Estimating Measurement Error in SIPP Annual Job Earnings: A Comparison of Census Survey and SSA Administrative Data*

John M. Abowd and Martha H. Stinson†

January 2005

Abstract

We quantify sources of variation in annual job earnings data collected by the Survey of Income and Program Participation (SIPP) to determine how much of the variation is the result of measurement error. Jobs reported in the SIPP are linked to jobs reported in a new administrative database, the Detailed Earnings Records (DER) drawn from the Social Security Administration's Master Earnings File, a universe file of all earnings reported on W-2 tax forms. As a result of the match, each job potentially has two earnings observations per year: survey and administrative. Unlike previous validation studies, both of these earnings measures are viewed as noisy measures of some underlying true amount of annual earnings. While the existence of survey error resulting from respondent mistakes or misinterpretation is widely accepted, the idea that administrative data is also error-prone is new. Possible sources of employer reporting error, employee under-reporting of compensation such as tips, and general differences between how earnings may be reported on tax forms and in surveys, necessitates the discarding of the assumption that administrative data is a "true" measure of the quantity collected by the survey. In addition, errors in matching SIPP and DER jobs, a necessary task in any use of administrative data, also contribute to measurement error in both earnings variables. Exploiting the presence of individuals with multiple jobs and shared employers over time, we estimate an econometric model that includes random person and firm effects as well as a common error component shared by SIPP and SSA earnings. We do not impose ancillary orthogonal design assumptions. Hence, our model is more general than conventional fixed effect estimators for this problem. The estimation equation includes two independent error components that represent the variation unique to each earnings measure. All fixed effects (except the mean), random person and firm effects, and the shared residual are interpreted as components of "true" variation that represent differences in earnings across people, firms and time periods due to underlying economic reasons. The independent error components are interpreted as measurement error. The ratio of true variation to total variation for the SIPP earnings measure is between .85 and .87, indicating that 13-15 percent of the variation in SIPP annual job earnings can be attributed to measurement error. In contrast the ratio of true to total variation for the DER earnings measure is between 0.73 and 0.80. We estimate a model that allows for independent $AR(1)$ processes in all three error terms and find auto-correlation parameters of 0.58 for the common component, .38 for the SIPP measurement error, and 0.62 for the DER measurement error. These relative magnitudes imply that for first-differenced earnings, in contrast to earnings levels, the reliability ratio will be lower for the SIPP than the DER.

1 Introduction and Background

Economists and statisticians have long recognized that survey data are prone to measurement error. Responses to questions about earnings, education levels, and job characteristics are not measured exactly but instead contain some truth and some error. The classical measurement error model as described by Fuller (1987) defines a dependent variable Y_t that is a linear function of a covariate x_t.

$$Y_t = \beta_0 + \beta_1 x_t + e_t$$

However x_t is not observed directly, and instead we see

*We would like to thank Julia Lane, George Jakubson, Donald Kenkel, Simon Woodcock, Kevin McKinney, Kristin Sandusky, Lars Vilhuber, and Marc Roemer for helpful comments and support. 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 Grant SES-9978093 to Cornell University (Cornell Institute for Social and Economic Research), the National Institute on Aging, and the Alfred P. Sloan Foundation. The views expressed herein are attributable only to the author(s) and do not represent the views 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 is preparing to support external researchers' use of these data; please contact U.S. Census Bureau, LEHD Program, Demographic Surveys Division, FOB 3, Room 2138, 4700 Silver Hill Rd., Suitland, MD 20233, USA.

†John M. Abowd is Professor of Industrial and Labor Relations at Cornell University, Distinguished Senior Research Fellow at the U.S. Census Bureau, Research Associate NBER, Research Affiliate CREST/INSEE, and Research Fellow IZA. Martha Stinson is an economist at the U.S. Census Bureau. contacts: john_abowd@cornell.edu, martha.stinson@census.gov

$$X_t = x_t + u_t$$

where x_t is the true value of the covariate and u_t is the measurement error. By assuming that the measurement error, the true values, and the errors are independently distributed as

$$\begin{bmatrix} x_t \\ e_t \\ u_t \end{bmatrix} \sim N \left\{ \begin{bmatrix} \mu_x \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \sigma_{xx} & 0 & 0 \\ 0 & \sigma_{ee} & 0 \\ 0 & 0 & \sigma_{uu} \end{bmatrix} \right\}$$

the joint distribution of Y and X can be written as

$$\begin{aligned} E\{Y, X\} &= (\beta_0 + \beta_1 \mu_x, \mu_x) \\ Var(Y, X) &= \begin{bmatrix} \beta_1^2 \sigma_{xx} + \sigma_{ee} & \beta_1 \sigma_{xx} \\ \beta_1 \sigma_{xx} & \sigma_{xx} + \sigma_{uu} \end{bmatrix} \end{aligned}$$

When Y is regressed on X, the expected value of the estimated regression coefficient $\widehat{\beta_1}$ is attenuated.

$$E\{\widehat{\beta_1}\} = \beta_1 \frac{\sigma_{xx}}{(\sigma_{xx} + \sigma_{uu})}$$

The ratio

$$\kappa_{xx} = \frac{cov(u, X)}{Var(X)} = \frac{\sigma_{xx}}{(\sigma_{xx} + \sigma_{uu})}$$

is often called the reliability ratio and it defines the ratio of $\widehat{\beta_1}$ to β_1. The proportional attenuation bias resulting from measurement error is defined as $\frac{\beta_1 - \widehat{\beta_1}}{\beta_1} = 1 - \kappa_{xx}$.

The bias resulting from measurement error can be exacerbated if one is using first differenced data. As Angrist and Krueger (1999) describe, the reliability ratio for a variable $\triangle X = (X_t - X_{t-1}) = (x_t - x_{t-1}) + (u_t - u_{t-1})$ is equal to

$$\kappa_{\triangle x \triangle x} = \frac{\sigma_{xx}}{\sigma_{xx} + \sigma_{uu}(\frac{(1-\tau)}{(1-\rho)})}$$

where τ is the auto-correlation coefficient of the measurement error and ρ is the auto-correlation coefficient of the true value of earnings. If $\rho > \tau$ then $\frac{(1-\tau)}{(1-\rho)}$ is greater than one and the signal to noise ratio declines. Thus determining the extent to which measurement error persists over time is important in assessing the impact on the estimated coefficient.

If the variance and structure of the measurement error is known, then unbiased or consistent estimators of β_1 can be obtained. Hence those studying measurement error have focused on estimating κ_{xx} and testing whether the assumptions of classical measurement error were violated. Studies that obtain a second report for the mismeasured variable of interest in order to calculate σ_{uu} and σ_{xx} have been termed validation studies. The most common approach is to view this second report as "truth" and calculate the measurement errors directly as $u = X - x$. The properties of these errors can then be investigated and researchers have often concluded that the assumptions of classical measurement errors were violated and that the errors were correlated with the true values, i.e. $\sigma_{xu} \neq 0$. However they acknowledge that their models were driven by the assumption that they obtained a true measure of x. Without this assumption, there would be no way to determine the relationship between the errors and the true values. This assumption is also fundamentally untestable and is justified solely by the authors' knowledge of the quality of the secondary data source.

One of the first earnings validation studies was done by Mellow and Sider (1983) using a special supplement to the January 1977 CPS that obtained name and address information of employers from the survey respondents[1]. Matched pairs with both employer and employee wage reports totaled 3,612. In this data set, employer-reported wages exceeded worker reports by 4.8% on average. In order to test the sensitivity of statistical models to the source of the variables used, the authors estimated four different wage regressions. In the first two wage regressions, respondent-reported variables for union status, industry and occupation were regressed on worker and employer reported wages, respectively. In the second two wage regressions, employer-reported union status, industry and occupation were regressed alternatively on worker and employer reported wages. Returns to education and experience were strikingly constant across these four equations. The nonwhite-white differential was smaller when using employer-reported wages while the female differential was higher. The union wage-premium was smaller when using employer-reports of union coverage. Occupation and industry differentials were very similar across the

[1] Mellow and Sider also evaluate a second matched data set: the Employment Opportunity Pilot Project (EOPP). However this data set contains only general firm data such as industry and union status matched to specific workers and hence it is not possible to compare earnings reports from both the employer and employee using this data set.

different specifications. The authors concluded that the wage regressions were generally not that sensitive to the source of information: worker versus employer.

During the 1980s, a validation study at a large anonymous manufacturing company was undertaken. Results from this study were reported in Duncan and Hill (1985) and Bound et al. (1994). Workers at the company were interviewed using a PSID survey instrument and then information for these workers was obtained from company records. Bound et al. provided a comprehensive report on both waves. The first wave of data was collected in the summer of 1983 and included 418 workers and the second wave was conducted in 1987 with 341 of the originally interviewed workers and an additional 151 new workers. The authors treated the company reports of annual earnings as measures of true earnings values and considered any differences between worker and employer reports to be errors on the part of the workers. According to the authors, "We do this because of our confidence in the accuracy and recording of the company records, in part because of the extraordinary cooperation of the company involved. This is crucial, because if there were significant errors in the company records, one would have no way of knowing how they were correlated with other variables." By their own acknowledgement, the results in this paper were completely driven by this assumption.

The authors reported a noise to total variance ratio ($\frac{\sigma_{uu}}{(\sigma_{xx}+\sigma_{uu})}$ in the notation above) of .302 for annual earnings in 1986 and .151 for annual earnings in 1982. They argued that this ratio was misleading because the errors in earnings were correlated with the true levels of earnings. In this case the true variance ratio should be

$$\frac{cov(X,u)}{var(X)} = \frac{\sigma_{uu} + cov(u,x)}{\sigma_{xx} + \sigma_{uu} + 2cov(u,x)}$$

This ratio was calculated by regressing the errors on the employee-reported annual earnings and was 0.239 in 1986 and 0.076 in 1982. Thus the authors claimed that when earnings measures are used as independent variables in regression analyses, the bias resulting from measurement error will be mitigated by correlation between errors and true values.

Generally measurement error in a dependent variable will not cause bias in the regression coefficients but will make them less precise by increasing the overall variance of Y. However the correlation between the true value and the error of a dependent variable will introduce bias even if the independent variables are measured without error. The authors described this result in the following way:

$$\begin{aligned}
Y &= (1+\delta)y + v = x\beta + \varepsilon \\
\widehat{b} &= \frac{(1+\delta)cov(y,x)}{var(x)} \\
\frac{\widehat{b}}{\beta} &= (1+\delta)
\end{aligned}$$

Thus the proportional attenuation bias in the coefficient is δ which was estimated as -0.172 for 1986 and -0.104 for 1982. Again the calculation of these results was completely dependent on the strategy used to identify the errors separately from the true value of earnings.

The authors concluded by estimating two earnings equations, one using employee reported measures of earnings and tenure and the other using company recorded measures of the same variables. Education and experience were also included in the regressions. Since only one measure of education and experience was available (employee interview responses), these variables were considered measured without error. Regression coefficients from the worker-reported equation were measured against the "true" coefficients from the company-reported equation. According to this standard, the interview data overstated the return to education by 40% and understated the return to tenure by 20%.

Bound and Krueger (1991) conducted a similar validation study using linked CPS-Social Security Earnings Records. March 1978 CPS respondents were asked to report their Social Security Numbers and, using SSN, name, age, sex, and race, respondents were linked to SSA records. About 50% of respondents who were in both the 1977 and 1978 March CPS were successfully linked to SSA data. This study was complicated by the fact that SSA earnings reports were truncated at the maximum Social Security taxable earnings amount ($16,500 in 1977 and $15,300 in 1976). The authors made the same error-identifying assumptions as Bound, *et al.* Administrative records were viewed as truth with the exception that the truth was sometimes truncated. Thus the authors first estimated the relationship between the SSA and CPS earnings using a Tobit maximum likelihood approach which accounted for the truncation. The results from this estimation were used to calculate the variance/covariance matrix between CPS earnings and true SSA earnings. This matrix in turn was used to compute a variance/covariance matrix between x_t and u_t. The authors reported large negative correlations between measurement error and true earnings for both 1976 and 1977 (-0.46 and -0.42, respectively). They reported reliability ratios which did and did not take account of these correlations as 0.844 and 1.016 respectively for 1976 and 0.819 and 0.974 for 1977. They also noted that the reporting errors appeared to be positively correlated over time but "with only 2 years of data it is impossible to distinguish an autoregressive process in the measurement error from a person fixed effect or from other time-series processes."

Bound, Brown, and Mathiowetz (2001) summarized earnings validation studies and stated that the ideal information for correcting measurement error would be to know the joint distribution of all the true and observed variables, i.e. $f(y, x, Y, X)$.

However the authors recognized that information about this joint distribution has often come at the cost of assuming that validation data is truth. They write, "Those collecting validation data usually begin with the intention of obtaining 'true' values against which the errors of survey reports can be assessed; more often than not we end up with the realization that the validation data are also imperfect. While much can still be learned from such data, particularly if one is confident the errors in the validation data are uncorrelated with those in the survey reports, this means replacing one assumption (e.g. errors are uncorrelated with true values) with another (e.g. errors in survey reports uncorrelated with errors in validation data)."

Bound, Brown, and Mathiowetz also expressed the hope that future validation studies would be able to obtain secondary data reports for multiple consecutive years. Past validation studies have been able to create panels of earnings measures for at most two consecutive years. Thus it has been difficult to calculate the correlation of errors over time, an important component to assessing the impact of measurement error on panel data. Due to the high cost of validating panel data, the authors foresee the future of validation studies as being critically enhanced by opportunities to "merge administrative data to existing panel data."

Our research follows in the tradition of validation studies but with four major innovations. First, we use a new linked survey-administrative database. Second, we model the administrative and survey data symmetrically–neither is viewed as a measure of true earnings–allowing us to develop statistical methods that are robust to this assumption. Third, we link earnings records at the job level, allowing us to use information about the identity and characteristics of the employer. Fourth, we link earnings records from up to four consecutive years, which provides valuable new insight into the time series properties of measurement errors.

2 Statistical Models of Measurement Error

The goal of this paper is to estimate levels of measurement error in SIPP survey data using an alternative source of earnings data: Social Security Administration records called Detailed Earnings Records (DER). Unlike past studies, the administrative records are not treated as a measure of true earnings. Instead, both sources of earnings data are treated as noisy measures of some underlying true value of earnings. We estimate measurement error by decomposing both measures of earnings into shared effects and separate effects. The shared effects include the observable effects of general labor force experience and time as well as the unobservable effects of individual and firm heterogeneity. In addition, there is a shared error component that can be thought of as a nested individual-job-time period random effect. This effect is estimable due to the presence of two earnings observations for each year of the panel. It represents "economic" noise, or fluctuations in annual earnings due to unobservable economic factors which influence true earnings as opposed to reported SIPP or DER earnings. The separate effects are then attributed to measurement error, as these effects are presumably due to things that do not influence the underlying true value of earnings. Using the estimated variance components, the reliability ratios of both the SIPP and DER earnings variables can then be calculated as the ratio of true to total variance.

This modeling follows the spirit of Abowd and Card (1989). Using several different long panel data sets, they first-differenced earnings and hours and adjusted for experience. They then examined the variance-covariance matrix of these differences and tested the fit of various structural models, all of which included a measurement error component. This model will rely on random person and firm effects instead of first-differencing and has the advantage of a second source of data to identify the measurement error but the parsing of variance among structural components is similar.

