures

Abstract
Earnings mobility has been studied at both the macro level (the amount of mobility in an economy) and the micro level (the correlates of individuals' income changes). While measurement error is recognized as potentially important at both these levels, very little is known about the degree to which earnings mobility estimates are affected by measurement error. We compare micro and macro earnings mobility estimates for the U.S. during the 1990s using both survey-based earnings and administrative-based earnings. We find that measurement error in survey-based earnings has little qualitative effect on mobility estimates, but often has a large quantitative effect.

Keywords
earnings, mobility, measurements, census, survey

Comments

Suggested Citation
Dragoset, L. M., & Fields, G. S. (2007). *Validating U.S. earnings mobility measures* [Electronic version]. Washington, DC: United States Census Bureau.

Validating U.S. Earnings Mobility Measures[1]

Lisa M. Dragoset[2] and Gary S. Fields[3]

Abstract

Earnings mobility has been studied at both the macro level (the amount of mobility in an economy) and the micro level (the correlates of individuals' income changes). While measurement error is recognized as potentially important at both these levels, very little is known about the degree to which earnings mobility estimates are affected by measurement error. We compare micro and macro earnings mobility estimates for the U.S. during the 1990s using both survey-based earnings and administrative-based earnings. We find that measurement error in survey-based earnings has little qualitative effect on mobility estimates, but often has a large quantitative effect.

[1] This document reports the results of research and analysis undertaken by the U.S. Census Bureau staff. It has undergone a Census Bureau review more limited in scope than that given to official Census Bureau publications. This document is released to inform interested parties of ongoing research and to encourage discussion of work in progress. This research is a part of the U.S. Census Bureau's Longitudinal Employer-Household Dynamics Program (LEHD), which is partially supported by the National Science Foundation Grants SES-9978093 and SES-0427889 to Cornell University (Cornell Institute for Social and Economic Research), the National Institute on Aging Grant R01~AG018854-02, and the Alfred P. Sloan Foundation. The views expressed on statistical and methodological issues are those of the author(s) and not necessarily those of the U.S. Census Bureau, its program sponsors or data providers. Some or all of the data used in this paper are confidential data from the LEHD Program. The U.S. Census Bureau supports external researchers' use of these data through the Research Data Centers (see www.ces.census.gov). For other questions regarding the data, please contact Jeremy S. Wu, Program Manager, U.S. Census Bureau, LEHD Program, Demographic Surveys Division, FOB 3, Room 2138, 4700 Silver Hill Rd., Suitland, MD 20233, USA. (Jeremy.S.Wu@census.gov http://lehd.dsd.census.gov). We thank Peter Gottschalk, Erica Groshen, George Jakubson, and John Abowd for their helpful comments, as well as the members of the Cornell University Research Seminar on Labor Markets in Comparative Perspective and participants at the 2006 Society of Labor Economists annual meeting.

[2] lmd49@cornell.edu, Cornell University.

[3] gsf2@cornell.edu, Cornell University.

1 Introduction

Income mobility is defined as the change in income from one period to another for the same individual; compensation mobility concerns the change in incomes from the labor market (labor earnings plus benefits). Earnings mobility concerns only the change in labor earnings, excluding all benefits such as employer contributions to 401(k) plans and health insurance plans. The empirical literature on income and earnings mobility in various countries around the world is voluminous; see Atkinson, Bourguignon, and Morrisson (1992), Baulch and Hoddinott (2000), and Chronic Poverty Research Centre (2004) for surveys.

It is widely recognized that incomes and earnings are measured with error, the existence of which casts doubt on two main conclusions in the mobility literature (Duncan and Hill, 1985; Deaton, 1997; Bound, Brown, and Mathiowetz, 2001; Fields et al., 2003 (ii)). One is measuring the amount of mobility in a country over time ("macro mobility") – for example, gauging the extent of movement between income groups such as quintiles or calculating the correlation between initial income and final income. Mistakenly interpreting movements in measured earnings that are purely due to measurement error as movements in actual earnings would produce more apparent changes between income groups than in fact took place and likewise a lower correlation between initial and final income than truly occurred. A second potentially problematical area has been that of determining the correlates of individual mobility within the income distribution ("micro mobility"), or which income groups experience the most positive or negative income changes. Measurement error in initial income produces a spurious link between income change and initial income level, producing the appearance of convergent mobility, i.e. high-income people gaining less in dollars or percentages than low-income people.

Researchers have responded to the concern about measurement error in several ways. One is to note the concern and proceed to use measured incomes despite it. This is by far the most common approach to the measurement error issue. A second response is to use administrative records rather than survey reports. This approach has dominated research on income mobility in France, in which a whole series of studies have been conducted using administrative data; see, for example, Bigard, Guillotin, and Lucifora (1998), Buchinsky, Fougère, and Kramarz (1998), and Buchinsky, Fields, Fougère, and Kramarz (2003), among others. A third response, found in the U.S. literature, is to measure the differences between results obtained using survey data compared with the results using administrative records. Such studies are called "validation studies" and are surveyed in Bound, Brown, and Mathiowetz (2001).

In this paper, we conduct a more comprehensive validation study than heretofore, addressing the specific issues cited above as well as many others. We gauge the effect of measurement error in survey-based earnings on mobility estimates for the United States by comparing the estimates of macro and micro mobility obtained when using earnings reported by respondents in household surveys with the estimates obtained using an independent administrative measure of earnings. We use a new dataset that contains individually reported total annual labor earnings from the Survey of Income and Program Participation (SIPP) linked to employer-reported total annual labor earnings from the Social Security Administration's Detailed Earnings Record (DER). Treating the administrative DER earnings as equal to "true" latent earnings, we ask how much mobility estimates are affected by measurement error in survey-based earnings. The specific research questions are as follows: For the United States in the 1990s, how much are macro mobility estimates, micro mobility profile estimates, and micro mobility regressions affected by measurement error in survey-based earnings? Do those individuals who do best (worst) in the univariate profile results also do best (worst) when holding other things equal in the regression results? We use several different concepts and measures of macro mobility and use mobility profiles and regression models to analyze micro mobility.

Our general findings are twofold. The first is that measurement error in survey-based earnings has little qualitative effect on mobility estimates. The great majority of qualitative results hold when administrative records are used instead of survey responses. In particular, we find evidence of convergent mobility using both survey-based and administrative-based data, and using both unconditional (univariate) analysis and conditional (multiple regression) models. The second is that in several cases, measurement error in survey-based earnings has a large effect on the quantitative magnitudes of mobility estimates, though not in a systematic direction. It is impossible to know how typical the U.S. results are compared to what would be found in other countries. Nonetheless, from these two findings we conclude that analysts should continue doing research using survey data when only survey data are available.

The balance of the paper is organized as follows. We review the previous literature in Section 2, describe the data in Section 3, discuss the empirical methodology and results in Sections 4 and 5 respectively, and conclude in Section 6.

2 Previous Evidence

A great many measures have been used to gauge how much mobility there is in an economy. These measures include the correlation between initial income and final income, the elasticity of final income with respect to initial income, the proportion of income recipients who change income quintile, the average change in log-income, the average absolute value of income change, the chi-squared value in a contingency table, the average number of dollars gained by the winners and lost by the losers, and many others.

It is now understood that these measures gauge different mobility concepts. These six concepts are: time dependence, which measures the degree to which individuals' earnings in one year are determined by their earnings in the previous year; positional movement, which is measured by observing individuals' changes in economic positions in earnings distributions (either ranks, centiles, deciles, or quintiles); share movement, which happens when individuals' shares of total earnings in the population change; earnings flux, which concerns the size of changes in individual's earnings levels but not their sign; directional earnings movement, which measures how many people move up or down the earnings distribution and by how much; and mobility as an equalizer of longer-term earnings, which compares the inequality of earnings at a point in time with the inequality of earnings over a longer time period (Fields 2001, 2004).

There is very little evidence concerning how much estimates of these six different macro mobility concepts may be affected by measurement error. A very large literature uses only survey-based data to study macro and micro mobility in the U.S. See Atkinson, Bourguignon, and Morrisson (1992) for an excellent review of the earlier literature. Later studies include Gottschalk and Moffitt (1994), Gittleman and Joyce (1995, 1996), Buchinsky and Hunt (1996), Burkhauser, Holtz-Eakin, and Rhody (1997), Fields and Ok (1999), and Hisnanick and Walker (2004).

A much smaller literature uses only administrative-based data to study mobility. In an attempt to work with an error-free measure of earnings, a number of researchers working on France have used administrative-based earnings measures rather than survey-based earnings measures. This data set, the DADS (Declarations Annuelles de Données Sociales) from the French national statistical office INSEE, was used for example by Buchinsky et al. (2003) to study income mobility in France. The French data set is administrative only and does not include a survey-based measure of earnings. Therefore, researchers working on France have not been able to examine how mobility estimates change when using survey-based versus administrative-based earnings data.

The previous literature offers a small number of studies that make selective comparisons of survey-based versus administrative-based results, but no previous study has made such comparisons as comprehensively as we do in this paper; see Bound, Brown, and Mathiowetz (2001) for a complete survey of this literature through the 1990s and Abowd and Stinson (2005) and Gottschalk and Huynh (2006) for more recent contributions. A few of these validation studies also look at measurement error in earnings changes, defined as the difference between survey-based earnings changes and administrative-based earning changes. Duncan and Hill (1985) find no statistically significant difference between mean earnings changes based on the individual survey reports versus the employer records for a single large U.S. manufacturing firm. However, these earnings changes are obtained by differencing reports of earnings in two calendar years from the same interview, rather than differencing reports of annual earnings from two different interviews in a longitudinal study.

Duncan and Hill (1985), Bound and Krueger (1991), Bound et al. (1994), and Pischke (1995) all find evidence of "mean-reverting measurement error," defined as low earners tending to overstate earnings in surveys relative to administrative reports and high earners tending to understate them. Bound and Krueger (1991) report that for men nearly 65% of the observed variation in earnings changes is true variation, while for women the corresponding percentage is 80%. These four studies also find that measurement error is positively correlated over time.

Pischke (1995) was the first to establish the relationship between measurement error and earnings dynamics. Pischke proposed a simple model in which annual earnings are composed of a permanent (random-walk) component and a transitory (white noise) component and measurement error is composed of a person-specific component which is constant over time, a component which is correlated with the transitory component of earnings, and white noise. When this model was applied to the Panel Study of Income Dynamics Validation Study (PSIDVS) data, he found that the white-noise error more than offset the underreporting of transitory earnings, resulting in a slight understatement of the permanence in earnings changes in the survey-based data, relative to the administrative-based data. The PSIDVS sample was small and not representative. Our study builds on Pischke's work by using a larger and nationally representative sample to study the effect of measurement error on earnings changes estimates.

Abowd and Stinson (2005) used the same SIPP-SSA public use file that we use in this study. They created a person-job level dataset from the SIPP-SSA file by matching each SIPP respondent's reported jobs to his/her jobs from the Detailed Earnings Record (taken from Box 1 on the W-2 form) by employer name. They assumed that neither survey-based nor administrative-based earnings equal "true" earnings, but that both are measured with error, and estimated the ratio of error to total variance to be 0.67 for survey-based earnings changes and 0.71 for administrative-based earnings changes.

As stated above, several studies find evidence of "mean-reverting" measurement error, or a negative correlation between the measurement error and the value of earnings as given by the employer-recorded or administrative earnings. To formalize how this finding will affect estimates of micro mobility, we follow Kim and Solon (2005) and consider the textbook model of errors-in-variables:

$$(1)\ y_{it} = y_{it}^{*} + w_{it},$$

where y_{it} is observed earnings, y_{it}^{*} is true earnings, and the measurement error w_{it} is assumed to have zero mean and to be orthogonal to y_{it}^{*}. This model can be viewed as a restricted version of a more general model of measurement error:

$$(2)\ y_{it} = n_i + \lambda y_{it}^{*} + w_{it},$$

where n_i is an individual-specific effect for reporting error and w_{it} is again uncorrelated with y_{it} and each of its determinants. The textbook model of measurement error is the case where $n_i = 0$ and $\lambda = 1$. The evidence of "mean-reverting" measurement error found in the literature corresponds to a value of λ that falls between 0 and 1. Differencing equation (2) leads to

$$(3)\ \Delta y = \lambda \Delta y^{*} + \Delta w.$$

Now suppose the earnings mobility equation we wish to estimate takes the following unconditional form:

$$(4)\ \Delta y^{*} = \rho \Delta x + \varepsilon,$$

where x is a vector of determinants and ε is independently and identically distributed and orthogonal to Δx. What the researcher is actually able to estimate is the following:

$$(5)\ \Delta y = \rho_1 \Delta x + \varepsilon_2.$$

Least squares will provide a consistent estimate of $\rho_1$since both components of the error term (ε_2 and Δw) are orthogonal to the regressors. But if $0 < \lambda < 1$, then least squares provides estimates of ρ that are biased downward by λ (i.e., plim $\hat{\rho}_1 = \lambda\rho$). Bound et al. (1994) estimate equation (3) and obtain a value for λ of 0.779 (with standard error 0.041) using least squares.