Given this statistical model, the SIPP earnings equation for a given individual i is:

$$\begin{aligned} \ln(SIPPEARN_{ist}) &= \beta_{oSIPP} + \beta_1 Race.Gender + \beta_2 Race.Gender.Educ + \beta_3 Race.Gender.Exp_{it} + \\ &\quad \beta_4 Time_{it} + \beta_5 \left[P_{1990}, P_{1991}, P_{1992}, P_{1993}\right] + \theta_{1i} + \theta_{2i} Exp_{it} + \psi_j + \eta_{ist} + \omega_{ist} \end{aligned} \tag{1}$$

and the DER earnings equation for the same individual is identical except for the last component:

$$\begin{aligned} \ln(DEREARN_{ist}) &= \beta_{oDER} + \beta_1 Race.Gender + \beta_2 Race.Gender.Educ + \beta_3 Race.Gender.Exp_{it} + \\ &\quad \beta_4 Time_{it} + \theta_{1i} + \theta_{2i} Exp_{it} + \psi_j + \eta_{ist} + \upsilon_{ist} \end{aligned} \tag{2}$$

where i subscripts the individual, j subscripts the firm, s subscripts the person-firm match or job, and t subscripts the year. The variables are defined as follows:

$$
\begin{aligned}
P_{1990}, P_{1991}, P_{1992}, P_{1993} &= \text{vector of 4 indicator variables specifying the SIPP panel of the individual;} \\
& \quad \text{the 1996 panel is the excluded group} \\
Race.Gender &= \text{full interaction of male and white produces four categories:} \\
& \quad \text{white male, non-white male, white female, non-white female} \\
Educ &= \text{four levels of education fully interacted with race and gender:} \\
& \quad \text{high school degree, some college, college degree, graduate degree} \\
& \quad \text{separately for each demographic group} \\
Exp_{it} &= \text{general labor market experience, fully interacted with race and gender:} \\
& \quad \text{actual exp. calculated using employment history collected in the SIPP;} \\
& \quad \text{experience enters as a piecewise linear spline with nodes at} \\
& \quad \text{2 years, 5 years, 10 years, and 25 years;} \\
& \quad \text{separate effects for each demographic group} \\
Time_{it} &= \text{calendar time, base year is 1990} \\
\text{Person heterogeneity, slope and intercept} &= (\theta_1, \theta_2) \sim N(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, G_1) \\
\text{Firm heterogeneity} &= \psi \sim N(0, G_2) \\
\text{Common error component} &= \eta \sim N(0, G_3) \\
\text{Measurement error, SIPP and DER} &= (\omega, \upsilon) \sim N(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, R) \\
& \quad G_1, G_2, G_3, R \text{ are defined below.}
\end{aligned}
$$

The total number of jobs held by all individuals is N, the total number of individuals is I, the total number of firms employing individuals in the sample is J, the number of covariates included in X is k, and the total number of time periods is 10. The maximum number of time periods a job may be observed depends upon the origin SIPP panel. In the 1990, 1991, and 1993 SIPP panels, there are 2 years of complete earnings data. In the 1992 panel there are 3 years and in the 1996 panel, 4 years. Thus a job may be observed anywhere from 1 to 4 years depending on the tenure of the job and the source panel.

Written in matrix notation, the model is

$$Y = X\beta + Zu + e$$

where Y is an $(N \times 10 \times 2) \times 1$ vector of stacked SIPP and DER earnings, X is an $(N \times 10 \times 2) \times k$ matrix of covariates treated as fixed effects, β is a $k \times 1$ vector of fixed effects coefficients, Z is an $(N \times 10 \times 2) \times (I + J + N \times 10)$ design matrix of the random effects, u is a $(I + J + N \times 10) \times 1$ vector of random effects and e is an $(N \times 10 \times 2) \times 1$ vector of residuals.

The fixed effects represent shifts in the conditional mean of the distribution of SIPP or DER earnings, or both. For example, the race, gender, education, experience, and time effects shift the mean of both earnings distributions, whereas the intercept terms shift only one of the means. The β_{0SIPP} term shifts the mean of the entire SIPP earnings distribution and β_{0DER} shifts the mean of the DER earnings distribution. The vector β_3=[$\beta_{31990}, \beta_{31991}, \beta_{31992}, \beta_{41993}$] captures shifts in the mean of the SIPP panel-specific earnings distributions due to differences in earnings measures across SIPP panels (the 1996 panel is the excluded case).

The random effects capture variation in the data due to individual, firm, or time period heterogeneity that is left after controlling for observed characteristics. In other words, there is variation around the mean earnings due to unobservable characteristics of the person, employer, or time period for every category of individuals defined by the variables treated as fixed (X). The random effects quantify the amount of variance due to the different sources. The random effects vector, u, contains the stacked random effects, $\theta_1 ... \theta_I$, $\psi_1 ... \psi_J$, $\eta_{111990} ... \eta_{IN1999}$. The design matrix of the random effects, Z, contains one column for each individual, one column for each firm, and one column for each individual-firm-time period match. The variance matrices for the random person, firm, and shared error component effects, respectively, can be written as

$$
\begin{aligned}
G_1 &= I_{I\times I} \otimes \begin{bmatrix} \sigma^2_{\theta 1} & \\ \sigma_{\theta 1\theta 2} & \sigma^2_{\theta 2} \end{bmatrix} \\
G_2 &= I_{J\times J} \otimes \sigma^2_{\psi} \\
G_3 &= I_{NxN} \otimes \sigma^2_{\eta} \begin{bmatrix} 1 & \rho & \rho^2 & ... & \rho^9 \\ \rho & 1 & \rho & ... & ... \\ \rho^2 & \rho & ... & ... & \rho^2 \\ ... & ... & ... & 1 & \rho \\ \rho^9 & ... & \rho^2 & \rho & 1 \end{bmatrix}_{10\times 10} \\
\text{where } \sigma^2_{\eta} &= \frac{\sigma^2_{\varsigma}}{(1-\rho^2)}
\end{aligned}
$$

The shared error component is modeled as an $AR(1)$ process where errors are correlated within the same job for a given individual but not across jobs nor across individuals. This effect is identified by the fact that there are two observations for each time period. Individual heterogeneity is modeled with two random effects: a person-specific intercept term, θ_1, and a person-specific growth term, θ_2, that allows individual earnings to grow with experience at different rates.

The error vector, e, contains the stacked measurement error terms, $\omega_{111990}...\omega_{IN1999}$,$\upsilon_{111990}...\upsilon_{IN1999}$. The SIPP and DER errors follow separate $AR(1)$ processes with the covariance between them constrained to be zero. These errors are identified by differences in the SIPP and DER earnings reports for each year. The variance matrix for the residuals can be written as

$$
R = I_{(N\times 1)\times(N\times 1)} \otimes \begin{bmatrix} \sigma^2_{\omega} \begin{bmatrix} 1 & \rho_{sipp} & \rho^2_{sipp} & \cdots & \rho^9_{sipp} \\ \rho_{sipp} & 1 & \rho_{sipp} & \cdots & \cdots \\ \rho^2_{sipp} & \rho_{sipp} & \cdots & \cdots & \rho^2_{sipp} \\ \cdots & \cdots & \cdots & 1 & \rho_{sipp} \\ \rho^9_{sipp} & \cdots & \rho^2_{sipp} & \rho_{sipp} & 1 \end{bmatrix}_{10\times 10} & 0_{(N\times 10)\times(N\times 10)} \\ 0_{(N\times 10)\times(N\times 10)} & \sigma^2_{\upsilon} \begin{bmatrix} 1 & \rho_{der} & \rho^2_{der} & \cdots & \rho^9_{der} \\ \rho_{der} & 1 & \rho_{der} & \cdots & \cdots \\ \rho^2_{der} & \rho_{der} & \cdots & \cdots & \rho^2_{der} \\ \cdots & \cdots & \cdots & 1 & \rho_{der} \\ \rho^9_{der} & \cdots & \rho^2_{der} & \rho_{der} & 1 \end{bmatrix}_{10\times 10} \end{bmatrix}
$$

where ρ_{sipp} and ρ_{der} are the autocorrelation terms of the SIPP (ω) and DER (υ) measurements errors, respectively.

Estimates of β_{0SIPP}, β_{0DER}, β_1, β_2, β_3, β_4, β_5, the variance components ($\sigma^2_{\theta 1}$, $\sigma^2_{\theta 2}$, $\sigma_{\theta 1\theta 2}$, σ^2_{ψ}, σ^2_{η}, ρ, σ^2_{ω}, ρ_{sipp}, σ^2_{υ}, ρ_{der}), and realizations of the random effects (θ_1, θ_2, ψ, η) and the residuals (υ, ω) can be obtained by solving the mixed model equations.

$$
\begin{bmatrix} X'R^{-1}X & X'R^{-1}Z \\ Z'R^{-1}X & Z'R^{-1}Z + G^{-1} \end{bmatrix} \begin{bmatrix} \widehat{\beta} \\ \widehat{u} \end{bmatrix} = \begin{bmatrix} X'R^{-1}Y \\ Z'R^{-1}Y \end{bmatrix}
$$

The estimation is done by restricted maximum likelihood (REML) using an average information (AI) algorithm, developed and programmed by Gilmore, Thompson, and Cullis (1995). This method closely follows the Fisher scoring algorithm proposed by Patterson and Thompson (1971). Parameters are chosen to maximize the log likelihood function by satisfying a set of first order conditions, or score equations. Solutions to the score equations are calculated iteratively. The user furnishes a set of starting values for the variance components and the algorithm calculates the log likelihood and produces initial estimates of the fixed effects ($\beta's$) and the realized random effects. The information matrix is calculated using an averaging method that simplifies the process for large data sets with multiple random effects. The information matrix is then used to update the variance component estimates. The process is repeated until the estimates converge.

After estimates of the fixed effects and variance components have been obtained, the magnitude of the measurement error will be assessed in several ways. First, the reliability ratio commonly used in the literature will be calculated using the following two formulas.

earnings levels (3)

$$
\begin{aligned}
\kappa_{SIPP} &= \frac{\sigma^2_{\eta} + \sigma^2_{\theta 1} + (\overline{Exp^2}) * \sigma^2_{\theta 2} + 2 * (\overline{Exp}) * \sigma_{\theta 1\theta 2} + \sigma^2_{\psi}}{\sigma^2_{\eta} + \sigma^2_{\theta 1} + (\overline{Exp^2}) * \sigma^2_{\theta 2} + 2 * (\overline{Exp}) * \sigma_{\theta 1\theta 2} + \sigma^2_{\psi} + \sigma^2_{\omega}} \\
\kappa_{DER} &= \frac{\sigma^2_{\eta} + \sigma^2_{\theta 1} + (\overline{Exp^2}) * \sigma^2_{\theta 2} + 2 * (\overline{Exp}) * \sigma_{\theta 1\theta 2} + \sigma^2_{\psi}}{\sigma^2_{\eta} + \sigma^2_{\theta 1} + (\overline{Exp^2}) * \sigma^2_{\theta 2} + 2 * (\overline{Exp}) * \sigma_{\theta 1\theta 2} + \sigma^2_{\psi} + \sigma^2_{\upsilon}}
\end{aligned}
$$

$$\text{first-differenced earnings} \quad (4)$$
$$\kappa_{\triangle SIPP} = \frac{\sigma_\eta^2 + \sigma_{\theta_2}^2}{\sigma_\eta^2 + \sigma_{\theta_2}^2 + \sigma_\omega^2(\frac{(1-\rho_{SIPP})}{(1-\rho)})}$$
$$\kappa_{\triangle DER} = \frac{\sigma_\eta^2 + \sigma_{\theta_2}^2}{\sigma_\eta^2 + \sigma_{\theta_2}^2 + \sigma_\upsilon^2(\frac{(1-\rho_{DER})}{(1-\rho)})}$$

These ratios will give an estimate of the proportion of overall variance that is “true" while the remainder will be attributed to measurement error. High reliability ratios indicate good quality data with low amounts of measurement error. These ratios will be calculated for the sample as a whole and also for individual demographic groups in order to determine whether some groups are more affected by measurement error than others.

Next, we modify our base estimation equations in order to investigate the sources of measurement error. For both the fixed effects and the person and firm random effects, we wish to estimate an “average" effect and then a deviation from that average due to measurement error. We accomplish this by interacting source with each of the fixed effects. Thus the estimation equations become

$$\begin{aligned} \ln(SIPPEARN_{ist}) = {} & \beta_{oSIPP} + \beta_{1SIPP}Race.Gender + \beta_{2SIPP}Race.Gender.Educ + \\ & \beta_{3SIPP}.Race.Gender.Exp_{it} + \beta_4 Time_{it} + \beta_5\left[P_{1990}, P_{1991}, P_{1992}, P_{1993}\right] \\ & + (\theta_{1SIPPi} + \theta_{2SIPPi}Exp_{it}) + \psi_{jSIPP} + \eta_{ist} + \omega_{ist} \end{aligned} \quad (5)$$

$$\begin{aligned} \ln(DEREARN_{ist}) = {} & \beta_{oDER} + \beta_{1DER}Race.Gender + \beta_{2DER}Race.Gender.Educ + \\ & \beta_{3DER}Race.Gender.Exp_{it} + \beta_4 Time_{it} \\ & + \theta_{1DERi} + \theta_{2DERi}Exp_{it} + \psi_{jDER} + \eta_{ist} + \upsilon_{ist} \end{aligned} \quad (6)$$

The SIPP and the DER equation now have separate coefficients to describe the effect of race and gender, education, and experience. We calculate the average effects as $\beta_{AVG} = \frac{(\beta_{SIPP}+\beta_{DER})}{2}$ and the SIPP deviations from the averages as $\beta_{dev} = \frac{(\beta_{SIPP}+\beta_{DER})-2\beta_{SIPP}}{2} = \frac{\beta_{DER}-\beta_{SIPP}}{2}$. The deviation for the DER has the opposite sign from the deviation for the SIPP. Because we are focused on determining what is a true economic effect and what is due to data differences between the SIPP and the DER, we report β_{SIPP}, β_{DER}, β_{AVG} and the absolute value of β_{dev}. We interpret β_{AVG} to be capturing true cross-sectional variation found in both the SIPP and DER and β_{dev} to be capturing cross-sectional variation that is unique to either the SIPP or the DER and is the result of measurement error.