Many earnings mobility studies in the United States and elsewhere also seek to estimate the following type of conditional model which includes lagged earnings as an explanatory variable:

$$(6)\ \Delta y^{*} \equiv y_{it}^{*} - y_{it-1}^{*} = \rho x + \delta y_{it-1}^{*} + \varepsilon.$$

What the researcher is actually able to estimate is the following:

$$(7)\ \Delta y = \rho_1 x + \delta_1 y_{it-1} + \varepsilon,$$

where plim $\hat{\rho}_1 = \lambda\rho$ and

$$(8)\ \text{plim}\ \hat{\delta}_1 = \frac{\delta Var(y^*_{it-1})}{Var(y^*_{it-1}) + (1/\lambda^2)Var(w_{it-1})}.^4$$

See the appendix for the derivation of (8). It is easy to see that certain types of measurement error will cause biased estimates of micro mobility.

Gottschalk and Huynh (2006) derive the analytical link between mean-reverting measurement error and two measures of macro mobility - the elasticity of log earnings with respect to lagged earnings and the correlation between current log earnings and lagged log earnings - and show that the various biases from mean-reverting measurement error act in offsetting directions. Specifically, their decomposition equation is of the form

$$\begin{aligned}(9)\ \hat{\beta}_{yy_{-1}} \quad = \quad & \beta_{yy_{-1}}(1 + \{(\beta_{wy^*} - \beta_{w_{-1}y^*_{-1}})\frac{var(y^*_{-1})}{var(y_{-1})}\}) + \\ & \{(\beta_{wy^*} - \beta_{w\varepsilon})\frac{var(\varepsilon)}{\beta_{yy_{-1}}var(y_{-1})}\} + \{[\beta_{ww_{-1}} + \beta_{\varepsilon w_{-1}} - \beta]\frac{var(w_{-1})}{var(y_{-1})}\},\end{aligned}$$

where $\beta_{yy_{-1}}$ is the slope coefficient from a regression of log earnings on lagged log earnings

$$(10)\ y_{it} = \beta_{yy_{-1}} y_{it-1} + \varepsilon,$$

and the measurement error in log earnings and lagged log earnings takes the following textbook model form:

$$(11)\ y_{it} = y^*_{it} + w_{it} \text{ and}$$

$$(12)\ y_{it-1} = y^*_{it-1} + w_{it-1}.$$

Using the SIPP-SSA linked data, which is what we also use, Gottschalk and Huynh find that the mean-reverting measurement error in SIPP earnings almost completely offsets the bias of classical measurement error, resulting in very similar macro mobility estimates using survey-based and administrative-based earnings.

Next, let us turn to the previous evidence on the comparison of univariate profile results with multivariate regression results. Previous work on mobility in several other countries, namely Argentina, Mexico, Venezuela, Indonesia, Spain, and South Africa, found that the effect of certain variables on mobility was reversed when moving from univariate mobility profiles to multivariate mobility regressions using survey-based earnings (Fields et al. 2003 i, ii, and iii, Fields et al. 2005). Therefore, one might expect to find that for the U.S., the

[4] We estimate equation (7) in the empirical work where x includes dummies for gender, race, age, and education.

signs of some variables may change in the univariate versus the multivariate results, at least when using survey-based earnings. As will be shown in Section 5, we do not find this to be the case for either survey-based or administrative-based earnings.

Before concluding this literature review, we would note that previous studies have attempted to correct in other ways for the possible bias introduced into mobility estimates by measurement error. Fields et al. (2003 i, ii, and iii) and Fields et al. (2005) studied income mobility in Indonesia, South Africa, Spain, and Venezuela and in Argentina, Mexico, and Venezuela respectively. They note that the problem of measurement error in the income variable could lead to overstatements of the income gains of the poor relative to the rich. To correct for measurement error in income, they ran earnings change regressions which use period t-1 predicted income in place of period t-1 reported income as an explanatory variable. In some countries, the estimates using predicted income confirm the results obtained when using reported income, while in others statistically significant results using initial reported earnings become insignificant when predicted earnings are used instead. Antman and McKenzie (2005) also attempted to correct for the possible measurement error bias in mobility estimates when studying earnings mobility in Mexico using the Encuesta Nacional de Empleo Urbano (ENEU). They used a pseudo-panel approach to obtain a consistent estimate of macro mobility, which they measured by the slope coefficient from a regression of cohort-specific mean current earnings on cohort-specific mean lagged earnings. Our work does not employ any of these methods, but rather seeks to compare the mobility estimates when using administrative-based versus survey-based earnings in an attempt to gauge the possible measurement error bias in the latter.

In summary, our review of the literature has found scattered evidence concerning how much estimates of macro and micro mobility may be affected by measurement error. Therefore, the results presented below are more complete than the existing literature in the sense of using a larger and nationally representative sample, including more macro mobility concepts and measures of them, presenting earnings mobility profiles, and estimating multivariate earnings mobility functions comparing administrative-based and survey-based earnings mobility estimates.

3 Data Description

In this research, we use a new dataset called the Survey of Income and Program Participation-Social Security Administration Public Use File (SIPP-SSA PUF), Version 3.1, which was created by the Longitudinal Employer Household Dynamics (LEHD) program at the U.S. Census Bureau. The dataset contains individually reported total annual labor earnings from the SIPP linked by Social Security Number (SSN) to employer reported total annual labor earnings subject to income tax from the Social Security Administration's Detailed Earnings Record (DER). The SIPP-SSA PUF actually contains two files, one person-level file and one person-job-level file.

The SIPP-SSA person-level file contains five different stacked SIPP panels (1990, 1991, 1992, 1993, and 1996). The 1990 and 1991 panels are two years long (e.g., the 1990 panel includes earnings data for 1990 and 1991), the 1992 and 1993 panels are three years long, and the 1996 panel is four years long. For this research, the three-year and four-year panels are divided into two-year-long panels for each set of two consecutive years from 1992-1993 through 1998-1999. Stacked together, these panels include a total of 353,120 individuals. However, each individual only has reported SIPP earnings for the years covered by the particular panel in which s/he was interviewed. The dataset also includes several key variables reported on the SIPP survey (race, age, gender, marital status, etc) and a flag variable indicating whether the individual has a validated social security number (SSN) and was thus able to be matched to his/her record in the SSA data. The method for validating SSNs for these five SIPP panels was as follows: If a SIPP respondent refused to provide an SSN, then no attempt was made to obtain a match for that person in the administrative data. For respondents who provided an SSN, a clerk used their name, address, and personal information to look them up in the SSA master file of all applications for Social Security cards (called the Numident file). If the Numident SSN matched the self-reported SSN, then the record was labeled as having a validated SSN. In cases where the Numident SSN was different from the self-reported SSN, the clerk filled in the correct SSN from the Numident file and the record was labeled as having a validated SSN.

For those individuals who do have a validated SSN, the person-level dataset includes annual earnings subject to FICA as reported on the Social Security Administration's Summary Earnings Record (SER), which are capped at the FICA taxable maximum, and the annual detailed earnings records (DER) as reported in the Social Security Administration's Master Earnings File, which are taken directly from Box 1 on the W-2 form and are not capped. The person-job-level dataset includes job-level detailed earnings records (DER) for each worker-employer combination for every year from 1978 through 2003. These job-level earnings may be summed across employers to obtain total annual DER earnings for each individual. If an individual does not have a validated SSN, then his/her SSA annual earnings (both SER and DER) are imputed using a multiple imputation technique for nonresponse in surveys. We exclude these individuals from our sample.

All of the individuals with validated SSNs have non-missing SER and DER earnings. However, some of these individuals have missing SIPP data. All SIPP data that were originally missing were completed using multiple imputation methods originally proposed by Rubin (1993) and updated by Raghunathan et al (2003). This imputation resulted in eight completed datasets which each contain the "true" underlying microdata where they were available (or non-missing) and imputed missing data. These eight completed datasets are analyzed by first analyzing each completed dataset separately and then combining results (such as regression coefficients) using formulas presented in Rubin (1987). Because we are using multiply completed data, we believe that our mobility estimates do not suffer from attrition or self-selection biases. We have essentially replaced one type of problem (sample attrition and item non-response) with another (the quality of the imputed values). This paper is part of a larger Census Bureau project to assess the analytic validity of the multiply completed SIPP-SSA file.

The SIPP interviews respondents at four-month intervals and collects earnings information for each of the previous four months. The annual earnings measure used in this study was created by first imputing earnings at a monthly level, and then summing earnings across twelve months. It should be noted that because the SIPP annual earnings measure aggregates twelve reports of monthly earnings, it could have very different measurement error properties from those of PSID or CPS earnings reports, which are directly for annual earnings in the preceding calendar year and could involve reference by respondents to their tax returns.

To create our final sample, we first chose the set of individuals aged 25-60 with validated SSNs who were dual labor force participants in both years for each set of consecutive years. An individual was defined as a labor force participant if he or she either a) had positive SIPP earnings for the year, b) had positive DER earnings for the year, or c) reported in the SIPP that s/he was actively looking for work during at least one month of that year. We next trimmed the sample using the following method; see Chen and Dixon (1972) and Yale and Forsythe (1976) for details and usage. We fit a mixed effect model for year-specific SIPP earnings with fixed personal characteristics and random person and employer effects using only SIPP earnings data that were within five standard deviations of the year-specific SIPP earnings mean. Then we created a residual for every observation, including those not used to fit the model. We repeated this process using DER earnings. Using the residual variances from these two models, we dropped year-individual observations with either the SIPP residual or the DER residual (or both) greater than five residual standard deviations. Finally, we stacked all the years for 1990-1999. This resulted in a final sample size of 229,578 person-year observations. Some individuals appear more than once in our sample (for example, if they were dual labor force participants in 1996-1997 and in 1997-1998). Note that all earnings variables are expressed as real earnings in January 1995.

It is probable that the set of individuals who have a validated social security number differs systematically from the set of individuals who do not. We feel that the advantage of having actual (as opposed to imputed) DER earnings that is gained by using only individuals with a validated social security number far outweighs the disadvantage of having a sample that is only representative of the population of individuals with validated social security numbers, as opposed to the entire population. Therefore we use only those individuals with validated social security numbers and claim that our sample is representative of the civilian non-institutionalized U.S. population of individuals with validated SSNs. One way to test this claim is to see whether the percentage of people with validated SSNs is the same for key personal characteristics as it is for the whole sample. Appendix Table A1 shows the percentage of observations who have validated SSNs broken into groups by demographic variables and other key variables in the data. For most groups, the percentage of observations with validated SSNs is close to 84%, which is the percentage of observations in the whole sample with validated SSNs. For a few groups (Hispanic, never married, and born outside the U.S.) the percentage is slightly smaller (around 75%). Because we are including only those individuals with validated SSNs, our sample probably includes fewer illegal immigrants than a representative sample would. The fact that there are fewer persons born outside the U.S. in our sample than in the whole sample supports this view. Appendix Table A2 shows the means and variances of several key variables for both the entire sample and for our sample. For no variable do we reject the hypothesis that the means are equal for the two samples. Thus, Tables A1 and A2 together provide evidence that the set of individuals with validated Social Security numbers is for the most part representative of all individuals in the sample.

All of the SIPP panels are stratified multistage probability samples rather than simple random samples. The results presented in this paper take into account the SIPP sampling error resulting from this multistage sampling design by clustering on the primary sampling unit, which is the first-stage cluster in the SIPP sampling design.

4 Empirical Methodologies for Macro Mobility and Micro Mobility Estimates

4.1 Methods Used to Analyze Both Macro and Micro Mobility

We define "true" latent earnings as the earnings obtained from the labor market, exclusive of other compensation such as benefits. "True" earnings include pre-tax employee contributions to deferred compensation plans, such as 401(k) retirement plans, and pre-tax employee-paid health insurance plan premiums. "True" earnings do not include any type of benefits, such as employer contributions to health insurance plans and deferred compensation plans, Medical Savings Accounts, educational assistance above a certain monetary level, fringe benefits, etc.