For the random effects, we follow a similar approach. We estimate the two equations with separate person and firm effects for the SIPP and the DER and allow these effects to be correlated. The variance/covariance matrix for the thetas becomes

$$G_1 = I_{I\times I} \otimes \begin{bmatrix} \sigma^2_{\theta 1SIPP} & & & \\ \sigma_{\theta 1SIPP\theta 1DER} & \sigma^2_{\theta 1DER} & & \\ \sigma_{\theta 1SIPP\theta 2SIPP} & \sigma_{\theta 1DER\theta 2SIPP} & \sigma^2_{\theta 2SIPP} & \\ \sigma_{\theta 1SIPP\theta 2DER} & \sigma_{\theta 1DER\theta 2DER} & \sigma_{\theta 2SIPP\theta 2DER} & \sigma^2_{\theta 2DER} \end{bmatrix}$$

Again we wish to report the variation associated with an average person effect and the variance associated with deviations from that average. We define $\theta_{1AVGi} = \frac{\theta_{1SIPPi}+\theta_{1DERi}}{2}$ and $\theta_{2AVGi} = \frac{\theta_{2SIPPi}+\theta_{2DERi}}{2}$ as the average between the SIPP and the DER random person effects. We define deviations from these averages as $\theta_{1DEVi} = \frac{\theta_{1DERi}-\theta_{1SIPPi}}{2}$ and $\theta_{2DEVi} = \frac{\theta_{2DERi}-\theta_{2SIPPi}}{2}$. We use the estimated variance components from G_1 to calculate variances and covariances of the average effects and the deviations according to the formulas given below.

$$\begin{aligned} \sigma^2_{\theta_{1AVG}} &= .25*\sigma^2_{\theta_{1SIPP}} + .25*\sigma^2_{\theta_{1DER}} + .5*\sigma_{\theta_{1SIPP}\theta_{1DER}} \\ \sigma^2_{\theta_{2AVG}} &= .25*\sigma^2_{\theta_{2SIPP}} + .25*\sigma^2_{\theta_{2DER}} + .5*\sigma_{\theta_{2SIPP}\theta_{1DER}} \\ \sigma_{\theta_{1AVG}\theta_{2AVG}} &= .25*\sigma_{\theta 1SIPP\theta 2SIPP} + .25*\sigma_{\theta 2SIPP\theta 2DER} + .25*\sigma_{\theta 1DER\theta 2SIPP} + .25*\sigma_{\theta 1SIPP\theta 2DER} \\ \sigma^2_{\theta_{1DEV}} &= .25*\sigma^2_{\theta_{1SIPP}} + .25*\sigma^2_{\theta_{1DER}} - .5*\sigma_{\theta_{1SIPP}\theta_{1DER}} \\ \sigma^2_{\theta_{2DEV}} &= .25*\sigma^2_{\theta_{2SIPP}} + .25*\sigma^2_{\theta_{2DER}} - .5*\sigma_{\theta_{2SIPP}\theta_{2DER}} \\ \sigma_{\theta_{1DEV}\theta_{2DEV}} &= .25*\sigma_{\theta 1SIPP\theta 2SIPP} + .25*\sigma_{\theta 2SIPP\theta 2DER} - .25*\sigma_{\theta 1DER\theta 2SIPP} - .25*\sigma_{\theta 1SIPP\theta 2DER} \end{aligned}$$

For specifications where the person random effects are not interacted with source, we report $\sigma^2_{\theta 1}$, $\sigma^2_{\theta 2}$, and $\sigma_{\theta 1\theta 2}$. For specifications where there are separate person intercepts and slopes for the two data sources, we report $\sigma^2_{\theta_{1AVG}}$, $\sigma^2_{\theta_{1DEV}}$,

$\sigma^2_{\theta_{2AVG}}$, $\sigma^2_{\theta_{2DEV}}$, $\sigma_{\theta_{1AVG}\theta_{2AVG}}$, $\sigma_{\theta_{1DEV}\theta_{2DEV}}$. In all specifications we report the total amount of variance due to unobserved person heterogeneity which we calculate as either $\sigma^2_{\theta 1} + \overline{Exp^2} * \sigma^2_{\theta 2} + 2 * \overline{Exp} * \sigma_{\theta 1 \theta 2}$ or $\sigma^2_{\theta 1AVG} + \overline{Exp^2} * \sigma^2_{\theta 2AVG} + 2 * \overline{Exp} * \sigma_{\theta 1AVG\theta 2AVG}$. For specifications where the person effect is interacted with data source, we interpret this variance to be the total true variance of the person effect and compare this to the total error variance of the person effect which is defined as $\sigma^2_{\theta 1DEV} + \overline{Exp^2} * \sigma^2_{\theta 2DEV} + 2 * \overline{Exp} * \sigma_{\theta 1DEV\theta 2DEV}$.

We interact the random firm effect with source in the same way as the person effect and estimate the variance/covariance matrix for the firm effect as

$$G_2 = I_{J\times J} \otimes \begin{bmatrix} \sigma^2_{\psi SIPP} & \\ \sigma_{\psi SIPP \psi DER} & \sigma^2_{\psi DER} \end{bmatrix}$$

We report $\sigma^2_{\psi_{AVG}} = .25 * \sigma^2_{\psi_{SIPP}} + .25 * \sigma^2_{\psi_{DER}} + .5 * \sigma_{\psi_{SIPP}\psi_{DER}}$ and $\sigma^2_{\psi_{DEV}} = .25 * \sigma^2_{\psi_{SIPP}} + .25 * \sigma^2_{\psi_{DER}} - .5 * \sigma_{\psi_{SIPP}\psi_{DER}}$.

The interaction of the data source indicator with the person and firm effects changes the way the reliability ratio for earnings levels is calculated. There are now three sources of measurement error: deviations from the average person effect, deviations from the average firm effect, and variance captured by the error terms, ω and υ, which essentially represents deviations from the common time period effect, η. In calculating the reliability ratio, the variance of the average person and firm effects is included in the numerator as "true variance" along with the variance of η. The variation associated with deviations from the average person and firm effects is now included in the denominator along with the variance of ω and υ. However these two new sources of measurement error are somewhat different from σ^2_ω and σ^2_υ because they affect both the SIPP and the DER reliability ratios in the same way and hence $\sigma^2_{\theta 1DEV} + \overline{Exp^2} * \sigma^2_{\theta 2DEV} + 2 * \overline{Exp} * \sigma_{\theta 1DEV\theta 2DEV}$ and $\sigma^2_{\psi_{DEV}}$ enter the denominator of both ratios.

$$\begin{aligned} \kappa_{SIPP} &= \frac{\sigma^2_\eta + \sigma^2_{\theta 1AVG} + (\overline{Exp^2}) * \sigma^2_{\theta 2AVG} + 2 * (\overline{Exp}) * \sigma_{\theta 1AVG\theta 2AVG} + \sigma^2_{\psi AVG}}{\begin{array}{c}\sigma^2_\eta + (\sigma^2_{\theta 1AVG} + (\overline{Exp^2}) * \sigma^2_{\theta 2AVG} + 2 * (\overline{Exp}) * \sigma_{\theta 1AVG\theta 2AVG}) + \sigma^2_{\psi AVG} + \\ (\sigma^2_{\theta 1DEV} + (\overline{Exp^2}) * \sigma^2_{\theta 2DEV} + 2 * (\overline{Exp}) * \sigma_{\theta 1DEV\theta 2DEV}) + \sigma^2_{\psi DEV} + \sigma^2_\omega\end{array}} \\ \kappa_{DER} &= \frac{\sigma^2_\eta + \sigma^2_{\theta 1AVG} + (\overline{Exp^2}) * \sigma^2_{\theta 2AVG} + 2 * (\overline{Exp}) * \sigma_{\theta 1AVG\theta 2AVG} + \sigma^2_{\psi AVG}}{\begin{array}{c}\sigma^2_\eta + (\sigma^2_{\theta 1AVG} + (\overline{Exp^2}) * \sigma^2_{\theta 2AVG} + 2 * (\overline{Exp}) * \sigma_{\theta 1AVG\theta 2AVG}) + \sigma^2_{\psi AVG} + \\ (\sigma^2_{\theta 1DEV} + (\overline{Exp^2}) * \sigma^2_{\theta 2DEV} + 2 * (\overline{Exp}) * \sigma_{\theta 1DEV\theta 2DEV}) + \sigma^2_{\psi DEV} + \sigma^2_\upsilon\end{array}} \end{aligned} \tag{7}$$

3 Data Description

The fundamental unit of observation in this paper is a job, defined as a match between an individual and a firm. Data on jobs come from two sources: five Survey of Income and Program Participation (SIPP) Panels conducted during the 1990's and the Detailed Earnings Records (DER) extracted from the Social Security Administration Master Earnings File for the respondents in each of the five panels. In both sources, data on earnings were reported on a sequential, calendar-year basis and job records had to be created by combining earnings records over time that belonged to the same job. Hence appropriately grouping earnings records and defining jobs was the first fundamental difficulty that we addressed in each data source. After job records were created, individuals in each data set were linked by Social Security Number. Finally, jobs for each individual from the two data sources were matched to each other. Each step of this process is described below.

3.1 Creating a SIPP Jobs Data Set

All the SIPP Panels conducted in the 1990s interviewed respondents every 4 months over the course of 2 1/2 to 4 years. Interviewees were divided into 4 rotation groups and one group was interviewed each month. At the time of the interview, retrospective information about the previous 4 months was collected[2]. Respondents were asked to report detailed information

[2] The SIPP panels generally began by interviewing the second rotation group in February of the panel reference year and collecting information about the previous 4 months. Interviews with the third, fourth, and first rotation groups followed and the reference period for collecting information moved forward each time so as to correspond to the preceding four months. For example, in the 1990 panel, the first interview cycle went as follows:

Second rotation group: interviewed in Feb.1990 and answered questions about Oct.1989-Dec.1990.
Third rotation group: interviewed in Mar. 1990 and answered questions about Nov.1989-Feb.1990
Fourth rotation group: interviewed in Apr.1990 and answered questions about Dec.1989-Mar.1990
First rotation group: interviewed in May 1990 and answered questions about Jan.1990-Apr.1990

This cycle continued until Sept. 1992 when the 8th interview with the first rotation group was completed, giving 8 complete waves of data for each rotation group. Thus the 1990 panel contains information on Oct.1989-Aug.1992. The following list gives the number of waves for each SIPP panel and the span of months referenced by the interview questions.

1990 panel: 8 waves, Oct.1989-Aug.1992
1991 panel: 8 waves, Oct.1990-Aug.1993
1992 panel: 10 waves, Oct.1991-Apr.1995

for up to 2 jobs they held during this time period. The employer name, industry, occupation, union status, usual weekly hours, and monthly earnings of each job were recorded, as well as any applicable start and end dates. Each job was also assigned a unique identification number, or job ID, with the intent that this identifier be time-invariant and allow the linking of job information across survey waves. For the first four panels (1990-1993), a paper and pencil survey instrument was used. During each interview, the Census field representative (FR) was required to record an employer name and assign a job ID for each job reported by the respondent, even if the job was a continuation of a job reported in a previous wave. While the FR was supposed to assign the same job ID to a continuing job and a new job ID to a newly begun job, there was no quality check to ensure that this procedure was followed. Beginning in 1996, a major survey redesign was implemented and information was collected using a Computer Assisted Personal Interview (CAPI) system. As a result, as long as the individual did not miss an interview, during the second and subsequent interviews, the CAPI instrument automatically assigned the same employer name and job ID each time further information about a continuing job was collected. When the respondent reported that a new job had started, the CAPI instrument assigned the next available job ID.

Job records were indexed by the longitudinal SIPP person ID, the wave (interview) number, and the job ID. We combined these records to create one observation per job that contained some time invariant job information, such as industry, and some time-varying information such as annual earnings. Table 1 shows the total number of respondents in each SIPP panel, the number that report holding at least one job over the course of the SIPP panel, the total number of person-wave-job records, and the total number of jobs reported, using the assigned SIPP job ID to count jobs. A careful examination of the person-wave-job records revealed serious problems with the SIPP job ID coding process. Because the definition of a "job" was so crucial to our comparison of job earnings from the SIPP and the DER data, we investigated the nature and causes of the job ID coding problems and developed an editing procedure that would resolve some of the inconsistencies we found. This section describes some of the problems we found and gives a summary of how we repaired the job id variables. Details about our edits are contained in Appendix A.

In the 1996 SIPP panel, the largest problem with jobs arose when the jobs had start dates prior to the beginning of the first wave in which they were reported and prior to the beginning of the previously held job. Table 2 gives one generic example of the cause of this problem. In this case, the individual was interviewed in waves 1 through 4 and reported a job which began February 1, 1996. However, the individual missed the fifth interview. When the next interview was conducted in wave 6, a new job was reported but the start date was prior to the beginning of wave 6 and prior to the beginning of job 1. The CAPI system was not designed to allow job IDs to be carried forward through missed interviews. Consequently, when this person temporarily dropped out of the panel, she was automatically given a new job ID at the time of the next interview, regardless of whether the job had actually begun in wave 6 or not. However, there were no restrictions placed on the start date she reported and hence this discrepancy arose. The case illustrated in Table 2 was the most common cause of the early start date problem. However it was not the only cause. The problem affected 21.6% of all jobs (29,520) and about 40% of the time there appears to have been a missing wave problem, while the rest of the time, the cause could not be determined. Whatever the reason, it was clear that the survey job IDs sometimes failed to link job records.

The problems encountered in the early SIPP panels (1990-1993) were considerably more complicated. There were two major types of problems - improper re-use of job IDs and improper assigning of new job IDs. Tables 3A and 3B give generic examples of these problems. In Table 3A, the SIPP respondent held the same job throughout the first four waves of the survey. However, in wave 3, the job ID was incorrectly changed, causing it to appear as if there had been a job transition. This error is identifiable because the name of the employer stays the same across the waves. Table 3B shows the second type of problem. In this case, the person changed jobs between waves 3 and 4 but the job ID was not changed. Thus, it appears that the person remained at the same job through all four waves and, consequently, a job transition was missed. Again, the true work history is apparent only through scrutinizing the employer names.

We developed an editing procedure that used employer name and person-level total job counts from the DER data to identify and correct SIPP job ID coding errors. This procedure involved several phases. First, employer names from job records were compared using statistical name matching software and those names deemed to be the same were grouped together by assigning new job IDs. Second, using these new job IDs, the total number of jobs held over the course of the survey was counted for each individual. A similar count was performed in the administrative DER data. A comparison was made between these two counts to identify cases where the name matching software had failed to correct or had introduced new job history errors. In the 1990-1993 panels, the group of problem observations was large enough to warrant further editing efforts. A second pass with the name matching software was performed and then a series of clerical edits was undertaken. In the 1996 panel, the job count comparison showed that the name matching step had corrected the most obvious problems and further editing was deemed unlikely to provide enough improvements to be worth the resource cost involved. More detail on the name matching procedure and the clerical edits can be found in Appendix A. A detailed description of the full process for the 1990-1993 panels can be found in Stinson (2003)[3]. Tables 4A and 4B provide a summary of the process for

1993 panel: 9 waves, Oct.1992-Dec.1995

1996 panel: 12 waves, Dec.1995-Feb.2000

The cycle of interviews is slightly different for the 1996 panel. The first rotation group was interviewed in April 1996 followed by the second, third, and fourth rotation groups in May, June, and July. Thus the span of referenced months begins in December instead of October and ends in February instead of April, August, or December.

[3]Recognizing that a similar edit could not be performed by researchers lacking access to the DER data, the Census Bureau has publicly released the revised SIPP job ids for the 1990-1993 SIPP panels on the SIPP website, http://www.sipp.census.gov/sipp/access.html.

all the SIPP panels, showing how many unique jobs resulted from each step of the editing procedure and how many records were affected at each step. Row 5 of Table 4A and row 2 of Table 4B show the final number of jobs defined by the revised set of job IDs and row 6 and row 3, respectively, show the number of jobs belonging to people who still have discrepancies between job counts in the DER and the SIPP. We are confident that the majority of these cases are the result of reporting differences between the survey and the administrative data and not failure to link job records in the SIPP.

Once we had defined a set of jobs for each SIPP panel, we created annual earnings measures by summing monthly earnings reports. It is important to understand the concept of earnings as used during the SIPP interview. During the 1990-1993 SIPP panels, respondents were asked about earnings from a specific employer in the following way: "The next question is about the pay ... received from this job during the 4-month period. We need the most accurate figures you can provide. Please remember that certain months contain 5 paydays for workers paid weekly and 3 paydays for workers paid every 2 weeks. Be sure to include any tips, bonuses, overtime pay, or commissions. What was the total amount of pay that ... received BEFORE deductions on this job in ...?" The field representative reads the name of each month and separately records earnings for that month. A special caveat is added for members of the Armed Forces, "Be sure to include cash housing allowances and any other special types of pay." The intent of the survey question was to collect gross earnings and if the person responded that he or she did not know the earnings amounts, the field representative asked if the person could provide the information during a follow-up phone call.

The 1996 survey instrument asked, "Each time he/she was paid by [Name of Employer] in [MonthX], how much did he/she receive BEFORE deductions?" The field representative then followed up with questions about whether there were any other payments such as tips, bonuses, overtime pay, or commissions. The FR was trained to probe several times to make sure all the payments from an employer in a given month were accurately reported. There were also consistency checks built into the CAPI instrument that were meant to spot earnings amounts that seemed unreasonable and provide the FR with the opportunity to make corrections. Respondents were also asked to refer to earnings records if possible so as to give accurate responses. Thus in the best case, these earnings reports most likely reflected the gross pay from monthly pay stubs.