We have several reasons to believe that the DER earnings measure is as close to "true" latent earnings as it is possible to get, and we will assume in this study that it is completely free of measurement error. First, the DER earnings measure is not capped at the FICA taxable maximum amount as is the SER earnings measure used in many previous earnings validation studies. Second, we are able to distinguish between self-employment DER earnings and employer DER earnings in the job-level dataset. This study will use only those jobs that represent wage and salary earnings and will exclude self-employment income. Hence, summing the DER earnings measure across jobs for each individual provides a measure of total employer-reported annual labor earnings from all jobs. This measure is directly comparable to the SIPP measured of annual labor earnings constructed by summing twelve monthly values of wage and salary earnings reported by the SIPP respondent.

There are several circumstances under which DER earnings may not equal "true" earnings. The first arises when an employee underreports tips and other earnings to the employer. We would prefer to drop all occupations which are likely to have large portions of their earnings in the form of tips, but the occupation variable available on the SIPP-SSA public use file is too coarse for this, with only five categories. Second, there are two items which may be reported under "gross earnings" on an employee's pay stub and which we include in our definition of "true" earnings, but which are not included in Box 1 on the W-2 form: pre-tax health insurance plan premiums and pre-tax contributions to deferred compensation plans, such as 401(k) retirement plans. Health insurance plan premiums are not likely to be missing from the DER earnings measure in a way that varies systematically with any of our explanatory variables, and hence will not bias our macro or micro mobility estimates. Pre-tax contributions to deferred compensation plans are reported elsewhere on the W-2 form (for example in Box 13 in 1999) and we add them to Box 1 to obtain gross earnings. Thirdly, DER earnings can include the following items, all of which employers are required to report as part of taxable income: employer contributions to health insurance plans, Medical Savings Accounts, educational assistance above a certain monetary level, certain types of fringe benefits, etc.

DER earnings may differ from SIPP reported earnings in the following circumstances, even though these differences are not a result of measurement error in either SIPP or DER earnings. First, SIPP respondents are only asked to report earnings on at most two jobs in any given month. If the respondent held more than two jobs in that month, then the DER annual earnings measure will include earnings from all employers for that month, while the SIPP annual earnings measure will not include earnings from the additional jobs. Second, annual SIPP earnings are topcoded (at $150,000 for the 1996 panel and at $100,000 for the earlier panels) while DER earnings are not. However, the individuals affected by this topcoding are not included in our sample as a result of the trimming described above in section 3.

For a number of reasons - because the DER earnings are not capped, because we are not including self-employment income, because we can add pre-tax contributions to deferred compensation plans onto Box 1 earnings, and because health insurance plan premiums missing from DER earnings are not likely to be correlated with other variables in the dataset - we believe that the DER earnings measure is as close to "true" earnings as it is possible to get. Therefore, we will assume in this study that DER earnings are without measurement error (i.e., they are equal to "true" earnings). We will compare the answers to macro and micro mobility questions using both SIPP and DER earnings to gauge the possible effect of measurement error in survey-based earnings on mobility estimates.

We use dollar earnings, rather than log earnings, in all of our main estimations. The reason for this is as follows. We are particularly interested in whether or not the finding of convergent mobility holds in the administrative data. If we find convergence using dollar earnings, it means that the lowest initial earners gained more in dollars than the highest initial earners. This result (called strong convergence) implies that the lowest initial earners gained more in percentage terms than the highest initial earners (called weak convergence). In other words, a finding of strong convergence using dollar earnings implies a finding of weak convergence using log earnings.

4.2 Macro Mobility Methodology

Macro mobility asks the question: how much earnings mobility was there in the United States during the 1990s? Many papers, including Hungerford (1993), Gittleman and Joyce (1995, 1996), Sawhill and Condon (1992), Burkhauser, Holtz-Eakin, and Rhody (1997), Buchinsky and Hunt (1996), and Gottschalk and Huynh (2006), gauge just one or two mobility concepts, which vary from study to study. However, Buchinsky et al. (2003) and Fields (2004) examine all six of the concepts of mobility that have been used in the literature. As stated earlier, the six concepts are: time dependence, which measures the degree to which individuals' earnings in one year are determined by their earnings in the previous year; positional movement, which is measured by observing individuals' changes in economic positions in earnings distributions (either ranks, centiles, deciles, or quintiles);

share movement, which happens when individuals' shares of total earnings in the population change; earnings flux, which concerns the size of changes in individual's earnings levels but not their sign; directional earnings movement, which measures how many people move up or down the earnings distribution and by how much; and mobility as an equalizer of longer-term earnings, which compares the inequality of earnings at a point in time with the inequality of earnings over a longer time period (Fields 2004). For the United States from 1970-1995 and France from 1967-1999, the studies mentioned above find that the answers to macro mobility questions depend dramatically on which mobility concept the researcher chooses to measure. For France, five of six mobility concepts showed that mobility had fallen over time, but the sixth did not. For the U.S., four of six mobility concepts showed that mobility first rose and then fell back to its previous level, while the remaining two concepts showed that mobility rose, fell, and then rose again over time.

Given that measures of the different mobility concepts have been shown to produce different time patterns, we too use all six different concepts of mobility to answer the above question concerning the extent of mobility in the U.S. in the 1990s. In this paper, we choose two measures of time-independence: one minus the coefficient from a regression of current earnings on earnings in the previous year, and the minus chi-squared statistic from a quintile transition matrix for earnings.[5] Each of the remaining five concepts is measured using a single measure: per-capita centile movement to gauge positional movement, the mean absolute value of share changes to gauge share movement, average absolute value of change in earnings to measure earnings flux, average change in earnings to measure directional movement, and Fields' equalization index to measure mobility as an equalizer of longer-term income. Their specific definitions appear in Table 1.

[5] We use the negative of the chi-squared statistic so that a more positive number represents more mobility, as it does for the other mobility measures.

Table 1	
Measures of Six Mobility Concepts Used in the Empirical Work	
Mobility Concept	**Measure of that Concept Used in this Research**
Time-Independence	$\chi^2 = \sum_i \sum_j \frac{(OBS_{ij}-EXP_{ij})^2}{EXP_{ij}}$, where OBS_{ij} is the number of individuals observed in a particular cell of a quintile transition matrix and EXP_{ij} is the number that would be expected in that cell if initial earnings and final earnings are statistically independent.
Time-Independence	$1-\beta_{yy_{-1}}$, the coefficient from a regression of current earnings on earnings in the previous year.
Positional Movement	$(1/n)\sum \lvert c(y_{2i})-c(y_{1i})\rvert$, where $c(.)$ denotes i's centile in the initial or final year earnings distribution.
Per-Capita Share Movement	$(1/n)\sum \lvert s(y_{2i})-s(y_{1i})\rvert$, where $s(.)$ denotes i's share of total earnings in the initial or final year.
Per-Capita Earnings Flux	$(1/n)\sum \lvert y_{2i}-y_{1i}\rvert$.
Per-Capita Directional Movement	$(1/n)\sum(y_{2i}-y_{1i})$.
Mobility as an Equalizer of Longer-Term Earnings	$E \equiv 1-(I(a)/I(y_1))$, where a is the vector of average earnings, y_1 is the vector of base-year earnings, and $I(.)$ is an inequality measure (either the Gini coefficient or the Theil index).

We do not think that one concept or measure of macro mobility is necessarily more important than another for understanding the amount of mobility taking place in a country over time. There is no single "best" measure of macro mobility. Each concept measures something quite different, and it is important to look at all of them to gain a more complete understanding of how much mobility there is in any given year and how the amount of mobility has changed over time.

For each of the six mobility concepts, we calculate mobility from one year to the next for the relevant individuals from 1990-1991 through 1998-1999. (Note: it is not possible to calculate mobility between 1995 and 1996 because none of the SIPP panels interviewed individuals in both of those years; the last full year of interviews for the 1993 panel was 1995 and the first full year of interviews for the 1996 panel was 1996).[6]

4.3 Micro Mobility Methodology

Micro mobility focuses on mobility of the individual and answers the question: which individuals moved up/down in the earnings distribution over time and by how much? As stated in Section 4.1, we exclude self-employment income from our analysis and instead look only at changes in wage and salary earnings.[7] To begin answering this question, we first present a mobility profile which shows the mean and standard deviation of one-year earnings changes for different subgroups of individuals. We present these statistics for individuals broken down by initial earnings quintile, gender, age, race, and education. We then use multivariate regression models to study the correlates of earnings changes while holding other variables constant. The regression model we focus on in this study specifies earnings changes as a function of initial earnings in steps and a linear function of gender, race, age, and education. We estimate equation (7) where yit-1 is lagged (or initial) earnings broken into five dummy variables for earnings quintile and x includes dummies for gender, race, age, and education. We do not interpret this as a causal model of earnings changes, but rather a way of answering the question of which individuals experience the most positive earnings changes, holding other things equal.

5 Results

Overall, the tables presented below produce two major results. First, qualitatively, measurement error in survey-based earnings has little effect on macro and micro mobility estimates. Second, quantitatively, measurement error in survey-based earnings often has a large effect on mobility estimates, although in some cases, the effects are minor. Moreover, measurement error does not affect mobility estimates in a systematic direction. In other words, the administrative-based results are neither systematically larger nor systematically smaller than the survey-based ones. We also find that the same groups have better earnings changes, both

[6] Actually, twelve months of SIPP data were collected for only two of the four rotation groups in the year 1996. One month (Jan.) for rotation group 3 and two months (Jan. and Feb.) for rotation group 4 were treated as missing data and were multiply imputed.

[7] It would have been possible to look also at individuals' changes in positions or shares, but we have not done so.

in the univariate micro mobility profile results and in the multivariate regression results. We will now discuss in turn the results for macro mobility rates, micro mobility profiles, micro mobility regressions, comparisons of profiles and regressions, and robustness tests.

5.1 Macro Mobility Results

Qualitatively, measurement error in survey-based earnings does have some effect on macro mobility estimates. Table 2 shows mobility estimates for six different mobility concepts, with two measures calculated for time-independence and one measure calculated for each of the other five concepts. Note that the measures of positional movement, share movement, and earnings flux are positive by definition and that the chi-squared statistic for time-independence is negative by definition, so there will be no qualitative (sign) differences for these measures. The remaining three measures of macro mobility can be either positive or negative. Table 2 shows that for the sixth concept (mobility as an equalizer of longer-term earnings), the administrative-based estimate is positive while the survey-based estimate is negative. That is, mobility had an equalizing effect on earnings in the administrative-based data, but a disequalizing effect on earnings in the survey-based data. There are no other qualitative differences when using administrative-based versus survey-based earnings.

Quantitatively, we see that the administrative-based results are neither systematically larger nor systematically smaller than the survey-based ones. In other words, measurement error does not affect macro mobility estimates in a systematic direction. Administrative-based estimates of macro mobility are smaller than survey-based estimates for four out of six mobility concepts (time independence, positional movement, share movement, and earnings flux) For these four concepts, administrative-based estimates are on average 64% of survey-based estimates. For a fifth concept (directional earnings movement), the administrative-based estimate is 54 times as large as the survey-based estimate.

5.2 Micro Mobility Profile Results

Qualitatively, we find that measurement error in survey-based earnings has little effect on micro mobility profiles. Table 3 shows the means and standard deviations of one-year earnings changes for fifteen groups within five different categories: initial earnings quintile (5 groups), gender (2 groups), race (2 groups), age (3 groups), and education (3 groups). The following qualitative results arise in both the administrative-based data and the survey-based data: 1) The hypothesis that mean earnings changes are equal for different groups within categories (for example, for the two racial groups within the category "race") is rejected at the 1% significance level for all five categories. 2) One might expect that it is always the most advantaged individuals who do better, perhaps as a result of human capital accumulation and the theory of comparative advantage. On the contrary, we find that neither the most-advantaged nor the least-advantaged workers (in terms of initial average earnings) experience the most positive earnings changes. The more advantaged do better in the case of race (non-blacks) and education (the better-educated). The less advantaged do better in the case of initial earnings (the lowest earnings groups) and age (the young). Using both earnings measures, we find convergent mobility, i.e., those in the lowest initial earnings quintile experience the most positive earnings changes while those in the highest initial quintile experience the least positive (or most negative) earnings changes. 3) Mean earnings changes are monotonically decreasing by initial earnings quintiles using both data sources. We find only one qualitative difference between administrative-based data and survey-based data: men do better on average than women in the administrative-based data, while women do better on average than men in the survey-based data. Overall, the micro mobility profile results agree qualitatively across the two data sets for four of the five categories (initial earnings quintile, race, age, and education) and disagree qualitatively for one (gender).