3.2 Creating a DER Jobs Data Set

The second source of data, Detailed Earnings Records (DER) from SSA, contained earnings histories for each SIPP respondent in the 1990, 1991, 1992, 1993, and 1996 panels with a validated SSN (for a definition and discussion of validation see section 3.3: "Matching SIPP and DER Jobs"). These histories included reports of annual earnings, by employer, from 1978-2000. For the purposes of this earnings comparison study, however, only non-self-employment jobs held during the time period covered by the survey questions were used[4]. Employers on this administrative data were identified by an IRS-assigned Employer Identification Number (EIN). Table 5 gives the total number of jobs that appear in the DER for SIPP respondents in all five panels and the total number of unique EINs, followed by the time period covered by the survey and the total number of jobs and unique EINs for this time period.

The earnings data contained in these DER files have as their source the W-2 records filed by employers on behalf of each employee. The primary earnings variable comes from Box 1 of the W-2 Form: wages, tips, and other compensation. This earnings variable in uncapped and represents all earnings that were taxable under federal income tax. There are at least two parts of earnings that would be reported on an employee's pay stub in "gross earnings" that are not included in Box 1: pre-tax health insurance plan premiums paid by the employee and pre-tax elective contributions made to deferred compensation arrangements such as 401(k) retirement plans. In the later case, these contributions are reported elsewhere on the W-2 form (for example Box 13 in 1999) and the DER file contains reports of these deferred earnings which can be added to Box 1 earnings to approximate gross earnings. However, pre-tax health insurance plan premiums are not reported on the W-2 Form and are not contained in the DER. This omission represents one important way in which administrative records may differ from survey records that is not the result of error in the survey data collection process. DER earnings will be lower than SIPP earnings if the respondent reported gross earnings during the survey that included health insurance plan premiums.

There are other possible differences between Box 1 on the W-2 Form and gross earnings reported in the survey, most of which involve some kind of employee benefit that the employee is unlikely to consider wages and may also be unlikely to be reported as such on a pay stub, but which the employer is nonetheless required to report as taxable income. These include educational assistance above a certain monetary level, business expense reimbursement above the amount treated as substantiated by the IRS, payments made by the employer to cover the employee's share of Social Security and Medicare taxes, certain types of fringe benefits such as the use of a company car, golden parachute payments, group-term life insurance over $50,000 paid for by the employer, potentially some portion of employer contributions to Medical Savings Accounts, non-qualified moving expenses, and, in some circumstances, sick pay from an insurance company or other third party payer. In all these cases, DER earnings are likely to be higher than SIPP earnings.

A final potential problem with DER employer reports is that EINs do not necessarily remain constant over time. Unlike Social Security Numbers which serve as good longitudinal identifiers for individuals, EINs can change for many reasons that

[4] The Detailed Earnings Records did contain reports of self-employment earnings. The SIPP also collected information about self-employment, but responses to these questions were treated separately from responses to the questions about jobs with employers. Self-employment reports from either source were not included in this study.

do not necessarily involve a person moving to a new employer. Company reorganizations that consist of mergers, acquisitions, or spin-offs of some parts of the company may result in a worker having two W-2 forms for a tax year, each with a different EIN, without having actually changed jobs. In cases such as these, the DER earnings will be lower than the SIPP earnings because a portion of the earnings for the year are missing. As part of the linking process between DER and SIPP earnings, we attempt to identify these kinds of successor-predecessor problems and merge the two DER jobs determined to be related to a single SIPP job (see Appendix C for details). However, at this early stage of research involving the administrative data, there is no way to know how well our method works.

The EIN linked employers to the Business Register, the master list of all businesses maintained by the Census Bureau that serves as the sampling frame for firm-level surveys. Using this link, we merged information from the Business Register about the industry and name of the employer to each relevant job report in the DER data. Details about this merge can be found in Appendix B. The employer name is the key linking element between the SIPP and DER job data.

3.3 Matching SIPP and DER Jobs

After the creation of the SIPP and DER jobs data sets, the next step was to take people who had job reports in both files and try to match each SIPP job record to a DER job record. Table 6 shows the total number of people and jobs that were potential matches following the job record creation process. Except for the 1996 SIPP jobs, the total number of jobs that were potential matches is the same as row 5 in Table 4A for the SIPP jobs and row 4 of Table 5 for DER jobs. In 1996, one final problem necessitated the dropping of a few additional jobs. Respondents were only allowed to report at most two jobs per interview. In cases where people had a series of short or part-time jobs, interviewers recorded a single job which was labeled as "various employers" or "work arrangement." There were 3,908 job records of this type in the 1996 SIPP data, representing possibly triple that many actual jobs. These jobs were essentially impossible to match to the DER because they do not represent earnings from a single employer. Hence, they were dropped, giving a new total of 121,450 jobs.

We began the job matching process by first linking at the person level. The unique identifier for a person on the DER was the SSN while the SIPP contained a longitudinal person identifier specific to the survey. A crosswalk file matched SSNs and SIPP person IDs. This crosswalk was developed using self-reported SSNs and a validation procedure. Each SIPP respondent was asked to provide an SSN. After this information was collected, SSA searched for each SSN in an administrative data base called the Numident, a universe file containing demographic information collected when every SSN was issued. Self-reported name, sex, race, and date of birth from the SIPP were compared to their administrative counterparts on the Numident. If a respondent's name and demographics were deemed close enough to the name and demographics associated with the SSN in the administrative data base, then the SSN was declared valid. For respondents who answered "do not know" to the SSN question, an attempt was made to find the missing SSN by locating the person in the Numident based on their reported name and demographic characteristics. When a respondent refused to provide an SSN, no attempt was made to link this person to any administrative data and the SSN was left missing. Validated SSNs were included in the crosswalk file and served as the basis for extracting Detailed Earnings Records from the SSA Master Earnings File. Hence in order for an individual to have any earnings reports in the DER, he or she, by necessity, must have a validated SSN.

The third column of Table 6 shows the number of people who matched between the SIPP and the DER. In all panels, some people were lost from both the SIPP and DER job data sets as a result of this match. On the SIPP side, there were two reasons why a person might not match. First, he or she might not have a validated SSN. The third column of Table 7A shows the number of people affected by this problem. The second possibility was that the person had a validated SSN and reported working in the SIPP, but did not have any earnings reports in the DER. This would be caused by the jobs being relatively informal (baby-sitting, yard work, household help) and hence not generating W-2 forms, or by over-reporting of jobs. On the DER side, the only reason for a person not to match was because the person did not report any jobs in the SIPP survey. As seen in the third column of Table 7B, it was far more likely for a person to have jobs in the DER and not the SIPP than the reverse. It would appear that overall, the SIPP undercounts employed people.

As shown in Table 6, even for those people who had employment reports in both the SIPP and the DER, the number of jobs reported was much higher in the administrative data compared to the survey data. At least one factor that influenced the job count on each side was the timing of the survey. In every SIPP panel, the survey asked employment questions of at least some respondents in the last few months of the year preceding the official beginning year. For example in the 1990 panel, the first interview reference period included between one and three months of 1989 for the second, third, and fourth rotation groups. Also the last interviews in the 1990 panel were conducted in September 1992, leaving the last quarter of 1992 uncovered for every rotation group. The 1991-1993 panels followed similar patterns. In the 1996 panel, the first interview reference period included December 1995 for the first rotation group and the last reference period included one or two months of 2000 for the third and fourth rotation groups. In order to attempt to match as many SIPP and DER jobs as possible, all DER jobs from the year preceding the main survey beginning year and jobs from the last survey year were included for the rotation groups that were surveyed at some point during these years. However, some of these DER jobs could clearly have ended before the survey began or started after the survey ended, thus artificially inflating the DER job counts. In the early SIPP panels (1990-1993), jobs that ended before the main survey year or began only in the survey end-year accounted for 27%-32% of all DER jobs. In the 1996 panel, these jobs only accounted for 11% of all DER jobs, largely because the timing of the survey conformed more closely to the calendar year. Another factor which artificially depressed the SIPP job counts is the fact that the survey only collected information covering a maximum of 2 jobs.

After we matched by SSN, a job-to-job match was performed again using the statistical name matching software Integrity. The matching was performed in several steps, called passes. The goal was to first link jobs that were almost certain matches based on the fields deemed to be the most reliable matching indicators and then to link jobs that were less certain matches using other fields. The primary basis for matching was self-reported name of the employer from the SIPP and administrative name of the employer from the Census Business Register. Earnings were not used in the match in order to prevent bias in the subsequent comparison of earnings. Appendix C gives the details of this match including which matching variables were used in each pass and how duplicate matches were handled. The first row of Table 8 gives the number of SIPP jobs that were successfully matched to a counterpart job in the DER.

In summary, in order for a job reported in a SIPP survey to be used in our earnings comparison analysis, a person and job level match must be completed and there must be earnings reports present during one of the full survey years. The process we have described matched all people with validated SSNs who appeared in both files. We then attempted to match all jobs associated with these people. Of the jobs that matched, we then restricted ourselves to comparing earnings only in the full survey years and dropped all jobs that did not have earnings reports in at least one full survey year.

The second row of Table 8 shows how many jobs were dropped because the earnings were out of scope and the third row shows the final total number of jobs per panel that were used in the analysis. At this point jobs from all panels were combined to give a total of 207,323 jobs, 137,361 people, and 115,958 unique employers. Table 9 gives annual summary statistics for earnings and labor force experience (as calculated from the SIPP survey) for all jobs from 1990 to 1999. As is clear from comparing rows 1 and 2 of this table, there were some jobs which matched but did not have the same number of years of reported earnings. For example a SIPP job could have earnings reports for 1996 and 1997 but not 1998 while the DER job could have reports for all three years. It was even possible for a job to be missing earnings in all sample years for the SIPP (or conversely the DER) and have earnings reports from the DER (SIPP). This resulted in slightly different sample sizes between the SIPP and the DER data for each year. Missing values were modeled in the maximization routine as conditionally missing at random and hence the panel was not required to be balanced. The decision not to require exact matching in the earnings years was based on the fact that earnings essentially reported as zero in one source and positive in another source was a type of measurement error that we did not wish to exclude.

Tables 10 and 11 describe the variance-covariance structure of the SIPP and DER earnings over time. The covariances are listed below the diagonal and correlations are listed above. In the SIPP data, the correlations between adjacent years range from .72-.76. In the DER, they are slightly higher, .77-.78. The variance of earnings is also higher in the DER than in the SIPP. Table 12 gives the correlations between each year of DER and SIPP data. The correlations between SIPP and DER earnings in the same year is quite high: .83-.87. The correlations between adjacent years of SIPP and DER data is not as high as between adjacent years of data from the same source, but they nonetheless range from .67-.72. It is also interesting to note that the correlations seem slightly higher in the early 1990s than in the later 1990s.

4 Results

Tables 13-16 present parameters from the estimation of equations described in Section 2. The first column in each table reports the results of estimating equations (1) and (2) while the last column reports results for equations (5), (6). The four columns in between show parameters estimated by making one major change at a time in order to get from the first to the second specification. In column 2, we show results from interacting the fixed effects with a data source indicator but leaving the random effects the same for each data source. In column 3, we interact the person random effects, $\theta_1$and θ_2, with the source indicator but leave the fixed effects and the firm random effect the same. In column 4, we interact the firm random effect and nothing else. In column 5, we interact the person and firm random effects but not the fixed effects. Finally in column 6, we interact the person and firm random effects, as well as the fixed effects.

In Table 13 we present the variances of the random effects. Since mixed models of this type distribute the variation unexplained by the fixed effects across the different factors that are included as random effects and the residuals, it is possible to compare the magnitude of the variance of the random effects and determine which effect explains a higher percentage of the variation. Heterogeneity at the person level produces a total person effect variance of .3-.32 while heterogeneity in employers produces a firm effect variance of .32-.36. Thus is appears that unobserved person-level skills and unobserved firm level qualities are both equally important in determining variation in earnings. The largest source of person-level heterogeneity comes from differences in the effects of experience. The fixed effects for experience give average experience effects for each demographic group and the parameter θ_2 gives individual deviations from these average effects. The sample variation in θ_2 is around .185, roughly double the sample variation in θ_1, which results from individual deviations from the average intercept terms. The interaction of the data source indicator with the person and firm random effects does not dramatically alter the results, although it does raise the variation of both types of random effects slightly. The variation due to error in the overall person effect (i.e. deviations in the SIPP and DER from the average person effect) is small, .009, meaning that the person effects in the SIPP and DER are highly correlated and the deviations from the average are small. However the small magnitude of the overall error disguises an interesting fact. The variance of the error in θ_1 is actually quite high, .085 compared to .09 for the true variance of θ_1. The low overall variance of the person effect is the result of low variance in the error of θ_2 and a negative correlation between errors in θ_1 and θ_2. Thus not only do high values of θ_1 tend to be associated with low values of θ_2 (i.e. people with high initial earnings have lower growth due to experience), but high errors in θ_1 tend to be associated with low errors in θ_2. In effect some of the error in the overall person effect cancels out. The variation due to error in the firm effect is also not large but is approximately double the person level error (.02), meaning that the SIPP and DER firm effects are slightly less correlated than the person effects.

The DER and SIPP measurement errors can be thought of as deviations from the average levels of earnings that are due to variation that was only found in one source and not the other. The SIPP measurement error captures variation due to fluctuations around mean earnings found only in the survey data. The DER measurement error does the same for the administrative data. In the first specification, any variation that is common to both sources is captured by the person, firm, or common error component, depending on the type of variation. In later specifications, some of the variation unique to one data source is attributed to differences in the person, firm, or fixed effects. Of the three error terms, ω (SIPP), υ (DER), and η (common time period error), υ has the highest autocorrelation coefficient, beginning at .73 in the first specification. The SIPP error ω has the lowest correlation across time with $\rho_\omega = .25$ initially. The common time period error η falls inbetween with a correlation coefficient of .59. However as the data source indicator is interacted with the fixed effects and then the random effects, the correlation of the DER error over time falls (.62 in specification 6), while the correlation of the SIPP error rises (.38 in specification 6) and the correlation of the common time period error stays approximately the same. A similar trend can be observed in the levels of variance of the three effects. The variance of the DER error is initially .51 but falls to .36 while the SIPP error has an initial variance of .21 and rises to .25. The variance of the common time period component stays at .76. Thus it appears that some of the error in the DER earnings can be quantified as appearing in the person or firm effects and being persistent over time. This may also be the result of mismatching between DER and SIPP jobs.

The fact that the variance of the DER measurement error is higher than the variance of the SIPP measurement error suggests that there is more variation that is unique to the administrative data than the survey data. As shown at the bottom of Table 13, the higher variance of the DER error compared to the SIPP error leads to a higher reliability ratio for the DER data compared to the SIPP data as calculated first using equation 3 and then equation 7. This difference declines as data source is interacted with the fixed and random effects, consistent with the variance of υ declining. The SIPP reliability ratio ranges from .87-.85 and the DER ratio ranges from .73 to .8. The SIPP reliability ratio is similar to those obtained by Bound *et al.*. in the PSID validation study (0.7 in 1986 and 0.85 in 1982) and Bound and Krueger in the CPS validation study (0.84 for 1976 and 0.82 for 1977). However our ratio was obtained in a much different manner. Our model estimates both the true and the measurement error variation, using the repeated earnings measures in a given time period to identify the model. These high reliability ratios for both the SIPP and the DER data give us confidence that the data collected by the survey and provided by the administrative records system of the Social Security Administration give meaningful information about the earnings of survey respondents.