Quantitatively, we find that measurement error in survey-based earnings has a large effect on mean earnings changes and on the inequality of earnings changes. The mean earnings change for a particular population group (such as blacks) is defined as the average earnings change for that group. The inequality of earnings changes for a particular demographic category (such as "race") is defined as the standard deviation of mean earnings changes across groups (e.g., blacks and non-blacks) within that demographic category.

Table 3 shows that for 13 out of 15 groups, we reject the hypothesis that mean earnings changes are equal when using administrative-based earnings versus survey-based earnings. Regarding magnitudes, administrative-based estimates of mean earnings changes are more positive than survey-based estimates for 12 out of 15 groups (the exceptions are the lowest three quintiles). On average, the administrative-based mean earnings changes are 766 dollars greater than the survey-based mean earnings changes.

Table 4 compares the inequality of mean earnings changes across groups within five categories when using administrative-based versus survey-based earnings. We use the standard deviation of mean earnings changes across groups within each category to measure inequality of earnings changes for that category. We find that the inequality of earnings changes is neither systematically larger nor systematically smaller in one data set than in the other. Specifically, the inequality of earnings changes within the initial quintile, race, and age categories is smaller in the administrative-based data than in the survey-based data, while the inequality of earnings changes within the gender and education categories is larger in the administrative-based data than in the survey-based data.

Tables 5 and 6 repeat the analysis in tables 3 and 4 using log earnings changes. When measuring the inequality of earnings changes across groups within each category, using log earnings allows the earnings changes across groups (such as black and non-blacks) to be compared in the same percentage terms. Table 6 shows that the inequality of earnings changes is neither systematically larger nor systematically smaller in one data set than in the other. Specifically, the inequality of earnings changes within the gender and race categories is smaller in the administrative-based data than in the survey-based data, while the inequality of earnings changes within the initial quintile, age, and education categories is larger in the administrative-based data than in the survey-based data.

Overall, Tables 3 through 6 provide evidence that measurement error in survey-based earnings often has a large quantitative effect on micro mobility profile estimates, though not in any systematic direction.

5.3 Micro Mobility Regression Results

Qualitatively, we find that measurement error in survey-based earnings has no effect on micro mobility regressions. Table 7 presents a regression which specifies earnings changes as a function of initial earnings, age, and education entered in steps and gender and race entered as dummies. There are no qualitative differences between using survey-based earnings and administrative-based earnings in this multiple regression. All 11 regression coefficients have the same sign using the two different earnings measures. Using both survey-based and administrative-based earnings, we find that other things equal, individuals in the lowest earnings quintiles do better than those in higher earnings quintiles, men do better than women, non-blacks do better than blacks, the youngest workers do better than the oldest workers, and more educated workers do better than less educated workers.

Quantitatively, though, we find that measurement error in survey-based earnings has a large effect on micro mobility regressions. For most of the regression variables, we reject the hypothesis that the two sets of coefficients are equal. Regarding magnitudes, administrative-based estimates are smaller (in absolute value) than survey-based estimates for 8 out of 10 regression variables (the exceptions are the two age dummies, which are not statistically significantly different from each other). On average, the administrative-based estimates are 64% of the survey-based estimates.

In summary, we have found for the regressions that a) qualitatively, measurement error in survey-based earnings has no effect on mobility estimates, b) quantitatively, measurement error in survey-based earnings has a large effect on mobility estimates, but not in a systematic direction.

We turn next to comparing the univariate mobility profile results with the multivariate regression results.

5.4 Comparing Univariate and Multivariate Results

Previous work on mobility in other countries (Fields et al. 2003 i, ii, and iii, Fields et al. 2005) found that the effect of certain variables on mobility was reversed when moving from univariate mobility profiles to multivariate mobility regressions. This is not the case for the United States. For four out of five categories (initial earnings quintile, race, age, and education), the univariate results are qualitatively the same as the multivariate regression results. For both the survey-based earnings data and the administrative-based earnings data, we find that both unconditionally and when holding other things equal, individuals in the lowest earnings quintiles do better than those in higher earnings quintiles, non-blacks do better than blacks, the youngest workers do better than the oldest workers, and more educated workers do better than less educated workers. However, the univariate results by gender are mixed (men do better than women in the administrative-based data while women do better than men in the survey-based data), but the regression results show that, other things equal, men do better than women in both the administrative-based and the survey-based data.

One result is particularly noteworthy. The U.S. Census Bureau reports constant earnings inequality in the U.S. for the early part of the 1990s and again in the later 1990s (U.S. Census Bureau, 2005).[8] However, from 1992 to 1993, earnings inequality jumped by three Gini points, the very same time when new methods were used to collect earnings data (U.S. Census Bureau, 2004). Though it is impossible to tell whether using the old methods would have produced constant or rising earnings inequality, there is no evidence whatsoever suggesting that earnings inequality fell in the United States over the period of our analysis; the Census Bureau evidence suggests that inequality either rose or remained constant. The 1990s was also a period of growth for the U.S.: real GDP per capita rose from $28,000 to $34,000 (Johnston and Williamson, 2006). The combination of growth with constant or rising inequality might lead one to expect that persons in the most advantaged groups would always be the ones who experienced the most positive earnings changes in dollars from one year to the next. However, this is not what we find. The groups who were the most advantaged to begin with were the most-educated, men, non-blacks, the non-young, and (of course) those in the highest initial earnings quintile. Using both survey-based

[8] These earnings inequality estimates were produced by the U.S. Census Bureau using cross-sectional data from the Current Population Survey, Annual Social and Economic Supplement (formerly known as the March Supplement), rather than from the SIPP panels.

and administrative-based earnings data and estimating both conditional and unconditional models for both data sets, we find that those in the lowest initial earnings quintile (the least advantaged group in terms of initial earnings) and the young (the least advantaged group in terms of age) experienced the most positive earnings changes while those in the highest quintile (the most advantaged group in terms of initial earnings) and the non-young (the most advantaged group in terms of age) experienced the least positive earnings changes. Thus, for initial earnings quintile and age, mobility in the U.S. was convergent, not divergent, in the 1990s (i.e., those who were initially least advantaged did the best and those who were initially most advantaged did the worst).

5.5 Robustness Checks

To check the robustness of our core results, we ran several checks. First, we restricted the sample to only the set of workers with positive survey-based earnings and positive administrative-based earnings in both years. Tables 8 through 11 repeat the analysis of Tables 2, 3, 4 and 7 using only dual positive earners, rather than dual labor force participants. The major results are essentially unchanged. Qualitatively, we find that measurement error in survey-based earnings has little effect on mobility estimates. Quantitatively, we find that the average effect of measurement error on mobility estimates is similar in magnitude to the average effect when using dual labor force participants. Once again, we find that measurement error in survey-based earnings often has a large quantitative effect on mobility estimates, but not in a systematic direction.

For four out of six macro mobility concepts, administrative-based estimates are on average 67% of survey-based estimates, which compares with an average of 64% using dual labor force participants. For the other two mobility concepts, administrative-based estimates are larger than survey-based estimates. For micro mobility, we again find that for both earnings measures and in both the univariate and the multivariate analysis, the less advantaged do better in terms of initial earnings and age and the most advantaged do better in terms of race and education. However, the results by gender are again mixed: administrative-based data show that unconditionally, men do better on average than women while survey-based data show the opposite, but the regression results show that, other things equal, men do better than women in both the administrative-based and the survey-based data. Concerning magnitudes, we find that the administrative-based mean earnings changes are on average 783 dollars greater than the survey-based mean earnings changes in the mobility profile, but that administrative-based coefficients are on average 92% of the survey-based coefficients in the mobility regression. The corresponding numbers using dual labor force participants were that the administrative-based mean earnings changes were on average 766 dollars greater than the survey-based mean earnings changes, and the administrative-based regression coefficients were on average 64% of the survey-based coefficients. In summary, the results using dual positive earners agree both qualitatively and quantitatively with those using dual labor force participants.

Second, we tried several different specifications for the multivariate model: entering initial earnings using different functional forms, checking the signs of demographic variables with initial earnings excluded, and estimating the model for each race/gender group separately. Our key results regarding the effect of measurement error on survey-based mobility estimates are unchanged. We find that qualitatively, measurement error in survey-based earnings has little effect on mobility estimates. Quantitatively, we find yet again that measurement error in survey-based earnings does not affect mobility estimates in a systematic direction. The average quantitative effects of measurement error are of similar magnitudes in each new regression model that includes initial earnings as they were in the base model. Table 12 shows a mobility regression model with initial earnings entered linearly, rather than in steps by quintiles. Our major qualitative result is unchanged: all of the regression coefficients are statistically significant and have the same signs in both data sources. Concerning magnitudes, on average, administrative-based coefficients are 62% of survey-based coefficients in absolute value. (In the base model, the corresponding number was 64%). Table 13 shows a mobility model with initial earnings entered as a spline function by initial earnings quintile. Our core qualitative result is again the same: in both data sets, we find that the least advantaged do better in terms of initial earnings and age, while the most advantaged do better in terms of gender, race, and education. Quantitatively, administrative-based coefficients are on average 94% of survey-based coefficients. Table 14 shows a model of earnings changes as a function of only demographic variables (gender, race, age, and education). We find that not all the regression coefficients are statistically significant, but where they are significant, all of the regression variables retain the same sign as in our core results: men do better than women, non-blacks do better than blacks, the young do better than the old, and the more educated do better than the less educated. Finally, Tables 15 through 18 show our main micro mobility regression specification (from Table 7) estimated separately for each race/gender group. Here, not all the regression coefficients are always statistically significant, but when they are significant, the regression variables follow the same patterns as before for all four race/gender groups. Quantitatively, administrative-based coefficients range on average from 68% of survey-based coefficients to 1.2 times the survey-based coefficients. (In the base model, the corresponding number was 64%). We find evidence of convergent mobility in every race/gender group: that is, the individuals in the highest initial quintile experienced smaller earnings changes than the individuals in the lowest initial quintile.

Third, because there is some speculation on the validity of multiply imputed data, we ran all of our analyses using only those individuals for whom all 24 months of SIPP earnings were available (i.e., non-imputed) for each set of two consecutive years. All of our major results hold using this sample of non-imputed earnings data. We do not include these results here, but they may be obtained from the authors on request.

In summary, our main results are robust to using dual positive earners rather than dual labor force participants and to using only the set of individuals with non-imputed SIPP earnings. Furthermore, all of our multiple regression results are robust qualitatively and quantitatively to entering initial earnings using different functional forms, excluding initial earnings, and estimating the model separately for each race/gender group.

5.6 On the Compatibility Between the Mobility Results and Inequality Patterns

We have found evidence of convergent mobility in every micro mobility profile and every micro mobility regression using both administrative-based and survey-based earnings for the United States in the 1990s. We also know that earnings inequality in the United States was either constant or rising and that real GDP per capita was rising over this same period of time. Before concluding, we wish to remark on how the two sets of results can be reconciled.

Table 19 presents the calculations of mean earnings by anonymous quintiles using our data. (The anonymous quintiles treat the initial year earnings and the final year earnings as variables from two different cross-sections.) The combination of growth with constant or rising inequality might lead one to expect that the anonymous persons in the most advantaged groups (such as the highest earnings quintile) would be the ones who experienced the most positive earnings changes in dollars from one year to the next. We see that when treating our data as a cross-section rather than a panel, we find exactly this: using both earnings measures, the mean earnings of the highest quintile rose the most while the mean earnings of the lowest quintile rose the least (or fell the most). However, we know from the results above that when we employ the panel aspect of the data to look at mean earnings changes for named individuals whom we follow over time, it is those in the lowest quintile who experienced the most positive earnings changes while those in the highest quintile experienced the least positive earnings changes.

Two things were happening at the same time. One is that the dollar differences between different percentiles of the earnings distribution were widening. The other is that the places in the different parts of the earnings distribution were being occupied by different individuals. This finding highlights the importance of conducting mobility studies alongside inequality studies for obtaining a more accurate picture of what individuals actually experienced during a given time period.

6 Conclusion

In this study, we have shown that for the U.S., measurement error in survey-based earnings makes some difference to mobility estimates, both qualitatively and quantitatively, but not a huge one. Most of the results obtained hold when administrative-based earnings are used instead of survey-based earnings. In particular: 1) Of the six macro mobility concepts studied, four are of similar magnitude for the two sets of data. 2) Regarding the micro mobility profiles, for four of the categories (initial earnings, race, age, and education), those groups that are found to be more mobile in one data set are also found to be more mobile in the other. 3) For the micro mobility regressions, all of the variables had the same sign and were statistically significant in the two data sets. 4) We find evidence of convergent mobility (high-income people gaining less in dollars than low-income people) using both data sources, both unconditionally and conditionally. However, there are a small number of differences between the survey-based and administrative-based results: 1) Two of the macro mobility measures produced different results: a) The average earnings change was much larger using administrative data than survey data, and b) The mobility that took place equalized longer-term earnings relative to initial earnings using one data set but disequalized using the other. 2) Survey-based data show that unconditionally, women did better on average than men, while administrative-based data show the opposite. Stated differently, the gender gap of average earnings decreased in the 1990s according to survey-based earnings, while the administrative-based earnings show that the gender gap increased during this period. 3) There are often large differences between administrative-based estimates and survey-based estimates, but the administrative-based results are neither systematically larger nor systematically smaller than the survey-based ones.

Some of our findings might be considered unexpected. First, it might have been expected that the income category with the best (worst) earnings changes would also be the education category with the best (worst) earnings changes. Therefore, given that individuals in the highest initial earnings quintile did the worst, it might be expected that the individuals in the highest education category did the worst. This is not what we find, though. Instead, we find that individuals in the highest education category experienced the most positive earnings changes. Second, one might expect the unconditional mobility profile results to differ from the conditional mobility regression results, since this has been found to be true in other countries. This is not the case for the U.S., though. Rather, we find in the administrative data that those groups of individuals who do best in the univariate profile results (non-blacks, men, the young, and the best educated) also do best when holding other things equal in the regression results.

As we see it, analysts can go on doing research using survey data when survey data are all that is available, but should be aware that the results one obtains from survey data are not necessarily the results one would obtain if one had access to administrative data. Furthermore, because we cannot conclude anything from our work on the U.S. about possible measurement error in surveys from other countries, it would also be worthwhile for this kind of validation study to be conducted for other countries that have matched survey-administrative earnings records.

7 References

1. Abowd, J. M. and M. H. Stinson. 2005. "Estimating Measurement Error in SIPP Annual Job Earnings: A Comparison of Census Survey and SSA Administrative Data." LEHD Technical Paper, U.S. Census Bureau. <http://instruct1.cit.cornell.edu/~jma7 /abowd-stinson-SOLE-2003.pdf>.

2. Abowd, J. M. and S. D. Woodcock. 2001. "Disclosure Limitation in Longitudinal Linked Data" in P. Doyle, J. Lane, L. Zayatz, and J. Theeuwes, eds., Confidentiality, Disclosure, and Data Access: Theory and Practical Applications for Statistical Agencies, Elsevier Science B.V., Amsterdam: 215-277.

3. Antman, F. and D. J. McKenzie. 2005. "Earnings Mobility and Measurement Error: A Pseudo-Panel Approach." World Bank, Policy Research Working Paper No. 3745.

4. Atkinson, A. B., F. Bourguignon, and C. Morrisson. 1992. Empirical Studies of Earnings Mobility, Harwood Academic Publishers.

5. Baulch, B. and J. Hoddinott. 2000. "Economic Mobility and Poverty Dynamics in Developing Countries." *Journal of Development Studies* 36: 1-24.

6. Bigard, A., Y. Guillotin, and C. Lucifora. 1998. "Earnings Mobility: An International Comparison of Italy and France." *Review of Income and Wealth* 44: 535-554.

7. Bound, J., C. Brown, G. J. Duncan, and W. L. Rodgers. 1994. "Evidence on the Validity of Cross-Sectional and Longitudinal Labor Market Data." *Journal of Labor Economics* 12: 345-368.

8. Bound, J., C. Brown, and N. Mathiowetz. 2001. "Measurement Error in Survey Data" in J. J. Jeckman and E. Leamer, eds., Handbook of Econometrics, Chapter 59, Elsevier Science B.V., Amsterdam: 3707-3745.

9. Bound, J. and A. Krueger. 1991. "The Extent of Measurement Error in Longitudinal Earnings Data: Do Two Wrongs Make a Right?" *Journal of Labor Economics* 9 (1): 1-24.

10. Buchinsky, M., D. Fougere, and F. Kramarz. 1998. "La mobilité salariale en France: 1967-1987." *Revue Economique* 49: 879-890.

11. Buchinsky, M., G. S. Fields, D. Fougere, and F. Kramarz. 2003. "Francs or Ranks? Earnings Mobility in France, 1967-1999." CEPR Discussion Papers 3937.

12. Buchinsky, M. and J. Hunt. 1996. "Wage mobility in the United States." *Review of Economics and Statistics* 81 (3): 351-368.

13. Burkhauser, R. V., D. Holtz-Eakin, and S. E. Rhody. 1997. "Labor Earnings Mobility and Inequality in the United States and Germany During the Growth Years of the 1980s." *International Economic Review* 38 (4): 775-794.

14. Chronic Poverty Research Centre. 2004. "The Chronic Poverty Report 2004-05." Manchester: Chronic Poverty Research Centre.

15. Deaton, A. 1997. The Analysis of Household Surveys. Baltimore: Johns Hopkins.

16. Duncan, G. J. and D. H. Hill. 1985. "An Investigation of the Extent and Consequences of Measurement Error in Labor-Economic Survey Data" *Journal of Labor Economics* 3(4):508-532.

17. Fields, G. S. 2001. Distribution and Development: A New Look at the Developing World. Russell Sage Foundation, New York.

18. Fields, G. S. 2004. "Economic and Social Mobility Really Are Multifaceted." Working Paper, Cornell University.

19. Fields, G. S. and E. A. Ok. 1999. "Measuring Movement of Incomes." *Economica* 66 (264): 455-472.

20. Fields, G.S., M. L. Sanchez Puerta, R. D. Hernandez, and S. Freije. 2005. "Earnings Mobility in Argentina, Mexico, and Venezuela: Testing the Divergence of Earnings and the Symmetry of Mobility Hypotheses." Working Paper, Cornell University.

21. Fields, G. S., P. L. Cichello, S. Freije, M. Menendez, and D. Newhouse. 2003(i). "Escaping From Poverty: Household Income Dynamics in Indonesia, South Africa, Spain, and Venezuela" in G. S. Fields and G. Pfeffermann, eds., Pathways Out of Poverty: Private Firms and Economic Mobility in Developing Countries, World Bank, Washington DC.

22. Fields, G. S., P. L. Cichello, S. Freije, M. Menendez, and D. Newhouse. 2003(ii). "Household Income Dynamics: A Four Country Story." *Journal of Development Studies* 40 (2): 30-54.

23. Fields, G. S., P. L. Cichello, S. Freije, M. Menendez, and D. Newhouse. 2003(iii). "For Richer or For Poorer? Evidence from Indonesia, South Africa, Spain, and Venezuela." *Journal of Economic Inequality* 1: 67-99.

24. Gittleman, M. and M. Joyce. 1995. "Earnings Mobility in the United States, 1967-91." *Monthly Labor Review* 118: 3-13.

25. Gittleman, M. and M. Joyce. 1996. "Earnings Mobility and Long-Run Inequality: An Analysis Using Matched CPS Data." *Industrial Relations* 35 (2): 180-196.

26. Gottschalk, P. and M. Huynh. 2006. "Impact of Non-Classical Measurement Error on Measures of Earnings Inequality and Mobility." Working Paper, Boston College and Social Security Administration.

27. Gottschalk, P. and R. Moffitt. 1994. "The Growth of Earnings Instability in the U.S. Labor Market." Brookings Papers on Economic Activity 2: 217-272.

28. Hisnanick, J. J. and K. G. Walker. 2004. "Dynamics of Economic Well-Being: Movements in the U.S. Income Distribution, 1996-1999." Current Population Reports, U.S. Census Bureau. <http://www.census. gov/prod/2004pubs/p70-95.pdf>.

29. Hungerford, T. L. 1993. "U.S. Income Mobility in the Seventies and Eighties." *Review of Income and Wealth* 39: 403-417.

30. Johnston, L. D. and S. H. Williamson. "The Annual Real and Nominal GDP for the United States, 1970 - Present." Economic History Services, April 1, 2006.
<http://www.eh.net/hmit/gdp/>.

31. Kim, B. and G. Solon. 2005. "Implications of Mean-Reverting Measurement Error for Longitudinal Studies of Wages and Employment." *Review of Economics and Statistics* 87 (1): 193-196.

32. Pishke, J. 1995. "Measurement Error and Earnings Dynamics: Some Estimates From PSID Validation Study." *Journal of Business and Economic Statistics* 13 (3): 305-314.

33. Rubin, D. B. 1987. Multiple Imputation for Nonresponse in Surveys. Wiley, New York.

34. Rubin, D. B. 1993. "Satisfying Confidentiality Constraints Through the Use of Synthetic Multiply-Imputed Micro-Data." *Journal of Official Statistics* 9 (2): 461-468.

35. Sawhill, I. V. and M. Condon. 1992. "Is U.S. Income Inequality Really Growing?" *Policy Bites* 13 (June).

36. U.S. Census Bureau. "Gini Coefficient for Earnings of Men and Women, Full-Time Year-Round Workers: 1967-1998." May 13, 2004. <http://www.census.gov/hhes/income /incineq /p60204/fig2.html>.

37. U.S. Census Bureau. "Measures of Individual Earnings Inequality for Full-Time, Year-Round Workers by Sex: 1967 to 2001." May 13, 2005. <http://www.census.gov/hhes /www/income/histinc/ie2.html>.

Table 2: One-year Macro Mobility During the Period 1990-1999

Wage and Salary Earnings Only

Mobility concept	**Mobility measure**	**Using survey-based real earnings**	**Using admin-based real earnings**	**Ratio of admin-based to survey-based**
Time Independence	One minus the coefficient from regression of current real earnings on real earnings in the previous year	0.20	0.08	0.40
Time Independence	Minus chi-squared statistic from transition matrix	-1.36	-1.66	0.82
Positional Movement	Per-capita centile movement	11.49	7.2	0.63
Share Movement	Per-capita change in real earnings share	0.32	0.21	0.66
Earnings Flux	Per-capita change in dollar real earnings (absolute value)	8190.15	5563.76	0.68
Directional Earnings Movement	Per-capita change in dollar real earnings	13.74	744.71	54.20
Equalizer of Longer-Term Earnings	Fields' Equalization Index	-0.042	0.064	opposite in sign

Notes: The total sample size is 229578 and corresponds to the set of individuals ages 25 to 60 with validated SSNs who were labor force participants in both years for each set of two consecutive years from 1990-1999. All calculations are weighted to reflect the corresponding Decennial Census population on April 1st, 2000. All calculations are averaged across eight completed datasets using Rubin's (1987) formulas for computing statistics from multiply imputed data. All earnings are expressed as real earnings in January 1995.

Table 3: Micro Mobility Profile for One-year Real Earnings Changes from 1990-1999

Means and standard deviations of one-year real earnings changes

Wage and salary earnings only

	Using Survey-based real earnings	**Using Admin-based real earnings**	**Obs.**	**Test of H_2**	**Admin-based minus Survey-based**
Total sample	13.74	744.71	229578	H_2: **	730.97
	(172.40)	(40.36)			
By Initial Real Earnings Quintile	H_1: **	H_1: **			
Lowest Quintile	2999.34	2490.65	45918	H_2: **	-508.69
	(165.36)	(62.26)			
Quintile 2	1264.02	1257.54	45916	H_2:	-6.48
	(83.85)	(54.02)			
Quintile 3	634.78	554.81	45915	H_2:	-79.97
	(161.49)	(47.06)			
Quintile 4	-181.36	440.87	45916	H_2: **	622.23
	(335.30)	(61.81)			
Highest Quintile	-4364.61	-891.23	45913	H_2: **	3473.38
	(502.56)	(141.80)			
By Gender	H_1: **	H_1: **			
Men	-20.61	850.11	119061	H_2: **	870.72
	(151.43)	(55.94)			
Women	50.88	630.71	110517	H_2: **	579.83
	(224.12)	(40.00)			
By Race	H_1: **	H_1: **			
Black	-1102.99	636.20	24404	H_2: **	1739.19
	(299.06)	(62.94)			
Non black	149.71	757.95	205174	H_2: **	608.24
	(203.93)	(42.97)			
By Age	H_1: **	H_1: **			
25-36 yrs	800.33	1308.88	94236	H_2: **	508.55
	(213.01)	(56.94)			
37-48 yrs	-169.18	762.28	86765	H_2: **	931.46
	(191.87)	(53.09)			
49-60 yrs	-1117.02	-330.30	48577	H_2: **	786.72
	(135.49)	(68.43)			
By Education	H_1: **	H_1: **			
Primary or less	-141.06	34.33	25642	H_2: **	175.39
	(90.56)	(51.76)			
Secondary	-42.08	491.49	141727	H_2: **	533.57
	(117.70)	(38.13)			
Higher	196.78	1458.88	62209	H_2: **	1262.10
	(396.83)	(79.58)			
			Average ratio:		**766.42**

Notes: The total sample size of 229578 corresponds to the set of individuals ages 25 to 60 with validated SSNs who were labor force participants in both years for each set of two consecutive years from 1990-1999. All calculations are weighted to reflect the corresponding Decennial Census population on April 1st, 2000. All means and variances are averaged across eight completed datasets using Rubin's (1987) formulas for computing statistics from multiply imputed data. All earnings are expressed as real earnings in January 1995. **Hypothesis 1**, equality of means within categories, is rejected at the 1% significance level for all five categories (initial real earnings quintile, gender, race, age, education) using both earnings measures. **Hypothesis 2**: Means are equal when using survey-based versus administrative-based real earnings. * H_2 rejected at 5% significance level; ** H_2 rejected at 1% significance level.