Notably, the reliability ratio for first-differenced DER earnings (.75) is higher than for first-differenced SIPP earnings (.71) as calculated using equation 4. The reversal of the reliability ratios is due to the fact that the auto-correlation in the DER errors is substantially higher than the auto-correlation in the other error components. High auto-correlation in the DER errors means that some of the error gets differenced away because it is non-transitory. At the same time, some of the signal in η gets differenced away as well, but since the auto-correlation in η is not as high, less of the signal is lost in the first differencing. The opposite is true for the SIPP and hence the DER reliability ratio is the larger of the two. This points to a very different error creation process in the SIPP and the DER. Things which make the DER earnings higher or lower than the economic "true" earnings are more systematic and last longer whereas the things which cause the SIPP earnings to deviate from the common component of earnings are more likely to be one-time shocks.

Table 14 reports the fixed intercepts (β_{0SIPP} and β_{0DER}), the fixed linear time trend (β_2), and the fixed panel effects (β_3). The fixed intercept and panel effects can be viewed as the mean effect of the source of earnings, where the source is either the administrative data (i.e. the DER) or a particular survey (i.e. one of the SIPP panels). For instance the source effect for the 1996 SIPP panel is β_{0SIPP} and for the 1990 SIPP panel the source effect is $\beta_{0SIPP}+\beta_{31990}$. Comparing β_{0SIPP} and β_{0DER} tells us how average earnings in the 1996 panel differ from average earnings in the DER. Average earnings in the SIPP are lower than average earnings in the DER (6.85 versus 6.89), especially for the 1990 and 1991 SIPP panels, which have even lower average earnings than the 1992, 1993, and 1996 panels. This result is possible evidence that earnings were under-reported in the SIPP but have been improving over time. Interestingly, the coefficients for the panel effects seem to converge slightly as data source is interacted with the other fixed effects and random effects. The 1990, 1991 and 1993 panel effects become slightly less negative meaning the difference between these three panels and the 1996 panel lessens. Only the 1992 panel effect becomes slightly more negative. Overall it would seem that interacting data source with the random effects lessens the panel fixed effects, indicating that major differences among panels may be associated with differences in the people and firms. Finally the time effect, β_2, stays roughly constant across specifications. Between the early and late 1990s, all else equal, real wages declined slightly.

Table 15 reports race and gender fixed effects for all the specifications. These effects allow the four main demographic groups to have different levels of average earnings. Relative to white males, the other three demographic groups (non-white males, white females, non-white females) all have lower average earnings. The earnings of non-white males are 21% lower on average, while those of non-white females are 14% lower, and earnings of white females are 10% lower. These averages are remarkably stable across specifications. Even when the race and gender effects are interacted with data source, the average effect for each group remains the same and the deviation from the average is small and not significantly different from zero. Thus despite the fact that the overall level of average earnings is higher in the DER than in the SIPP, differences in this average due to race and gender are very similar between the two data sources.

Table 16 reports education effects, split out by demographic group. Individuals were classified into five educational categories: no high school degree (omitted group), high school degree, some college, college degree, and graduate school degree. For a given demographic group, the average increase in earnings due to a particular level of education stays relatively constant across specifications. However when education is interacted with data source, differences between the return to education in the administrative and survey data are revealed. For example in specification 6, white men have a 2.7% difference between SIPP and the DER return to a college degree and a 2.1% difference between the return to a graduate school degree. White women have a 1.2% and 1.3% difference in the return to college and graduate degrees. For both white men and women, the return to education is higher in the DER, and the difference is significant. Although the differences in the return to a high school degree or some college remain significant, they are smaller. For white men and women, the difference in DER and SIPP return never exceeds 1.3% (females high school degree). Interestingly, non-white men and women have somewhat different patterns. For non-white men, the return to a high school degree, some college, or a college degree is slightly higher in the SIPP than in the DER while the return to a graduate degree is higher in the DER. However none of these differences are significant. Non-white women are similar to white men and women in that they always have higher returns to education in the DER. However only the difference between the return to a graduate degree is significant. Yet this difference in SIPP and DER coefficients is the largest of any of the differences for specification 6. The difference between the average return to a graduate degree and the SIPP return is a full 4.2%. Highly educated non-white women appear to earn significantly more in the administrative data than in the SIPP.

Table 16 also reports experience effects for each demographic group for all the specifications. General labor force experience was modeled as a linear spline with nodes at 2, 5, 10, and 25 years and interacted with gender and race to give separate experience profiles for our four demographic groups. We calculated labor force experience using survey responses to questions about the year entering the labor force and time taken off work.[5] To be parsimonious in reporting results, Table 16 contains the coefficient on the 5-10 year piece of the experience spline for every specification. We graph the full experience profile for each demographic group using the effects from specification 6 only. As can be seen from looking at the effect of experience during the 5-10 year period of a person's career, the results do not change substantially across specifications. For white men, an additional year of experience after at least 5 years in the labor force yields approximately a 12% return. For non-white males, white females, and non-white females, the return is 10%, 9%, and 8% respectively. In specifications 2 and 6 where the fixed effects are interacted with the data source indicator, there are significant differences between the SIPP and DER experience coefficients but these differences never reach even 1%. Non-white males have the highest difference: .0085 and .0063 in specifications 2 and 6 respectively. White males and non-white females have similar differences (.0038 and .0035 respectively in specification 6) and white females have the lowest difference: .0029. For all the groups, at this point in the profile, the return to experience in the DER is higher than in the SIPP.

Figures 1-4 present the full picture of how earnings grow with experience for the different groups and in the different data sources. For example, Figure 1 shows earnings growing with accumulated experience for white males in both the SIPP and the DER. In each figure, the SIPP intercept is calculated as $\beta_{0SIPP} + \beta_{1SIPP(GRP)} + \beta_{2SIPPHighSchool(GRP)}$ and the DER intercept is calculated as $\beta_{0DER} + \beta_{1DER(GRP)} + \beta_{2DERHighSchool(GRP)}$ where GRP indicates the coefficient for the particular demographic group. For white men and women, the DER intercept is higher than the SIPP intercept but SIPP earnings grow more quickly over the 0-2 year range. Beginning at the 2 year mark, DER earnings grow more quickly and eventually surpass SIPP earnings. This cross-over takes place around 10 years for white women and around 14 years for white men. The difference continues to grow with experience and by 35 years of labor force experience, average log earnings are 9.11 in the DER and 9.03 in the SIPP for white women and 9.61 and 9.54 for white men. For non-white men, the DER intercept is approximately equal to the SIPP intercept, largely because the effect of a high school degree is slightly higher in the SIPP than the DER. SIPP earnings initially grow very quickly but DER earnings once again catch up and overtake the SIPP although not until approximately 22 years of experience. Non-white women are similar to white women except for them the cross-over in DER and SIPP earnings takes place around 16 years of experience. By 35 years of experience, the difference between SIPP and DER earnings for non-whites is not nearly as pronounced as for whites. In general, if one believes that the "true" return to experience lies somewhere inbetween these two lines, then on average, the SIPP overstates the return to experience in the early years of a person's career and understates the return in the later years with the reverse being true for the DER.

The results in Table 16 and Figures 1-4 lead us to conclude that there are significant differences in the fixed effects between SIPP and DER data but these differences depend on the demographic group. We next considered whether differences in the random effects depended on the demographic group. To answer this question, we used the predicted random effects ("blups") and the estimated residuals to calculate the necessary components of the reliability ratio separately for each demographic group.[6] Table 17 displays the SIPP and DER reliability ratios by demographic group. With the exception of white women,

[5] These questions were asked as part of the SIPP employment history topical module. This module consisted of a set of one-time only questions asked at the end of either the first wave (1992, 1993, and 1996 panels) or the second wave (1990 and 1991 panels).

[6] We actually calculated $\sigma^2_{\theta_{1AVG}}$, $\sigma^2_{\theta_{1DEV}}$, $\sigma^2_{\theta_{2AVG}}$, $\sigma^2_{\theta_{2DEV}}$, σ_η, σ_ω, and σ_υ using the blups for θ_1, θ_2, and η and the estimated residuals for

more educated groups have higher ratios of true to total earnings variance in both the SIPP and the DER. Within racial category, women have data with less error than men and whites have less error than non-whites, again in both data sources. As we found in the full sample, the reliability ratio is higher in the SIPP than in the DER for every group. But although the levels are very different across demographic groups, the difference between the SIPP and DER ratios is remarkably constant across groups - approximately 5%-6%. Thus we can say that some demographic groups have more reliable data than others but we cannot say that the administrative data provide an advantage over the SIPP data (or vice versa) for any certain demographic group. The gender and race effects on the random effects seem to apply equally to each data source.

Perhaps our most striking finding is that the reliability of the SIPP earnings data is actually greater than the reliability of the DER earnings data in all of our specifications (for example .85 versus .80 in specification 6). From our discussion above of the sources of differences between these two concepts, one is led naturally to the question: "Can the federal tax information reported on the W-2 really be less reliable than the respondent's reported earnings?" The statistical answer is clearly: yes, although a better interpretation might be that they are equally reliable. As we discussed in section 3.2, there are good reasons why the employer's direct report might contain some different information from the respondent's. Moreover, as we showed above in excruciating detail, the match between respondent-provided SIPP jobs and EIN-based DER jobs is also subject to error–an error that is intrinsic to the process of evaluating the data quality. In order to determine whether this match process is the critical factor driving this major result, we summed the earnings from each job held by an individual in a given year to create annual total earnings for each person in our data set and then we re-did our analysis just using these person-level totals. This necessitated dropping the firm effect from the estimation equations but otherwise they remained unchanged. We estimated specifications equivalent to 1, 2, 3, and 6, i.e. no data source interaction initially, followed by source interaction with fixed effects, source interaction with person random effects, and finally source interaction with both fixed and person random effects. The results are presented in Tables 18-21, which have the same form as the job-level result tables.

The random effects in Table 18 follow very similar patterns to the random effects in Table 13. The variance of the DER error term and the level of autocorrelation in this error drop sharply across the specifications. By the point at which the fixed and random effects are both interacted with data source, the variance of the DER error has actually fallen slightly below the variance of the SIPP error and the reliability ratio for the DER data exceeds that of the SIPP data. Several things may be inferred from this result. First, the job-matching process does seem to have introduced some error that caused systematic differences in the DER and SIPP data for some jobs and this led to high error in the DER jobs. Second, at the person level there is still enough variation that is unique to the DER data that initially the DER error variance is larger than the SIPP error variance. Only when the person random effects are allowed to differ between the SIPP and the DER, does the variance of the DER error term drop below the variance of the SIPP error term. Thus there are systematic differences at the person level between the SIPP and the DER, i.e. some of the variation that was lumped into the measurement error term initially is actually measurement error in the person random effects. However this measurement error is not identified separately for the SIPP and DER person effects. We can simply calculate the average $\theta's$ and deviations from this average, but these deviations by definition have the same absolute value for the SIPP and the DER. Thus error in the person random effects enters both the SIPP and DER reliability ratios. When some of the variance of the DER error term, υ, is shifted to error in θ_1 and θ_2 then this raises the DER reliability ratio relative to the SIPP ratio if relatively less variance from ω is shifted into error in the person random effects.

In Table 19 we see that average earnings in the DER are again higher than in the SIPP. Surprisingly, the signs of the coefficients on the SIPP panel indicators have switched compared to Table 14. The 1990-1993 panels now have higher average earnings than the 1996 panel and the effects are much higher: .03-.046. It is unclear what exactly caused this change. In Table 20 we see that for non-white males the difference in the SIPP and DER race/gender effects is now significant and rather large, .05. In Table 21, we see that the DER education effects are again higher than the SIPP effects for every demographic group at every level, except for non-white males with college and graduate degrees where the SIPP effect is higher. For whites, the differences are all significantly different from zero with the highest one being the difference in the return to a college degree for a white male (.02). For non-white men, none of the differences are statistically significant while for non-white women, only the difference for a graduate degree is significant. These patterns are essentially the same as we saw in the job-level results.

The experience effect for the 5-10 years piece of the spline is also shown in Table 19. For whites, both male and female, the SIPP coefficient is slightly higher than the DER coefficient but it is only significant for white females. For non-whites, the DER coefficient is the larger of the two but for neither men nor women is it significant. Figures 5-8 show graphs of the same form as Figures 1-4. These graphs tell very different stories than the job-level ones. DER earnings begin higher for every demographic group and although SIPP earnings initially grow faster with experience, they never catch-up to DER earnings and at some point in the later career, the difference begins to widen. This happens earlier for non-whites than for whites so by 35 years the gap between DER and SIPP earnings is wider for non-whites than whites. Hence we conclude

ω and υ. We used the full sample estimate for $\sigma^2_{\psi AVG}$ and $\sigma^2_{\psi DEV}$ because we were concerned about the non-representative nature of a subset of the firms in our sample. We also used the full sample estimate for $\sigma_{\theta_{1AVG}\theta_{2AVG}}$ and $\sigma_{\theta_{1DEV}\theta_{2DEV}}$ because the estimation software did not provide the necessary output to correctly calculate the covariance of a subset of blups. We considered solving this second problem by estimating the model for each demographic group separately but were again concerned that the firm effects would not be reliable because they would be estimated using only a subset of firms and a subset of people who worked at those firms.

that at the person level higher initial returns to experience in the SIPP are not sufficient to counteract higher initial earnings in the DER and there is no cross-over of earnings from the two sources as we saw at the job-level. However, as with the job-level results, higher returns to experience in the DER later in people's careers leads to an increasing gap between DER and SIPP earnings for more experienced workers.

5 Conclusion

The levels of measurement error calculated in this paper give some cause for optimism. First, for jobs found in both the SIPP and the SSA DER records, the earnings reports from the two sources are highly correlated. Second, measurement error accounts for only 15% of the variation in SIPP annual earnings and 20% of the variation in DER annual earnings. However errors in both types of data are correlated over time and in the case of the SIPP, this correlation makes the attenuation bias resulting from measurement error worse than in the DER data. We find significant differences in returns to education and experience between the two data sources, although these differences depend on the demographic group considered. Generally returns to education are higher in the DER but this result is significant only for whites and non-white women with graduate degrees. Returns to experience are initially higher in the SIPP but at some point, the return becomes higher in the DER. Hence workers late in their careers will have a significant gap between DER and SIPP earnings than can be attributed to differences in the return to experience. We find that non-whites and men have lower reliability ratios than whites and women but that no demographic group has more reliable data in the DER than in the SIPP. When we do the analysis at the person-level (as opposed to the person-job level), we find that in some specifications the reliability ratio of the DER data is higher than that of the SIPP data, although the difference is small (.80 for the DER and .78 for the SIPP). In effect, doing the analysis at the person-level causes the reliability ratios to converge, perhaps because any error in the job match is no longer of consequence at the person-level.

The substantial advantage of our approach is the ability to extract the common "economic" component of the earnings measure without having to assert that either SIPP or DER earnings records were "true." Researchers with access to these components could, in principle, make better inferences about the relation between annual earnings and other survey variables than would have been possible with either variable alone.

References

[1] Abowd, John M. and David Card. 1989. "On the Covariance Structure of Earnings and Hours Changes." *Econometrica* 57(2):411-446.

[2] Abowd, John M. and Lars Vilhuber. 2002. "The Sensitivity of Economic Statistics to Coding Errors in Personal Identifiers." *Technical Paper TP-2002-17*, LEHD, U.S. Census Bureau.

[3] Angrist, Joshua D. and Alan B. Krueger. 1999. "Empirical Strategies in Labor Economics." In *Handbook of Labor Economics Vol. 3A*, ed. Orley Ashenfelter and David Card, 1277-1357. Amsterdam: Elsevier.

[4] Bound, John and Alan B. 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.

[5] Bound, John, Charles Brown, Greg J. Duncan, Willard L. Rodgers. 1994. "Evidence on the Validity of Cross-sectional and Longitudinal Labor Market Data." *Journal of Labor Economics* 12(3):345-368.

[6] Bound, John, Charles Brown, Nancy Mathiowetz. 2001. "Measurement Error in Survey Data." In *Handbook of Econometrics Vol. 5*, ed. James J. Heckman and Edward Leamer, 3705-3843. Amsterdam: Elsevier.