Table 4: Inequality of Mean Real Earnings Changes Across Groups Within Categories

Wage and Salary Earnings Only

	Survey-based	**Admin-based**	**Ratio of Admin-based to Survey-based**
Initial quintile	2452.35	1105.98	0.45
Gender	35.72	109.62	3.07
Race	386.12	37.53	0.10
Age	732.91	612.61	0.84
Education	116.96	483.89	4.14

Notes: The inequality measures reported are weighted standard deviations of mean real earnings changes across groups within each category. These numbers are calculated from Table 3. Example calculation: for initial quintile using survey-based real earnings, 2452.35 is the weighted standard deviation (weighted by sample sizes) of the following five numbers from Table 3: 2999.34, 1264.02, 634.78, -181.36, -4364.61. This is a measure of the inequality of mean real earnings changes across groups (quintiles) within that category.

Table 5: Micro Mobility Profile for One-year Log Real Earnings Changes from 1990-1999

Means and standard deviations of one-year log real earnings changes

Wage and salary earnings only

	Using Survey-based log real earnings	**Using Admin-based log real earnings**	**Obs.**	**Test of H_2**	**Admin-based minus Survey-based**
Total sample	-0.01	0.03	229578	H_2: **	0.04
By Initial Log Real Earnings Quintile	H_1: **	H_1: **			
Lowest Quintile	0.35	0.61	45918	H_2: **	0.26
Quintile 2	-0.03	-0.03	45916	H_2:	0.00
Quintile 3	-0.06	-0.04	45915	H_2:	0.02
Quintile 4	-0.08	-0.03	45916	H_2: **	0.05
Highest Quintile	-0.13	-0.05	45913	H_2: **	0.08
By Gender	H_1: **	H_1: **			
Men	-0.02	0.02	119061	H_2: **	0.04
Women	-0.01	0.03	110517	H_2: **	0.04
By Race	H_1: **	H_1: **			
Black	-0.06	0.05	24404	H_2: **	0.11
Non black	-0.01	0.03	205174	H_2: **	0.04
By Age	H_1: **	H_1: **			
25-36 yrs	0.02	0.06	94236	H_2: **	0.04
37-48 yrs	-0.02	0.03	86765	H_2: **	0.05
49-60 yrs	-0.06	-0.04	48577	H_2: **	0.02
By Education	H_1: **	H_1: **			
Primary or less	-0.03	0.03	25642	H_2: **	0.06
Secondary	-0.02	0.02	141727	H_2: **	0.04
Higher	-0.01	0.04	62209	H_2: **	0.05
			Average ratio:		**0.06**

Notes: The total sample size of 229578 corresponds to the set of individuals ages 25 to 60 with validated SSNs who were labor force participants in both years for each set of two consecutive years from 1990-1999. All calculations are weighted to reflect the corresponding Decennial Census population on April 1st, 2000. All means and variances are averaged across eight completed datasets using Rubin's (1987) formulas for computing statistics from multiply imputed data. All earnings are expressed as real earnings in January 1995. **Hypothesis 1**, equality of means within categories, is rejected at the 1% significance level for all five categories (initial log real earnings quintile, gender, race, age, education) using both earnings measures. **Hypothesis 2**: Means are equal when using survey-based versus administrative-based log real earnings. * H_2 rejected at 5% significance level; ** H_2 rejected at 1% significance level.

Table 6: Inequality of Mean Log Real Earnings Changes Across Groups Within Categories

Wage and Salary Earnings Only

	Survey-based	**Admin-based**	**Ratio of Admin-based to Survey-based**
Initial quintile	0.18	0.27	1.53
Gender	0.01	0.01	0.92
Race	0.02	0.01	0.38
Age	0.03	0.03	1.16
Education	0.01	0.01	1.12

Notes: The inequality measures reported are weighted standard deviations of mean log real earnings changes across groups within each category. These numbers are calculated from Table 5. Example calculation: for initial quintile using survey-based log real earnings, 0.18 is the weighted standard deviation (weighted by sample sizes) of the following five numbers from Table 5: 0.35, -0.03, -0.06, -0.08, -0.13. This is a measure of the inequality of mean log real earnings changes across groups (quintiles) within that category.

Table 7: Micro Mobility Multivariate Results for One-Year Real Earnings Changes from 1990-1999

Wage and Salary Earnings Only

Dependent Variable: One-year change in real earnings

* significant at 5%; ** significant at 1%

	Using Survey-based real earnings	**Using Admin-based real earnings**	**Test of H_2**	**Ratio of Admin-based to Survey-based**
Quintile 2	-1979.24**	-1136.02**	H_2: **	0.57
	(200.87)	(68.93)		
Quintile 3	-2926.70**	-2013.93**	H_2: **	0.69
	(288.02)	(71.20)		
Quintile 4	-4193.42**	-2415.27**	H_2: **	0.58
	(460.39)	(84.81)		
Quintile 5	-9052.75**	-4164.70**	H_2: **	0.46
	(624.37)	(159.09)		
Male	1440.84**	808.74**	H_2: **	0.56
	(220.71)	(56.31)		
Black	-1894.55**	-180.71**	H_2: **	0.10
	(369.85)	(69.55)		
Ages 37-48	-260.77*	-269.86**	H_2:	1.03
	(122.31)	(64.96)		
Ages 49-60	-1093.69**	-1323.06**	H_2:	1.21
	(199.56)	(87.56)		
Highschool	1347.94**	702.01**	H_2: **	0.52
	(120.99)	(59.16)		
College	3413.25**	2357.33**	H_2: **	0.69
	(214.44)	(87.43)		
Constant	1723.39**	1617.00**	H_2:	
	(137.23)	(85.43)		
Observations	229578	229578	**Average:**	**0.64**
R-squared	0.04	0.02		
H_1:	**	**		

Notes: Robust standard errors in parentheses. Excluded quintile is quintile 1; excluded age group is 25-36; excluded education category is no high school. The total sample size of 229578 corresponds to the set of individuals ages 25 to 60 with validated SSNs who were labor force participants in both years for each set of two consecutive years from 1990-1999. All calculations are weighted to reflect the corresponding Decennial Census population on April 1st, 2000. All means and variances are averaged across eight completed datasets using Rubin's (1987) formulas for computing statistics from multiply imputed data. All earnings are expressed as real earnings in January 1995. **Hypothesis 1**: equality of coefficients across quintiles. **Hypothesis 2**: equality of coefficients when using survey-based versus administrative-based real earnings. * rejected at 5%; ** rejected at 1%.

Table 8: One-year Macro Mobility During the Period 1990-1999: Dual Positive Earners

Wage and Salary Earnings Only

Mobility concept	**Mobility measure**	**Using survey-based real earnings**	**Using admin-based real earnings**	**Ratio of admin-based to survey-based**
Time Independence	One minus the coefficient from regression of current real earnings on real earnings in the previous year	0.19	0.08	0.42
Time Independence	Minus chi-squared statistic from transition matrix	-1.37	-1.63	0.84
Positional Movement	Per-capita centile movement	11.34	7.51	0.66
Share Movement	Per-capita change in real earnings share	0.31	0.22	0.71
Earnings Flux	Per-capita change in dollar real earnings (absolute value)	8065.08	5828.32	0.72
Directional Earnings Movement	Per-capita change in dollar real earnings	250.79	976.49	3.89
Equalizer of Longer-Term Earnings	Fields' Equalization Index	0.32	0.50	1.56

Notes: The total sample size is 229578 and corresponds to the set of individuals ages 25 to 60 with validated SSNs who had positive real earnings in both years for each set of two consecutive years from 1990-1999. All calculations are weighted to reflect the corresponding Decennial Census population on April 1st, 2000. All calculations are averaged across eight completed datasets using Rubin's (1987) formulas for computing statistics from multiply imputed data. All earnings are expressed as real earnings in January 1995.

Table 9: Micro Mobility Profile for One-year Real Earnings Changes from 1990-1999: Dual Positive Earners

Means and standard deviations of one-year real earnings changes

Wage and salary earnings only

	Using Survey-based real earnings	**Using Admin-based real earnings**	**Obs.**	**Test of H_2**	**Admin-based minus Survey-based**
Total sample	250.79	976.49	186732	H_2: **	725.70
	(253.40)	(46.61)			
By Initial Real Earnings Quintile	H_1: **	H_1: **			
Lowest Quintile	2709.66	2898.19	37347	H_2: *	188.53
	(91.40)	(60.22)			
Quintile 2	1645.78	1203.15	34347	H_2: **	-442.63
	(200.02)	(55.72)			
Quintile 3	962.31	809.44	34346	H_2:	-152.87
	(266.60)	(58.41)			
Quintile 4	291.95	755.06	34346	H_2:	463.11
	(404.95)	(76.13)			
Highest Quintile	-3953.89	-600.74	34346	H_2: **	3353.15
	(504.76)	(157.82)			
By Gender	H_1: **	H_1: **			
Men	183.38	1143.11	97992	H_2: **	959.73
	(220.05)	(62.94)			
Women	326.35	790.74	88740	H_2: **	464.39
	(313.36)	(44.79)			
By Race	H_1: **	H_1: **			
Black	-1122.97	787.70	19395	H_2: **	1910.67
	(311.51)	(76.39)			
Non black	412.33	998.76	167337	H_2: **	586.43
	(288.93)	(48.70)			
By Age	H_1: **	H_1: **			
25-36 yrs	1020.42	1528.92	77920	H_2: **	508.50
	(288.12)	(66.01)			
37-48 yrs	6.09	946.15	71162	H_2: **	940.06
	(265.93)	(57.51)			
49-60 yrs	-802.83	-49.95	37650	H_2: **	752.88
	(216.65)	(72.61)			
By Education	H_1: **	H_1: **			
Primary or less	21.82	422.57	17999	H_2: **	400.75
	(142.38)	(59.82)			
Secondary	175.50	663.83	115525	H_2: **	488.33
	(202.62)	(44.09)			
Higher	481.98	1808.31	53208	H_2: **	1326.33
	(451.19)	(88.63)			
			Average ratio:		**783.16**

Notes: The total sample size of 186732 corresponds to the set of individuals ages 25 to 60 with validated SSNs who had positive real earnings in both years for each set of two consecutive years from 1990-1999. All calculations are weighted to reflect the corresponding Decennial Census population on April 1st, 2000. All means and variances are averaged across eight completed datasets using Rubin's (1987) formulas for computing statistics from multiply imputed data. All earnings are expressed as real earnings in January 1995. **Hypothesis 1**, equality of means within categories, is rejected at the 1% significance level for all five categories (initial real earnings quintile, gender, race, age, education) using both earnings measures. **Hypothesis 2**: Means are equal when using survey-based versus administrative-based real earnings. * H_2 rejected at 5% significance level; ** H_2 rejected at 1% significance level.

Table 10: Inequality of Mean Real Earnings Changes Across Groups Within Categories:
Wage and Salary Earnings Only
Dual Positive Earners

	Survey-based	**Admin-based**	**Ratio of Admin-based to Survey-based**
Initial quintile	2240.26	1140.14	0.51
Gender	248.08	290.35	1.17
Race	525.91	240.89	0.46
Age	746.55	631.99	0.85
Education	280.48	579.99	2.07

Notes: The inequality measures reported are weighted standard deviations of mean real earnings changes across groups within each category. These numbers are calculated from Table 9. Example calculation: for initial quintile using survey-based real earnings, 2240.26 is the weighted standard deviation (weighted by sample sizes) of the following five numbers from Table 9: 2709.66, 1645.78, 962.31, 291.95, -3953.89. This is a measure of the inequality of mean real earnings changes across groups (quintiles) within that category.

Table 11: Micro Mobility Multivariate Results for One-Year Real Earnings Changes from 1990-1999:

Wage and Salary Earnings Only

Dual Positive Earners

Dependent Variable: One-year change in real earnings

* significant at 5%; ** significant at 1%

	Using Survey-based real earnings	**Using Admin-based real earnings**	**Test of H_2**	**Ratio of Admin-based to Survey-based**
Quintile 2	-1352.27**	-2172.76**	H_2: **	1.61
	(111.32)	(98.31)		
Quintile 3	-2405.7**	-3309.84**	H_2: **	1.38
	(191.27)	(88.76)		
Quintile 4	-3639.72**	-3847.55**	H_2:	1.06
	(314.60)	(109.24)		
Quintile 5	-8048.12**	-5525.75**	H_2: **	0.69
	(479.61)	(183.80)		
Male	1176.84**	1171.16**	H_2:	1.00
	(222.15)	(67.97)		
Black	-2058.73**	-277.71**	H_2: **	0.13
	(429.61)	(84.61)		
Ages 37-48	-293.26*	-129.09*	H_2:	0.44
	(125.48)	(71.26)		
Ages 49-60	-781.13**	-992.88**	H_2:	1.27
	(181.92)	(90.40)		
Highschool	1268.81**	914.92**	H_2: **	0.72
	(132.01)	(70.92)		
College	3264.19**	2844.25**	H_2: **	0.87
	(221.08)	(108.77)		
Constant	1909.77**	2743.21**	H_2: **	
	(249.77)	(94.00)		
Observations	186732	186732	**Average:**	**0.92**
R-squared	0.04	0.02		
H_1:	**	**		

Notes: Robust standard errors in parentheses. Excluded quintile is quintile 1; excluded age group is 25-36; excluded education category is no high school. The total sample size of 186732 corresponds to the set of individuals ages 25 to 60 with validated SSNs who had positive real earnings in both years for each set of two consecutive years from 1990-1999. All calculations are weighted to reflect the corresponding Decennial Census population on April 1st, 2000. All means and variances are averaged across eight completed datasets using Rubin's (1987) formulas for computing statistics from multiply imputed data. All earnings are expressed as real earnings in January 1995. **Hypothesis 1**: equality of coefficients across quintiles. **Hypothesis 2**: equality of coefficients when using survey-based versus administrative-based real earnings. * rejected at 5%; ** rejected at 1%.

Table 12: Micro Mobility Multivariate Results for One-Year Real Earnings Changes from 1990-1999:
Linear in Initial Real Earnings
Wage and Salary Earnings Only

Dependent Variable: One-year change in real earnings
* significant at 5%; ** significant at 1%

	Using Survey-based real earnings	Using Admin-based real earnings	Test of H_1	Ratio of Admin-based to Survey-based
Initial Real Earnings	-0.24**	-0.11**	H_1: **	0.46
	(0.01)	(0.01)		
Male	2562.13**	1366.27**	H_1: **	0.53
	(196.18)	(80.63)		
Black	-2196.71**	-323.75**	H_1: **	0.15
	(388.72)	(75.53)		
Ages 37-48	314.77**	41.11**	H_1:	0.13
	(107.30)	(74.02)		
Ages 49-60	-458.02**	-915.86**	H_1:	2.00
	(153.68)	(89.57)		
Highschool	2082.04**	1005.26**	H_1: **	0.48
	(245.23)	(73.14)		
College	5690.99**	3463.21**	H_1: **	0.61
	(533.17)	(134.50)		
Constant	2127.21**	1357.54**	H_1: **	
	(172.93)	(80.68)		
Observations	229578	229578	**Average:**	**0.62**
R-squared	0.11	0.04		

Notes: Robust standard errors in parentheses. Excluded age group is 25-36; excluded education category is no high school. The total sample size of 229578 corresponds to the set of individuals ages 25 to 60 with validated SSNs who were labor force participants in both years for each set of two consecutive years from 1990-1999. All calculations are weighted to reflect the corresponding Decennial Census population on April 1st, 2000. All means and variances are averaged across eight completed datasets using Rubin's (1987) formulas for computing statistics from multiply imputed data. All earnings are expressed as real earnings in January 1995. **Hypothesis 1**: equality of coefficients when using survey-based versus administrative-based real earnings. * rejected at 5%; ** rejected at 1%.

Table 13: Micro Mobility Multivariate Results for One-Year Real Earnings Changes from 1990-1999:
Initial Real Earnings Spline by Quintiles
Wage and Salary Earnings Only

Dependent Variable: One-year change in real earnings
* significant at 5%; ** significant at 1%

	Using Survey-based real earnings	**Using Admin-based real earnings**	**Test of H_3**	**Ratio of Admin-based to Survey-based**
Intercept quintile 1	-983.49**	1889.01**	H_3: **	1.92
	(308.98)	(117.31)		
Intercept quintile 2	-449.30*	877.33**	H_3: **	1.95
	(189.80)	(135.09)		
Intercept quintile 3	-1251.94**	-478.41**	H_3: **	0.38
	(295.03)	(110.35)		
Intercept quintile 4	-2476.49**	-1100.32**	H_3: **	0.44
	(409.23)	(129.70)		
Intercept quintile 5	-259.45	901.61**	H_3:	3.48
	(1249.47)	(252.58)		
Slope quintile 1	-0.45**	0.17**	H_3: **	0.38
	(0.06)	(0.02)		
Slope quintile 2	-0.12**	-0.14**	H_3:	1.17
	(0.03)	(0.02)		
Slope quintile 3	-0.17**	-0.07**	H_3: **	0.41
	(0.03)	(0.02)		
Slope quintile 4	-0.17**	-0.04**	H_3: **	0.24
	(0.03)	(0.01)		
Slope quintile 5	-0.43**	-0.20**	H_3: **	0.47
	(0.04)	(0.01)		
Male	2060.64**	1158.66**	H_3: **	0.56
	(150.14)	(67.19)		
Black	-1891.42**	-288.02**	H_3: **	0.15
	(319.70)	(66.55)		
Ages 37-48	98.65	-74.27	H_3:	0.75
	(106.31)	(69.94)		
Ages 49-60	-633.71**	-995.07**	H_3: **	1.57
	(172.46)	(89.45)		
Highschool	1381.81**	727.91**	H_3: **	0.53
	(121.12)	(60.17)		
College	4584.88**	3044.96**	H_3: **	0.66
	(256.22)	(102.98)		
Observations	229578	229578	**Average:**	**0.94**
R-squared	0.13	0.06		
H_1:	**	**		
H_2:	**	**		

Notes: Robust standard errors in parentheses. Excluded age group is 25-36; excluded education category is no high school. The total sample size of 229578 corresponds to the set of individuals ages 25 to 60 with validated SSNs who were labor force participants in both years for each set of two consecutive years from 1990-1999. All calculations are weighted to reflect the corresponding Decennial Census population on April 1st, 2000. All means and variances are averaged across eight completed datasets using Rubin's (1987) formulas for computing statistics from multiply imputed data. All earnings are expressed as real earnings in January 1995. **Hypothesis1**: equality of intercept coefficients across quintiles. **Hypothesis 2**: equality of slope coefficients across quintiles. **Hypothesis 3**: equality of coefficients when using survey-based versus administrative-based real earnings. * rejected at 5%; ** rejected at 1%.

Table 14: Micro Mobility Multivariate Results for One-Year Real Earnings Changes from 1990-1999:
Exclude Initial Real Earnings
Wage and Salary Earnings Only

Dependent Variable: One-year change in real earnings
* significant at 5%; ** significant at 1%

	Using Survey-based real earnings	**Using Admin-based real earnings**	**Test of H_2**	**Ratio of Admin-based to Survey-based**
Male	-137.73	193.13**	H_2: *	1.40
	(172.18)	(55.15)		
Black	-1318.52**	-32.67	H_2: **	0.02
	(417.25)	(67.12)		
Ages 37-48	-1038.61**	-579.14**	H_2: **	0.56
	(113.77)	(60.65)		
Ages 49-60	-1914.33**	-1658.63**	H_2:	0.87
	(177.66)	(84.58)		
Highschool	-2.44	103.27*	H_2:	42.32
	(149.58)	(57.40)		
College	171.48	1063.57**	H_2:	6.20
	(384.33)	(84.50)		
Constant	984.69**	856.30**	H_2: *	
	(202.87)	(72.31)		
Observations	229578	229578	**Average:**	**8.56**
R-squared	0.003	0.004		
H_1:	**	**		

Notes: Robust standard errors in parentheses. Excluded age group is 25-36; excluded education category is no high school. The total sample size of 229578 corresponds to the set of individuals ages 25 to 60 with validated SSNs who were labor force participants in both years for each set of two consecutive years from 1990-1999. All calculations are weighted to reflect the corresponding Decennial Census population on April 1st, 2000. All means and variances are averaged across eight completed datasets using Rubin's (1987) formulas for computing statistics from multiply imputed data. All earnings are expressed as real earnings in January 1995. **Hypothesis 2**: equality of coefficients when using survey-based versus administrative-based real earnings. * rejected at 5%; ** rejected at 1%.

Table 15: Micro Mobility Multivariate Results for One-Year Real Earnings Changes from 1990-1999: Black Males

Wage and Salary Earnings Only

Dependent Variable: One-year change in real earnings

* significant at 5%; ** significant at 1%

	Using Survey-based real earnings	**Using Admin-based real earnings**	**Test of H_2**	**Ratio of Admin-based to Survey-based**
Quintile 2	-2122.43**	-1258.27**	H_2:	0.59
	(478.18)	(253.84)		
Quintile 3	-2795.23**	-1820.66**	H_2:	0.65
	(513.10)	(274.25)		
Quintile 4	-4716.94**	-2347.4**	H_2: **	0.50
	(719.95)	(308.30)		
Quintile 5	-16305.18**	-4386.67**	H_2: **	0.27
	(2433.09)	(446.03)		
Ages 37-48	455.01	-211.96	H_2:	0.47
	(395.39)	(227.06)		
Ages 49-60	195.56	-1087.71**	H_2: *	5.56
	(570.00)	(370.47)		
Highschool	957.54*	650.44**	H_2:	0.68
	(509.44)	(226.34)		
College	3626.13**	1847.68**	H_2:	0.51
	(1097.04)	(419.97)		
Constant	1591.22**	2004.22**	H_2:	
	(626.23)	(244.05)		
Observations	10571	10571	**Average:**	**1.15**
R-squared	0.11	0.02		
H_1:	**	**		

Notes: Robust standard errors in parentheses. Excluded quintile is quintile 1; excluded age group is 25-36; excluded education category is no high school. The sample size of 10571 corresponds to the set of black, male individuals ages 25 to 60 with validated SSNs who were labor force participants in both years for each set of two consecutive years from 1990-1999. All calculations are weighted to reflect the corresponding Decennial Census population on April 1st, 2000. All means and variances are averaged across eight completed datasets using Rubin's (1987) formulas for computing statistics from multiply imputed data. All earnings are expressed as real earnings in January 1995. **Hypothesis 1**: equality of coefficients across quintiles. **Hypothesis 2**: equality of coefficients when using survey-based versus administrative-based real earnings. * rejected at 5%; ** rejected at 1%.

Table 16: Micro Mobility Multivariate Results for One-Year Real Earnings Changes from 1990-1999: Nonblack Males

Wage and Salary Earnings Only

Dependent Variable: One-year change in real earnings

* significant at 5%; ** significant at 1%

	Using Survey-based real earnings	**Using Admin-based real earnings**	**Test of H_2**	**Ratio of Admin-based to Survey-based**
Quintile 2	-2576.79**	956.16**	H_2: **	0.37
	(319.37)	(146.50)		
Quintile 3	-3604.90**	2099.41**	H_2: **	0.58
	(363.39)	(120.42)		
Quintile 4	-4726.32**	-2407.84**	H_2: **	0.51
	(502.61)	(137.09)		
Quintile 5	-9392.73**	-4020.60**	H_2: **	0.43
	(488.46)	(192.10)		
Ages 37-48	-473.86*	-619.03**	H_2:	1.31
	(256.74)	(109.48)		
Ages 49-60	-1700.03**	-2006.49**	H_2:	1.18
	(439.94)	(157.73)		
Highschool	1892.35**	752.08**	H_2: **	0.40
	(266.36)	(98.23)		
College	3866.82**	2759.83**	H_2: **	0.71
	(255.07)	(152.26)		
Constant	3433.69**	2533.04**	H_2: **	
	(228.65)	(135.75)		
Observations	108490	108490	**Average:**	**0.69**
R-squared	0.04	0.02		
H_1:	**	**		

Notes: Robust standard errors in parentheses. Excluded quintile is quintile 1; excluded age group is 25-36; excluded education category is no high school. The sample size of 108490 corresponds to the set of non-black, male individuals ages 25 to 60 with validated SSNs who were labor force participants in both years for each set of two consecutive years from 1990-1999. All calculations are weighted to reflect the corresponding Decennial Census population on April 1st, 2000. All means and variances are averaged across eight completed datasets using Rubin's (1987) formulas for computing statistics from multiply imputed data. All earnings are expressed as real earnings in January 1995. **Hypothesis 1**: equality of coefficients across quintiles. **Hypothesis 2**: equality of coefficients when using survey-based versus administrative-based real earnings. * rejected at 5%; ** rejected at 1%.