[7] Duncan, Greg J. and Daniel 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.

[8] Fellegi, I.P. and A.B.Sunter. 1969. "A Theory for Record Linkage." *Journal of the American Statistical Association* 64:1183-1210.

[9] Fuller, Wayne. 1987. *Measurement Error Models*. New York: John Wiley and Sons, Inc.

[10] Gilmour, Arthur R., Robin Thompson, Brian R. Cullis, "Average Information REML: An Efficient Algorithm for Variance Parameter Estimation in Linear Mixed Models," *Biometrics* 51(4):1440-1450.

[11] Mellow, Wesley and Hal Sider. 1983. "Accuracy of Response in Labor Market Surveys: Evidence and Implications." *Journal of Labor Economics*, 1(4):331-344.

[12] Newcombe, H.B., J.M. Kennedy, S.J. Axford, and A.P. James. 1959. "Automatic Linkage of Vital Records." *Science* 130:954-959.

[13] Patterson, H.D. and R. Thompson. 1971. "Recovery of Interblock Information when Block Sizes are Unequal." *Biometrika.* 58(3):545-554.

[14] Stinson, Martha H. 2003. "Technical Description of SIPP Job Identification Number Editing, 1990-1993 SIPP Panels." *LEHD Technical Paper*, U.S. Census Bureau.

6 Appendix A Editing of SIPP JOB ID Variable

As described in Section 3.1, the SIPP job id variable had difficulty in longitudinally tracking jobs (see tables 2, 3A, 3B for examples of the problems). In the early SIPP panels, the problems were mostly the result of field representatives being required to collect information about an on-going job over and over again. Inconsistencies crept in over time as the name of the employer was collected and written down separately at the time of each interview. Wave-specific names differed both across and within the original SIPP job IDs. Different spellings, use of abbreviations in later waves, and slightly different wording were the most common differences within job IDs. In contrast, the 1996 panel only recorded employer names when new jobs were begun and hence employer names differed only across job IDs and not within. In the 1990-1993 panels, the goal was to create an entirely new set of job IDs that was not derived from the old job IDs because these were deemed too unreliable. Hence it was necessary to compare all person-job-wave records for a given individual and group those with the same name. In 1996, however, the goal was simply to check and see if some jobs for an individual should be linked because an individual may have missed a survey wave and been incorrectly assigned a new job id when he or she was next interviewed. Hence person-job-wave observations with the same job ID assigned were accepted as belonging to the same job and job-level records were created. These job-level records were then compared and those with names deemed to be the same were grouped together.

Because of the spelling and wording differences across observations, we used probabilistic matching methods as developed by Newcombe, Kennedy, Axford, and James (1959) and Fellegi and Sunter (1969) and implemented in a commercial software program called Integrity. These methods have been used extensively at the U.S. Census Bureau to solve problems of miss-coded identifiers (for an example of an application of probabilistic person name matching to fix SSN miscodes, see Abowd and Vilhuber (2002)). This method involves grouping records into "blocks" of possible matches and then computing matching weights for pairs of records within the "block". Pairs with matching weights above a certain threshold, or cutoff point, are deemed to be matches and those with weights below another threshold are deemed to be non-matches. Those pairs with matching weights in between the two thresholds are termed uncertain and clerical review is suggested.

A matching weight for a pair of records is a composite score that is created by comparing the records across a variety of fields, assigning a weight to each field based on a determination of whether the field agrees or disagrees, and then summing the weights from all the fields involved in the comparison. Each field used in the matching is assigned an m and u probability. The m probability is the probability that the same field on two separate records agrees given that the two records were indeed a match. When this probability is set to less than one, it is assumed that there are some errors in the fields and that even if two records are a match, there is still some probability that the field is miscoded on one of the records and the two fields will disagree. The u probability is the probability that the same field on two separate records agrees given that the records are not a match. This is the probability that a field agrees at random. Given the m and u probabilities, agreement and disagreement weights for each field are calculated using the following formulas:

$$\begin{aligned}\text{agreement weight} &= \log_2(\frac{m}{u}) \\ \text{disagreement weight} &= -(\log_2(\frac{1-m}{1-u}))\end{aligned}$$

The decision of whether a field agrees or disagrees, and hence whether it receives the agreement or disagreement weight, can be implemented in a variety of different ways. One can be quite strict and insist on absolute identity in order to declare agreement or one can allow some level of discrepancy between fields without declaring disagreement. This flexibility is especially useful for name matching because it allows the user to take account of potential misspelling of words.

In our application, we blocked on the SIPP person identifier and hence only job records for the same person were compared. To create the fields for comparison we parsed the reported name into several pieces. Common words such as "Inc," "Company," or "Firm" were saved in one set of fields while geography words such as state names were saved in another set. The remaining words from the name were thought most likely to be unique to a particular employer and were saved in a third set of fields. We performed several sets of comparisons, or passes, using different fields in each pass. The choice of m and u probabilities and cutoff levels was determined both by knowledge about the fields and by experimentation. For the fields that contained unique name words, a high m probability and low u probability were chosen. Since these words were deemed to be the part of the employer name that was unique to that firm, matching values were essential to matching records, thus requiring the high m probability. At the same time, these words were unlikely to agree at random and hence produce false matches, so a low u probability was chosen. The result of these choices was that matching values of the unique word names received very high agreement weights and also very high disagreement weights. The fields that contained common words and geography words, on the other hand, had higher u probabilities. Agreement in one of these fields produced a lower agreement weight because matches were more likely to happen at random while disagreement produced a more negative disagreement weight because non-matches meant the companies were unlikely to be the same. Cutoff values were chosen by examining certain and uncertain matches and determining the range of their weights. Appendix Table A1 gives the exact blocking and matching fields used along with their m and u probabilities.

For the early SIPP panels, probabilistic name matching alone proved inadequate for creating a consistent set of job IDs. While the name matching procedure both separated jobs records originally assigned the same job ID and connected job records originally assigned different job IDs, the former was the most common outcome. This can be seen in Table 4A by the fact that the number of total jobs rose substantially after Integrity processing. This result was due to the fact that the most common problems in the survey were the re-use of job IDs, as described in Table 3B, and the high degree of irregularity in the spelling of job names and the common use of abbreviations in later waves. Universities and government agencies with common acronyms were especially problematic. For example, the Integrity software could not recognize the names "University of X" and "UofX" or "Department of Y" and "DOY" as being the same. Hence in these panels, a considerable amount of clerical review was undertaken in order to separate cases where Integrity correctly and incorrectly split job records. In cases where there were discrepancies between the total job count in the SIPP and the total job count in the DER, job records were output and reviewed by two separate individuals. When one of the reviewers discovered two jobs in a respondent's job history that appeared to be the same, she manually changed the job ID to reflect this determination. The second reviewer re-checked all these changes as a quality assurance measure. After this extensive manual review, a few final edits were performed to locate any final obvious cases where Integrity had erroneously failed to link job records. The work history of any person who had one job that consisted of at least four linked job records and a second job that consisted of only one job record was examined to see whether the single job record in fact belonged to the job with at least four linked records. Corrections to job IDs were made to link job records that were determined to belong to the same job.

7 Appendix B

The merge between the DER and the Business Register was somewhat complex because the Business Register had two parts. The first part was called the Single-unit file and contained records for all EINs that were either single-unit companies or sub-masters. Single-unit companies were firms with only one establishment that had a single EIN. Sub-masters were companies with multiple establishments that shared an EIN, *i.e.* multi-unit companies. For single-unit companies, the names and industries found on the Single-unit Business Register file were likely to correspond to the names and industries of employers reported in the SIPP. However for sub-masters, the name and industry were potentially quite different because these represented some aggregate concept - name of parent company or major industry out of a group of industries represented within a multi-unit company. Hence for sub-masters, we also searched for information about the EIN in the second part of the Business Register, the Multi-unit file. Here we obtained multiple records for each EIN representing the names and industries of all the different establishments associated with a sub-master record. For these multi-unit companies, we kept one record for each unique three-digit industry. Establishments within the same industry tended to have extremely similar names and hence this choice resulted in both a manageable number of observations to match to SIPP jobs while still providing additional information that might assist in the match.

The Business Register is maintained on a yearly basis. Initially an EIN from a job was sought in the Business Register year that corresponds to the first year the job was reported in the DER. If a job was already in progress at the time of the beginning of the survey, the start year was coded to be the first survey year since this was the first year the job was at risk to match to the SIPP. If the job was not found in the Business Register year corresponding to the start year, it was sought in the following two Business Register years. Appendix Table B1 presents a summary of the match rates between the DER and the Business Register. There are several interesting things to notice in this table. First, the match rates are extremely high, 98% for every panel except 1996. The low match rate for the 1996 panel relative to the other four panels can be explained by the fact that the latest year for which the Business Register was available at the time this research was conducted was 1999. Thus any job in the DER for a SIPP respondent from the 1996 panel that began in the year 2000 could not be matched to the Business Register. For the purposes of this study, this lack of data did not present a serious problem because so little SIPP data was collected in 2000 that annual earnings from jobs beginning in 2000 could not be accurately constructed for SIPP jobs. As described in Section 3.3., jobs beginning in the year 2000 were dropped from both the SIPP and the DER before comparing earnings. The second interesting thing to note is that although only 27%-32% of all EINs were multi-unit companies, these EINs accounted for 39%-44% of all jobs. SIPP respondents disproportionately work for large companies. Third, a small percentage of EINs and jobs were found in the Multi-Unit file but not in the Single-Unit file. The cause of this is unknown at this time and will need further research.

8 Appendix C

The job-level match between the SIPP and DER data compared employer name, calendar year indicators, and industry in order to link records from each source. On one side of the match were all the SIPP jobs deemed to be reports of employment at a single employer. On the other side of the match were all the records associated with the DER jobs deemed to have taken place during the at-risk time frame. Each DER record contained the name and industry of the EIN as found on the Single-unit part of the Business Register. When the EIN was also found on the Multi-unit part of the Business Register, the record contained a second name and industry representing information about a particular establishment of this EIN. When an EIN was associated with multiple establishments with different industries on the Multi-unit file, multiple records were created for this DER job. Each record contained the same Single-unit name and industry information but different Multi-unit name and industry information. This was done in an attempt to maximize the number of job matches obtained by using all possible name information associated with multi-unit companies. For example a person might report working for company X in the SIPP and have a job report in the DER with EIN A that is a multi-unit. The main company name of EIN A may be Y but one of the subsidiary establishments may be called X. By attaching names X and Y to EIN A in the DER, we increase the likelihood that this job will match to the SIPP job reported at company X.

Appendix Table C1 gives the blocking and matching fields for each pass along with the accompanying m and u probabilities for the matching fields. Several variables were also used in multiple passes, with the requirements for matching gradually relaxed. For example, in the third pass, three-digit Single-unit industry was used as a blocking variable and the four year-indicators were used as matching variables. Pass five was quite similar except that instead of requiring records to match on all four year-indicators, only start year was required to match. Start year was a field that indicated the first year that a record was found for this job with the first possible year being the year that data was first collected in the survey and the last possible year being the last year data was collected in the survey. Likewise in pass seven, only one-digit Single-unit industry was used as a blocking variable. This process enabled the detection of high-probability matches in early passes and then the addition of lower-probability matches in later passes.

Appendix Table C2 shows the results of the matching. Of the SIPP jobs, between 77%-79% were successfully matched to a DER job. Of these matches, 86%-88% were deemed high probability matches that surpassed the clerical editing threshold, while the remaining matches were between the clerical threshold and the no-match cutoff point or were duplicate matches. The majority of the matching took place in the first pass (between 75%-83% of all matches). The next most successful passes were 3 (5%-9%) and 7 (5%-6%).

Appendix Tables C3 and C4 highlight two problems that resulted from the matching. First, two different SIPP jobs could match to the same DER Job. An example of this type of case is illustrated in Table C3. There were several possible causes of this problem. First, it was possible that the two SIPP jobs were indeed the same and the SIPP job creation phase erroneously failed to link them. In this case the duplicate record was a "true" duplicate and both jobs were correctly matched to one DER job. However another possibility was that the matching software mistakenly matched a second SIPP job to the same DER job due to lack of differentiating information for the SIPP jobs. This was particularly likely in the later passes where matches were based on year and industry indicators alone. In this case, the duplicate was false and only one of the two matches was correct. Careful inspection of duplicate cases led to the adoption of the following rule: if the two SIPP jobs had been matched to the one DER job in either the first or second pass and there were 2 or fewer residual DER jobs left that had not matched to any SIPP job, then the second SIPP job was declared a true duplicate. It was combined with the first SIPP job to become one single SIPP job matched to the one DER job. Otherwise if the two SIPP jobs had matched to the one DER job in pass 3 or later or they had matched in pass 1 or 2 and there were 3 or more residual DER jobs, the duplicate was declared false and only the master record match was kept. The duplicate SIPP job was changed to be a residual, non-matching SIPP job. The total number of duplicates that were determined to be "true" and hence were subsequently combined is shown in row 3 of Table 9.

The second problem was the reverse duplication issue: two different DER jobs sometimes matched to the same SIPP job. Table C4 gives an example. This type of duplication was more common and it was more difficult to know the causes. The first possibility was that a company changed its EIN due to a change in ownership structure or some other reason. This is the successor-predeccessor problem described earlier. Another possibility was that SIPP respondents reported "lump" jobs, meaning that one SIPP job was really a combination of several jobs. Since administrative records pertained to the source of the earnings, it was possible that some individuals considered themselves as holding only one job but were paid from several different source EINs. It was also possible that individuals consciously grouped jobs in order to ease the burden of responding to the survey. These issues warrant further research.

We made a first attempt to tell whether two DER jobs that matched one SIPP job were indeed the "same" job by using some additional information from the Census Business Register. Previously we had augmented our list of EINs from the DER to include parent company information and name and industry information from one establishment of every unique three-digit industry group within the parent company. We then added annual geography information (a geocode created from the exact address) to each EIN at both the parent company level and the establishment level. We compared this geography information for each year of the survey across the two EINs and if the geocode was ever the same for the parent company or the establishment, we declared the two jobs to be duplicates. The intent of this geocode comparison exercise was to find cases where an EIN changed but the physical location did not change and hence it was likely that the SIPP respondent still considered himself to be at the same job. Since we did not keep every establishment within an industry group, we clearly did not compare every possible geocode. Hence our determination of how many DER jobs were duplicates and should be combined is probably an undercount.

We also added parent company identifiers to the EINs so we could tell if two EINs had some kind of ownership relationship. Two DER jobs that matched to one SIPP job but had the same parent company identifier were also declared to be a match. In this case it seemed possible that the SIPP respondent had kept the “same” job but had moved within the company or had simply experienced a company re-organization where the EIN tax reporting structure had changed. Row 4 of Table C2 shows how many DER job duplicates were determined to be legitimate.

Table 1: Original SIPP Job Summary

SIPP Panel	1990	1991	1992	1993	1996
Total SIPP respondents	69,432	44,373	62,412	62,721	116,636
Respondents who ever report a job	37,291	23,520	33,920	32,972	63,600
Person-job-wave observations	216,851	136,693	228,214	208,748	498,553
Jobs defined by original SIPP jobid	57,800	35,515	55,453	52,591	136,550

Table 2: SIPP Job ID Problems, 1996 Panel

Wave	Start date	Jobid
1	Feb. 1, 1996	1
2	Feb. 1, 1996	1
3	Feb. 1, 1996	1
4	Feb. 1, 1996	1
5		
6	Jan. 1, 1996	2

Table 3A: SIPP Job ID Problems, 1990-1993 Panels

Failure to link job across waves		
Wave	Firm Name	SIPP Jobid
1	AAAA	1
2	AAAA	1
3	AAAA	2
4	AAAA	1

Table 3B: SIPP Job ID Problems, 1990-1993 Panels

Failure to separate jobs across waves		
Wave	Firm Name	SIPP Jobid
1	AAAA	1
2	AAAA	1
3	AAAA	1
4	BBBB	1

Table 4A: Summary of the Job ID Editing Process, 1990-1993 SIPP Panels

	SIPP Panel	1990	1991	1992	1993
1	Number of Jobs, post name matching pass 1	78,225	46,316	74,078	68,803
2	Jobs belonging to people with conflict with DER	45,725	24,149	38,752	35,352
3	Number of Jobs, post name matching pass 2	69,138	41,814	66,602	62,251
4	Jobs belonging to people with conflict with DER	10,011	5,106	8,131	8,330
5	Number of Jobs, post clerical edits	66,991	40,818	65,278	61,094
6	Jobs belonging to people with conflict with DER	7,089	3,800	6,448	6,670

Table 4B: Summary of Job ID editing process, 1996 SIPP Panel

	SIPP Panel	1996
1	Jobs with startdate probs	29,520
2	Number of Jobs, post name matching	125,358
3	Jobs belonging to people with conflict with DER	15,331
4	Jobs with startdate probs	22,353

Table 5: Jobs from the DER

	SIPP Panel	1990	1991	1992	1993	1996
1	Total DER jobs	432,105	263,063	364,261	347,485	607,873
2	Total EINs	235,910	155,013	205,951	197,881	315,471
3	Years covered by survey	1989-1992	1990-1993	1991-1995	1992-1995	1995-2000
4	DER jobs in survey time	96,086	58,020	99,524	81,320	192,720
5	EINs	60,131	38,628	62,406	51,880	105,095

Table 6: Match Rates for People with Jobs in the SIPP and DER

SIPP Panel		SIPP	DER	Both SIPP	Both DER
1990	People with jobs	37,291	35,032	30,993	
	Total Jobs held	66,991	96,086	55,087	88,324
1991	People with jobs	23,520	21,729	19,056	
	Total Jobs held	40,818	58,020	32,447	52,797
1992	People with jobs	33,920	31,557	27,394	
	Total Jobs held	65,278	99,524	51,650	90,360
1993	People with jobs	32,972	29,831	26,267	
	Total Jobs held	61,094	81,320	47,723	74,317
1996	People with jobs	63,116	55,894	48,542	
	Total Jobs held	121,450	192,720	97,149	173,623

Table 7A: Reasons SIPP Workers Do Not Match DER

SIPP Panel	Total SIPP People	SIPP People without Valid SSNs	People Who Have Only SIPP Jobs	People in SIPP and DER
1990	37,291	4,856	1,442	30,993
1991	23,520	3,629	835	19,056
1992	33,920	5,477	1,049	27,394
1993	32,972	5,535	1,170	26,267
1996	63,116	12,425	2,149	48,542

Table 7B: Reasons DER Workers Do Not Match SIPP

SIPP Panel	Total DER people	DER People without Valid SSNs	People Who Have Only DER Jobs	People in SIPP and DER
1990	35,032	0	4,039	30,993
1991	21,729	0	2,673	19,056
1992	31,557	0	4,163	27,394
1993	29,831	0	3,564	26,267
1996	55,894	0	7,352	48,542

Table 8: Final Sample of Matched Jobs

SIPP Panel	1990	1991	1992	1993	1996	Total
Number of Matched Jobs after Combining Duplicates	41,885	25,258	39,729	36,469	75,110	218,451
Jobs w/out SIPP and DER Earnings in Sample Years	3,182	2,005	798	4,709	434	11,128
New Matched Job Total	38,703	23,253	38,931	31,760	74,676	207,323

Table 9: Annual Job Summary Statistics

	Year	1990	1991	1992	1993	1994	1996	1997	1998	1999
1	N SIPP	30,402	45,669	42,177	48,665	47,277	42,901	41,334	37,839	35,112
2	N DER	30,103	46,625	42,715	49,442	49,370	44,003	44,456	42,682	39,994
3	Mean Ln(SIPP real annual earnings)	9.04	9.10	9.12	9.13	9.14	9.08	9.14	9.21	9.27
4	St.Dev. Ln(SIPP real annual earnings)	1.46	1.45	1.44	1.43	1.42	1.45	1.44	1.42	1.38
5	Mean Ln(DER real annual earnings)	9.10	9.17	9.20	9.18	9.16	9.12	9.15	9.20	9.26
6	St.Dev. Ln(DER real annual earnings)	1.59	1.57	1.56	1.57	1.61	1.59	1.60	1.61	1.60
7	Mean years of experience	15.11	15.63	15.46	15.44	15.61	18.00	18.20	18.44	18.94
8	St. Dev. years of experience	11.68	11.69	11.48	11.41	11.41	12.44	12.48	12.54	12.61

Table 10: Covariance/Correlation Matrix for Ln(SIPP Job Annual Earnings)

	1990	1991	1992	1993	1994	1996	1997	1998	1999
1990	2.1172	0.75							
1991	1.1248	2.0894	0.76						
1992		1.1245	2.0789	0.75	0.69				
1993			1.1407	2.0571	0.74				
1994			0.8246	1.0992	2.0269				
1996						2.0980	0.72	0.66	0.63
1997						1.1462	2.0761	0.72	0.66
1998						0.8347	1.0898	2.0166	0.72
1999						0.7061	0.8087	1.0451	1.8923

Notes: Covariances on and below the diagonal; correlations above the diagonal.

Table 11: Covariance/Correlation Matrix for Ln(DER Job Annual Earnings)

	1990	1991	1992	1993	1994	1996	1997	1998	1999
1990	2.5230	0.77							
1991	1.4069	2.4564	0.78						
1992		1.3763	2.4457	0.78	0.72				
1993			1.4336	2.4741	0.78				
1994			1.1458	1.4259	2.5910				
1996						2.5372	0.77	0.73	0.69
1997						1.4944	2.5486	0.77	0.72
1998						1.2382	1.4634	2.5911	0.78
1999						1.0581	1.1911	1.4850	

Notes: Covariances on and below the diagonal; correlations above the diagonal.

Table 12: Correlation Matrix of SIPP/DER Job Annual Earnings

	Ln(SIPP Job Annual Earnings)								
Ln(DER Job Annual Earnings)	1990	1991	1992	1993	1994	1996	1997	1998	1999
1990	0.86	0.71							
1991	0.70	0.87	0.72						
1992		0.70	0.87	0.72	0.66				
1993			0.72	0.86	0.71				
1994			0.65	0.70	0.86				
1996						0.85	0.68	0.63	0.60
1997						0.70	0.84	0.68	0.63
1998						0.65	0.68	0.84	0.68
1999						0.61	0.63	0.67	0.83

Table 13: Variances of Random Effects

		Spec1	Spec2	Spec 3	Spec 4	Spec 5	Spec6
person effect	$\sigma^2_{\theta 1}$ (avg. intercept)	0.0997	0.0963	0.0866	0.1046	0.0932	0.0902
	$\sigma^2_{\theta 1 Error}$ (dev. from avg.)			0.0846		0.0869	0.0848
	$\sigma^2_{\theta 2}$ (avg. slope)	0.1849	0.1844	0.1884	0.1856	0.1861	0.1865
	$\sigma^2_{\theta 2 Error}$ (dev. from avg.)			0.0170		0.0175	0.0167
	$\sigma_{\theta 1 \theta 2}$ (covar of avgs.)	-0.1302	-0.1287	-0.1245	-0.1327	-0.1262	-0.1257
	$\sigma_{\theta 1E \theta 2E}$ (covar of devs.)			-0.0359		-0.0374	-0.0362
total true var. person eff.		0.3014	0.3018	0.3237	0.2995	0.3152	0.3155
total error var. person eff.				0.0110		0.0093	0.0085
firm effect	σ^2_{φ}	0.3267	0.3267	0.3245	0.3637	0.3606	0.3589
	$\sigma^2_{\varphi Error}$				0.0224	0.0218	0.0213
common error	σ^2_{η}	0.7743	0.7750	0.7685	0.7706	0.7681	0.7686
	ρ	0.5873	0.5882	0.5843	0.5859	0.5843	0.5848
SIPP error	σ^2_{ω}	0.2130	0.2121	0.2340	0.2433	0.2568	0.2538
	ρ_{SIPP}	0.2531	0.2519	0.3233	0.3487	0.3865	0.3798
DER error	σ^2_{ν}	0.5126	0.5087	0.4460	0.4011	0.3505	0.3549
	ρ_{DER}	0.7396	0.7365	0.7008	0.6633	0.6146	0.6191
SIPP Reliability Ratio		0.8681	0.8687	0.8526	0.8549	0.8444	0.8462
St err of RR		0.0012	0.0012	0.0013	0.0012	0.0014	0.0014
DER Reliability Ratio		0.7323	0.7340	0.7561	0.7814	0.8005	0.7988
St err of RR		0.0011	0.0011	0.0013	0.0013	0.0015	0.0015
SIPP First Diff. Rel. Ratio		0.7133	0.7135	0.7153	0.7142	0.7157	0.7158
DER First Diff. Rel. Ratio		0.7478	0.7467	0.7488	0.7457	0.7460	0.7458

Table 14: Fixed Effects: Intercepts, Panels, and Time

		Spec 1	Spec 2	Spec 3	Spec 4	Spec 5	Spec 6
SIPP Intercept	β_{0SIPP}	6.6781	6.6764	6.6479	6.6797	6.6578	6.6496
		0.0212	0.0221	0.0211	0.0212	0.0212	0.0225
DER Intercept	β_{0DER}	6.6948	6.6815	6.6841	6.6726	6.6660	6.6876
		0.0209	0.0234	0.0209	0.0210	0.0209	0.0239
time effect	β_2 (time)	-0.0223	-0.0231	-0.0216	-0.0212	-0.0207	-0.0216
		0.0009	0.0009	0.0009	0.0009	0.0009	0.0009
SIPP 1990 panel	β_{31990}	-0.0426	-0.0531	-0.0371	-0.0318	-0.0277	-0.0372
		0.0048	0.0048	0.0049	0.0048	0.0049	0.0049
SIPP 1991 panel	β_{31991}	-0.0523	-0.0591	-0.0459	-0.0429	-0.0381	-0.0448
		0.0056	0.0057	0.0057	0.0056	0.0057	0.0057
SIPP 1992 panel	β_{31992}	-0.0238	-0.0334	-0.0220	-0.0193	-0.0178	-0.0257
		0.0045	0.0046	0.0046	0.0045	0.0046	0.0046
SIPP 1993 panel	β_{31993}	-0.0125	-0.0190	-0.0070	-0.0072	-0.0031	-0.0091
		0.0049	0.0050	0.0050	0.0049	0.0050	0.0050

*coefficients in first row, standard errors in second row for each effect

Table 15: Fixed Effects: Race and Gender													
	Spec 1	Spec2				Spec3	Spec4	Spec5	Spec6			
	Main	DER	SIPP	Average	abs value dev.				DER	SIPP	Average	abs value dev.
Non-white Males	**-0.2133**	**-0.2104**	**-0.2089**	**-0.2096**	**0.0008**	**-0.2173**	**-0.2146**	**-0.2182**	**-0.2195**	**-0.2094**	**-0.2145**	**0.0051**
	0.0570	0.0644	0.0595	0.0573	0.0236	0.0568	0.0570	0.0568	0.0653	0.0607	0.0570	0.0268
White Females	**-0.1013**	**-0.1019**	**-0.0927**	**-0.0973**	**0.0046**	**-0.0995**	**-0.1010**	**-0.0981**	**-0.1021**	**-0.0972**	**-0.0997**	**0.0024**
	0.0287	0.0323	0.0300	0.0289	0.0117	0.0286	0.0287	0.0286	0.0329	0.0305	0.0287	0.0134
Non-white Females	**-0.1385**	**-0.1474**	**-0.1320**	**-0.1397**	**0.0077**	**-0.1389**	**-0.1396**	**-0.1380**	**-0.1440**	**-0.1318**	**-0.1379**	**0.0061**
	0.0524	0.0593	0.0546	0.0528	0.0216	0.0522	0.0524	0.0522	0.0601	0.0556	0.0525	0.0245

Table 16: Fixed Effects: Education and Experience

	Spec 1	Spec2				Spec3	Spec4	Spec5	Spec6			
	main	DER	SIPP	Average	abs value dev.	main	main	main	DER	SIPP	Average	abs value dev.
White Male												
high school	**0.3093**	**0.3304**	**0.3007**	**0.3156**	**0.0149**	**0.3067**	**0.3093**	**0.3077**	**0.3212**	**0.3034**	**0.3123**	**0.0089**
	0.0129	0.0143	0.0131	0.0130	0.0043	0.0128	0.0128	0.0128	0.0145	0.0130	0.0130	0.0044
some college	**0.3699**	**0.3924**	**0.3608**	**0.3766**	**0.0158**	**0.3660**	**0.3684**	**0.3659**	**0.3787**	**0.3622**	**0.3704**	**0.0083**
	0.0127	0.0142	0.0129	0.0129	0.0043	0.0127	0.0127	0.0127	0.0143	0.0128	0.0129	0.0044
college	**0.8560**	**0.9176**	**0.8338**	**0.8757**	**0.0419**	**0.8497**	**0.8531**	**0.8496**	**0.8924**	**0.8380**	**0.8652**	**0.0272**
	0.0155	0.0173	0.0158	0.0157	0.0051	0.0155	0.0155	0.0155	0.0174	0.0157	0.0157	0.0052
graduate school	**0.9482**	**0.9947**	**0.9316**	**0.9631**	**0.0316**	**0.9431**	**0.9480**	**0.9449**	**0.9779**	**0.9357**	**0.9568**	**0.0211**
	0.0162	0.0181	0.0165	0.0164	0.0054	0.0162	0.0162	0.0162	0.0182	0.0163	0.0164	0.0054
annual return	**0.1207**	**0.1287**	**0.1182**	**0.1234**	**0.0053**	**0.1193**	**0.1200**	**0.1191**	**0.1259**	**0.1182**	**0.1220**	**0.0038**
5-10 years exp	0.0035	0.0038	0.0036	0.0035	0.0012	0.0035	0.0035	0.0035	0.0039	0.0035	0.0035	0.0012
Non-white Male												
high school	**0.3712**	**0.3563**	**0.3760**	**0.3662**	**0.0099**	**0.3717**	**0.3723**	**0.3728**	**0.3518**	**0.3803**	**0.3660**	**0.0142**
	0.0313	0.0350	0.0318	0.0317	0.0108	0.0312	0.0312	0.0312	0.0351	0.0316	0.0316	0.0108
some college	**0.4265**	**0.4292**	**0.4248**	**0.4270**	**0.0022**	**0.4281**	**0.4244**	**0.4265**	**0.4109**	**0.4323**	**0.4216**	**0.0107**
	0.0315	0.0352	0.0320	0.0319	0.0108	0.0314	0.0314	0.0314	0.0353	0.0318	0.0318	0.0109
college	**0.8339**	**0.8418**	**0.8300**	**0.8359**	**0.0059**	**0.8335**	**0.8317**	**0.8324**	**0.8161**	**0.8387**	**0.8274**	**0.0113**
	0.0419	0.0468	0.0425	0.0424	0.0141	0.0418	0.0418	0.0417	0.0468	0.0423	0.0424	0.0140
graduate school	**1.1091**	**1.1506**	**1.0945**	**1.1226**	**0.0280**	**1.1025**	**1.1047**	**1.1020**	**1.1102**	**1.1025**	**1.1064**	**0.0039**
	0.0430	0.0482	0.0437	0.0436	0.0146	0.0430	0.0429	0.0429	0.0482	0.0435	0.0436	0.0145
annual return	**0.1003**	**0.1128**	**0.0958**	**0.1043**	**0.0085**	**0.0987**	**0.1000**	**0.0991**	**0.1094**	**0.0968**	**0.1031**	**0.0063**
5-10 years exp	0.0086	0.0096	0.0088	0.0087	0.0030	0.0086	0.0086	0.0086	0.0096	0.0088	0.0087	0.0030