Table 17: Micro Mobility Multivariate Results for One-Year Real Earnings Changes from 1990-1999: Black Females

Wage and Salary Earnings Only

Dependent Variable: One-year change in real earnings

* significant at 5%; ** significant at 1%

	Using Survey-based real earnings	**Using Admin-based real earnings**	**Test of H_2**	**Ratio of Admin-based to Survey-based**
Quintile 2	-2003.10**	-1443.74**	H_2:	0.72
	(372.17)	(161.08)		
Quintile 3	-3171.00**	-2188.36**	H_2:	0.69
	(529.75)	(240.48)		
Quintile 4	-5617.89**	-2720.07**	H_2: **	0.48
	(655.71)	(275.69)		
Quintile 5	-17363.03**	-5550.61**	H_2: **	0.32
	(1947.57)	(545.77)		
Ages 37-48	625.02	253.29	H_2:	0.41
	(411.75)	(210.15)		
Ages 49-60	227.08	-494.61	H_2: *	2.18
	(350.07)	(192.86)		
Highschool	873.39**	579.58**	H_2:	0.66
	(307.14)	(166.42)		
College	4105.23**	2370.99**	H_2: *	0.58
	(868.65)	(414.77)		
Constant	948.03*	1700.13**	H_2:	
	(424.97)	(152.53)		
Observations	13833	13833	**Average:**	**0.76**
R-squared	0.13	0.03		
H_1:	**	**		

Notes: Robust standard errors in parentheses. Excluded quintile is quintile 1; excluded age group is 25-36; excluded education category is no high school. The sample size of 13833 corresponds to the set of black, female individuals ages 25 to 60 with validated SSNs who were labor force participants in both years for each set of two consecutive years from 1990-1999. All calculations are weighted to reflect the corresponding Decennial Census population on April 1st, 2000. All means and variances are averaged across eight completed datasets using Rubin's (1987) formulas for computing statistics from multiply imputed data. All earnings are expressed as real earnings in January 1995. **Hypothesis 1**: equality of coefficients across quintiles. **Hypothesis 2**: equality of coefficients when using survey-based versus administrative-based real earnings. * rejected at 5%; ** rejected at 1%.

Table 18: Micro Mobility Multivariate Results for One-Year Real Earnings Changes from 1990-1999: Nonblack Females

Wage and Salary Earnings Only

Dependent Variable: One-year change in real earnings

* significant at 5%; ** significant at 1%

	Using Survey-based real earnings	**Using Admin-based real earnings**	**Test of H_2**	**Ratio of Admin-based to Survey-based**
Quintile 2	-1624.79**	-1253.79**	H_2: *	0.77
	(168.67)	(75.82)		
Quintile 3	-2431.14**	-1972.652**	H_2: *	0.81
	(233.95)	(73.65)		
Quintile 4	-3642.30**	-2334.54**	H_2: **	0.64
	(376.21)	(103.38)		
Quintile 5	-7309.36**	-4268.36**	H_2:	0.58
	(1831.38)	(291.95)		
Ages 37-48	-180.77	36.91	H_2:	0.20
	(154.00)	(99.54)		
Ages 49-60	-682.12**	-708.32**	H_2:	1.04
	(220.22)	(108.72)		
Highschool	935.32**	681.82**	H_2:	0.73
	(187.74)	(75.15)		
College	2880.35**	1959.95**	H_2: **	0.68
	(251.63)	(137.77)		
Constant	1440.26**	1474.12**	H_2:	
	(269.79)	(106.78)		
Observations	96684	96684	**Average:**	**0.68**
R-squared	0.04	0.02		
H_1:	**	**		

Notes: Robust standard errors in parentheses. Excluded quintile is quintile 1; excluded age group is 25-36; excluded education category is no high school. The sample size of 96684 corresponds to the set of nonblack, female individuals ages 25 to 60 with validated SSNs who were labor force participants in both years for each set of two consecutive years from 1990-1999. All calculations are weighted to reflect the corresponding Decennial Census population on April 1st, 2000. All means and variances are averaged across eight completed datasets using Rubin's (1987) formulas for computing statistics from multiply imputed data. All earnings are expressed as real earnings in January 1995. **Hypothesis 1**: equality of coefficients across quintiles. **Hypothesis 2**: equality of coefficients when using survey-based versus administrative-based real earnings. * rejected at 5%; ** rejected at 1%.

Table 19: Mean Real Earnings by Anonymous Quintiles
Wage and Salary Earnings Only

	Survey-based real earnings			Admin-based real earnings		
	Initial year	**Final year**	**Final minus initial**	**Initial year**	**Final year**	**Final minus initial**
Lowest Quintile	2861	2458	-403	1298	1297	-1
Quintile 2	11744	11218	-526	10930	11202	272
Quintile 3	20181	19964	-217	20531	21014	483
Quintile 4	30589	30721	132	31593	32334	741
Highest Quintile	57329	58492	1163	60502	62390	1888

Notes: The total sample size is 229578 and corresponds to the set of individuals ages 25 to 60 with validated SSNs who were labor force participants in both years for each set of two consecutive years from 1990-1999. All calculations are weighted to reflect the corresponding Decennial Census population on April 1st, 2000. All means are averaged across eight completed datasets using Rubin's (1987) formulas for computing statistics from multiply imputed data. All earnings are expressed as real earnings in January 1995.

8 Appendix

Derivation of equation (8) for the univariate case: From equation (6),

$$(13)\ \delta = \frac{Cov(\Delta y^*, y^*_{it-1})}{Var(y^*_{it-1})}.$$

From equation (7),

$$(14)\ \hat{\delta}_1 = \frac{Cov(\Delta y, y_{it-1})}{Var(y_{it-1})}$$

$$(15) = \frac{Cov(\lambda \Delta y^* + \Delta w, n_i + \lambda y^*_{it-1} + w_{it-1})}{Var(n_i + \lambda y^*_{it-1} + w_{it-1})} \text{ (plugging in from equations 2 and 3)}$$

$$(16) = \frac{Cov(\lambda \Delta y^*, \lambda y^*_{it-1})}{Var(\lambda y^*_{it-1}) + Var(w_{it-1})} = \frac{\lambda^2 Cov(\Delta y^*, y^*_{it-1})}{\lambda^2 Var(y^*_{it-1}) + Var(w_{it-1})} * \frac{Var(y^*_{it-1})}{Var(y^*_{it-1})}$$

$$(17) = \frac{Cov(\Delta y^*, y^*_{it-1}) V(y^*_{it-1})}{[Var(y^*_{it-1})] * [Var(y^*_{it-1}) + (1/\lambda^2) Var(w_{it-1})]} = \frac{\delta Var(y^*_{it-1})}{Var(y^*_{it-1}) + (1/\lambda^2) Var(w_{it-1})}.$$

Table A1: Representativeness of our sample

Our sample pools the years from 1990 to 1999 and is defined as the set of individuals ages 25-60 who were dual labor force participants for each set of two consecutive years and who have validated social security numbers. This table shows the percentage of observations by category who have validated SSNs out of the entire set of individuals ages 25-60 who were dual labor force participants.

Category	Sample Size	Percentage with validated SSNs	Category	Sample Size	Percentage with validated SSNs
Total	273689	83.98	Received welfare payments	22589	82.70
			Did not receive welfare payments	251100	84.10
Male	143011	83.40			
Female	130678	84.62	Received disability payments	6315	85.53
			Did not receive disability payments	267374	83.95
Black	29979	81.18			
Non black	243710	84.33	Total net worth below $100,000	69275	85.94
			Total net worth at least $100,000	204414	83.32
Hispanic	25585	76.97			
Non Hispanic	248104	84.71	Homeowner	176366	85.75
			Not homeowner	97323	80.74
25-36 years old	110169	82.14			
37-48 years old	102876	85.37	Born in country other than U.S.	29934	74.52
49-60 years old	60644	84.98	Born in U.S.	243755	85.15
By Education			Had a defined contribution pension plan	63455	85.90
Primary or less	30781	83.00			
Secondary	170450	83.33	Did not have a defined contribution pension plan	164641	84.00
Higher	72458	85.93			
Married	174623	85.98	Had a defined benefit pension plan	88490	85.43
Widowed	3834	84.03	Did not have a defined benefit pension plan	139606	83.96
Divorced/Separated	43195	83.78			
Never married	52037	77.46			
			Had health insurance coverage	235359	84.54
Reported job-limiting disability	19643	84.22	Did not have health insurance coverage	37650	80.78
Did not report job-limiting disability	246555	84.12			
By Number of Children					
0	138678	82.04			
1	53495	85.18			
2	52139	86.83			
3	20319	86.73			
4	6308	85.80			
5 or more	2750	81.03			

Notes: The total sample size of 273689 corresponds to the set of individuals ages 25 to 60 who were labor force participants in both years for each set of two consecutive years from 1990-1999. All calculations are averaged across eight completed datasets using Rubin's (1987) formulas for computing statistics from multiply imputed data.

Table A2: Representativeness of our sample

This table shows the means and variances of several key variables for both the entire sample and for our sample. The entire sample pools the years from 1990 to 1999 and includes individuals ages 25-60 who were dual labor force participants for each set of two consecutive years. Our sample further restricts the entire sample to include individuals who have validated social security numbers. H_1: Means are equal for the two samples: ** rejected at 1% level, * rejected at 5% level.

	Our Sample (229578)		Entire Sample (273689)		
Variable	**Mean**	**Std Dev**	**Mean**	**Std Dev**	**Test of H_1**
Male	0.52	0.00	0.52	0.00	H_1:
Black	0.11	0.01	0.11	0.01	H_1:
Hispanic	0.09	0.01	0.10	0.01	H_1:
Age (3 categories)	1.82	0.00	1.81	0.00	H_1:
Educ_3cat	2.16	0.01	2.15	0.01	H_1:
Marital status	1.90	0.01	1.94	0.01	H_1:
Reported job-limiting disability	0.07	0.00	0.07	0.00	H_1:
Number of children	0.97	0.01	0.95	0.01	H_1:
Received welfare payments	0.09	0.00	0.09	0.00	H_1:
Received disability payments	0.02	0.00	0.02	0.00	H_1:
Total net worth	99151.00	2575.48	97461.00	2763.66	H_1:
Homeowner	0.65	0.01	0.63	0.01	H_1:
Born in country other than U.S.	0.11	0.01	0.12	0.01	H_1:
Had a defined contribution pension plan	0.28	0.00	0.28	0.00	H_1:
Had a defined benefit pension plan	0.39	0.00	0.38	0.00	H_1:
Had health insurance coverage	0.86	0.00	0.85	0.00	H_1:
Weeks worked with pay	47.39	0.08	46.17	0.09	H_1:
Weeks worked part time	6.62	0.09	6.41	0.08	H_1:
Total annual work hours	1905.38	11.01	1880.73	11.94	H_1:
Total family income	50646.00	731.80	49756.00	768.30	H_1:
Total personal income	27926.00	690.32	27303.00	726.18	H_1:
Amount of welfare payments	2713.78	102.62	2667.30	109.54	H_1:
Amount of disability payments	3157.62	125.38	3109.95	118.94	H_1:
Total annual SIPP reported real earnings	24976.00	568.57	24309.00	595.99	H_1:
Change in total annual SIPP reported real earnings	13.74	172.40	33.77	165.41	H_1:

Notes: The total sample size of 273689 corresponds to the set of individuals ages 25 to 60 who were labor force participants in both years for each set of two consecutive years from 1990-1999. All calculations are weighted to reflect the corresponding Decennial Census population on April 1st, 2000. All calculations are averaged across eight completed datasets using Rubin's (1987) formulas for computing statistics from multiply imputed data. All earnings are expressed as real earnings in January 1995.