	Spec 1	Spec2				Spec3	Spec4	Spec5	Spec6			
	main	DER	SIPP	Average	abs value dev.	main	main	main	DER	SIPP	Average	abs value dev.
White Female												
high school	**0.3333**	**0.3595**	**0.3228**	**0.3411**	**0.0184**	**0.3304**	**0.3325**	**0.3305**	**0.3504**	**0.3239**	**0.3371**	**0.0133**
	0.0141	0.0158	0.0144	0.0143	0.0048	0.0141	0.0141	0.0141	0.0159	0.0143	0.0143	0.0049
some college	**0.4063**	**0.4239**	**0.3987**	**0.4113**	**0.0126**	**0.4041**	**0.4059**	**0.4045**	**0.4164**	**0.4005**	**0.4085**	**0.0080**
	0.0138	0.0154	0.0140	0.0139	0.0047	0.0138	0.0137	0.0137	0.0155	0.0139	0.0139	0.0048
college	**0.8472**	**0.8745**	**0.8360**	**0.8553**	**0.0192**	**0.8442**	**0.8462**	**0.8446**	**0.8633**	**0.8392**	**0.8512**	**0.0120**
	0.0168	0.0186	0.0170	0.0170	0.0056	0.0167	0.0167	0.0167	0.0188	0.0169	0.0170	0.0056
graduate school	**1.0605**	**1.0836**	**1.0513**	**1.0675**	**0.0161**	**1.0568**	**1.0617**	**1.0594**	**1.0809**	**1.0538**	**1.0673**	**0.0136**
	0.0179	0.0199	0.0182	0.0181	0.0059	0.0178	0.0178	0.0178	0.0201	0.0180	0.0181	0.0060
annual return	**0.0853**	**0.0914**	**0.0830**	**0.0872**	**0.0042**	**0.0844**	**0.0850**	**0.0846**	**0.0893**	**0.0835**	**0.0864**	**0.0029**
5-10 years exp	0.0033	0.0037	0.0034	0.0033	0.0011	0.0033	0.0033	0.0033	0.0037	0.0034	0.0033	0.0011
Non-white Female												
high school	**0.2707**	**0.2901**	**0.2631**	**0.2766**	**0.0135**	**0.2684**	**0.2670**	**0.2662**	**0.2748**	**0.2643**	**0.2696**	**0.0053**
	0.0297	0.0333	0.0302	0.0301	0.0102	0.0297	0.0297	0.0296	0.0333	0.0301	0.0301	0.0102
some college	**0.3957**	**0.4149**	**0.3878**	**0.4014**	**0.0135**	**0.3931**	**0.3892**	**0.3887**	**0.3919**	**0.3888**	**0.3903**	**0.0016**
	0.0289	0.0323	0.0294	0.0292	0.0099	0.0288	0.0288	0.0288	0.0324	0.0292	0.0292	0.0099
college	**0.8657**	**0.9137**	**0.8478**	**0.8807**	**0.0330**	**0.8601**	**0.8576**	**0.8556**	**0.8771**	**0.8504**	**0.8637**	**0.0134**
	0.0376	0.0420	0.0382	0.0381	0.0126	0.0375	0.0375	0.0375	0.0421	0.0379	0.0381	0.0125
graduate school	**1.1159**	**1.2041**	**1.0852**	**1.1446**	**0.0595**	**1.1086**	**1.1119**	**1.1088**	**1.1734**	**1.0902**	**1.1318**	**0.0416**
	0.0422	0.0473	0.0428	0.0428	0.0144	0.0421	0.0421	0.0420	0.0472	0.0426	0.0427	0.0143
annual return	**0.0753**	**0.0822**	**0.0735**	**0.0779**	**0.0044**	**0.0744**	**0.0748**	**0.0742**	**0.0805**	**0.0735**	**0.0770**	**0.0035**
5-10 years exp	0.0076	0.0085	0.0078	0.0077	0.0027	0.0076	0.0076	0.0076	0.0085	0.0078	0.0077	0.0027

Figure 1
Labor Force Experience Profiles, Job Level:
White Men

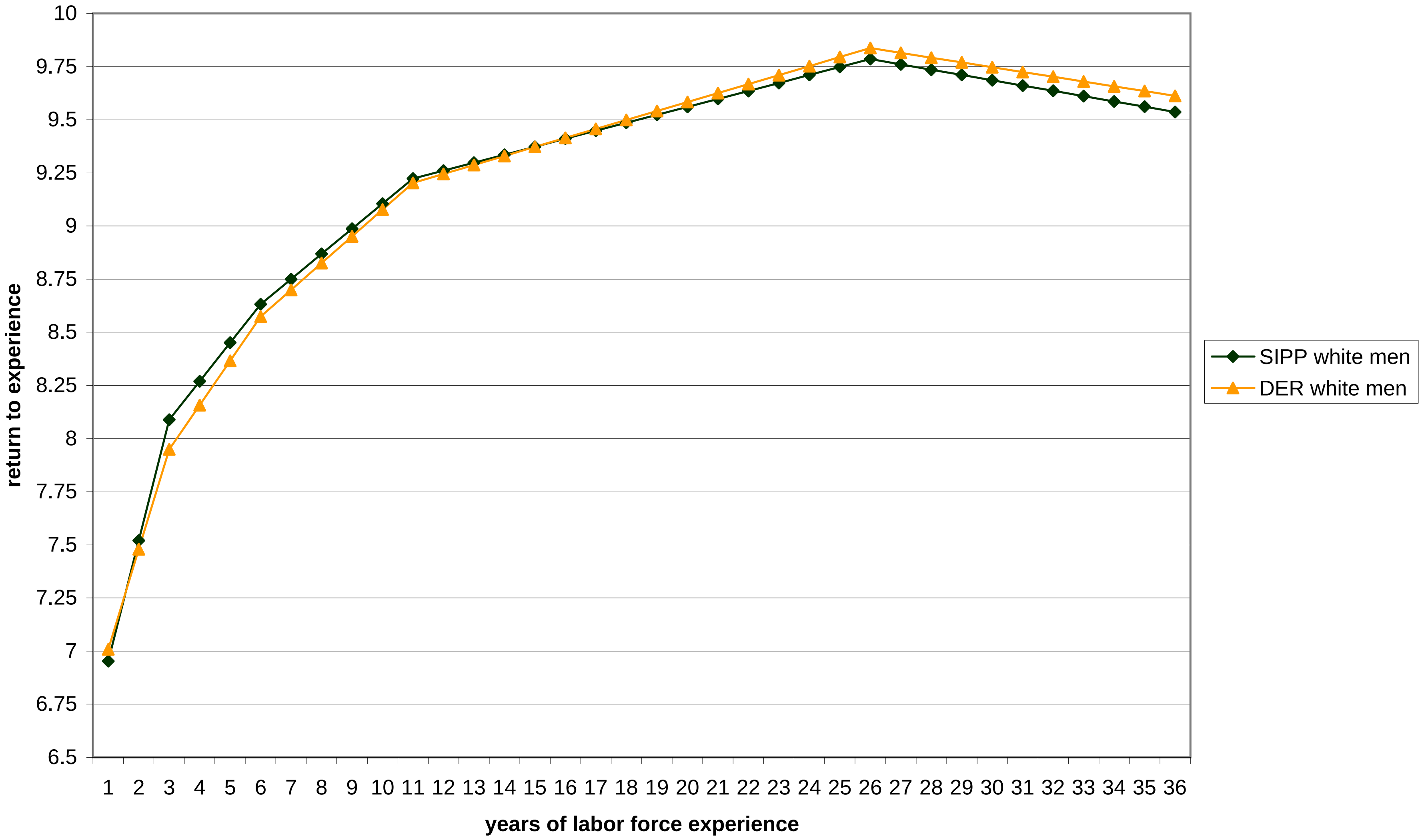


Figure 2
Labor Force Experience Profiles, Job Level:
White Women

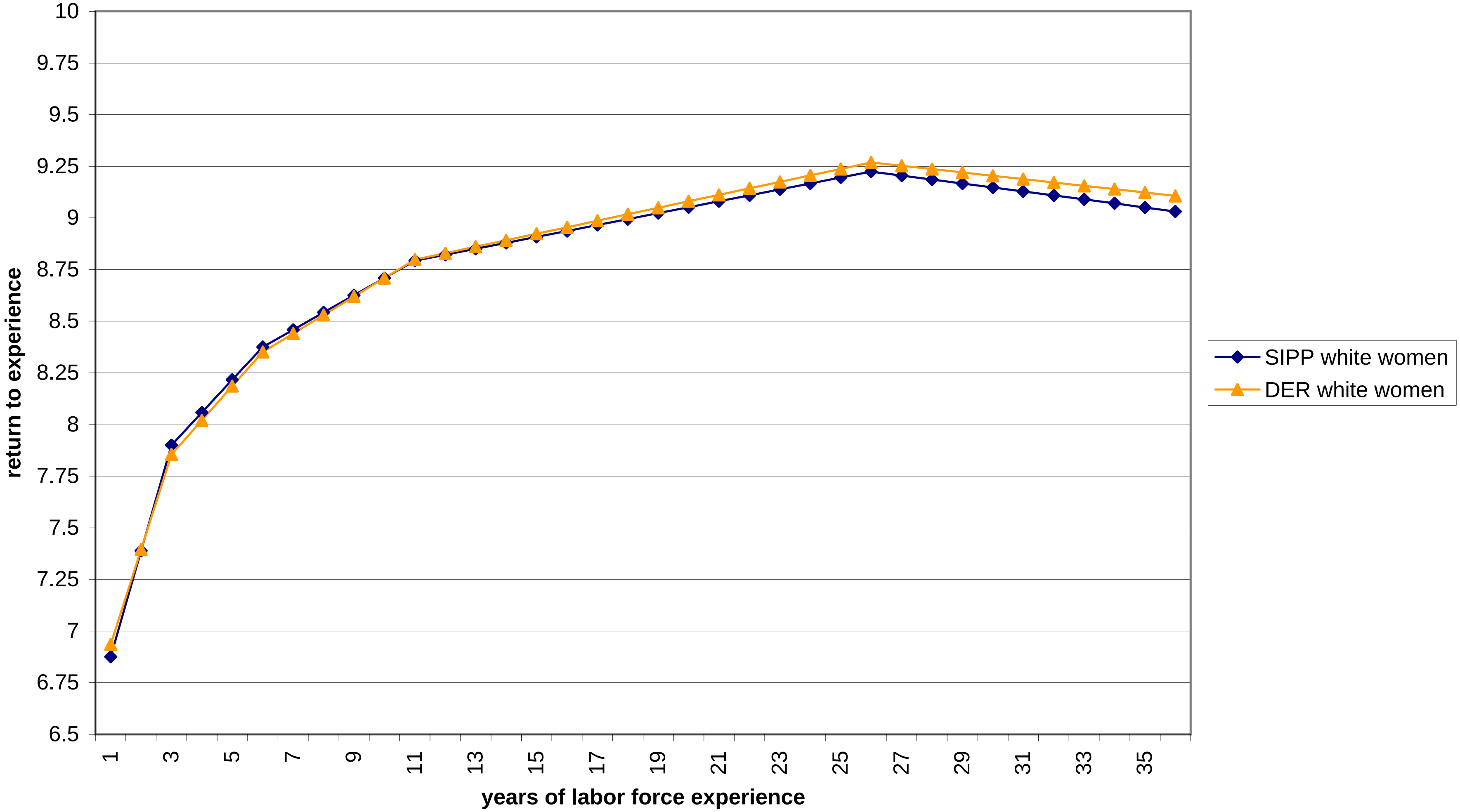


Figure 3
Labor Force Experience Profiles, Job Level:
Non-white Men

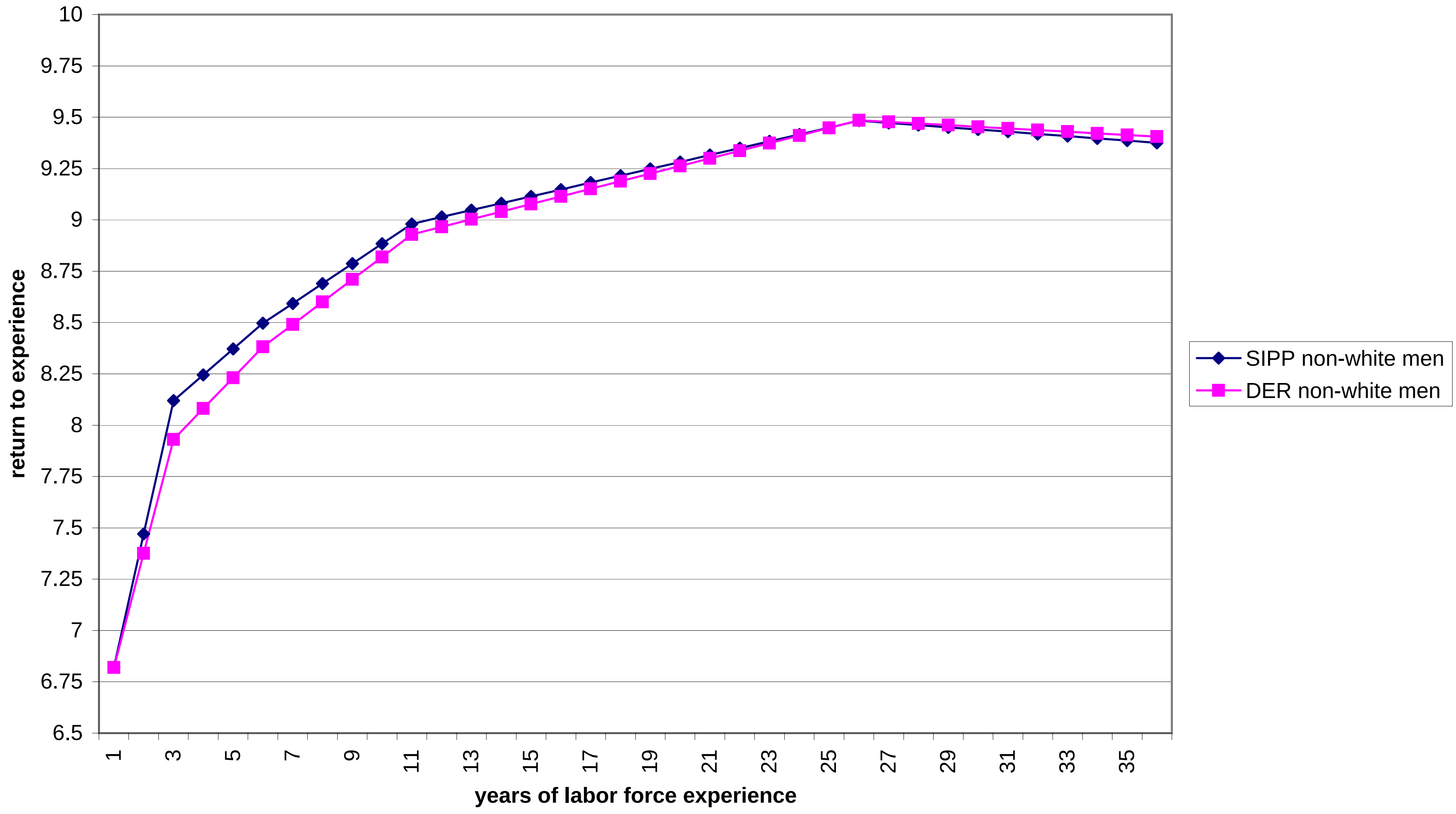


Figure 4
Labor Force Experience Profiles, Job Level:
Non-white Women

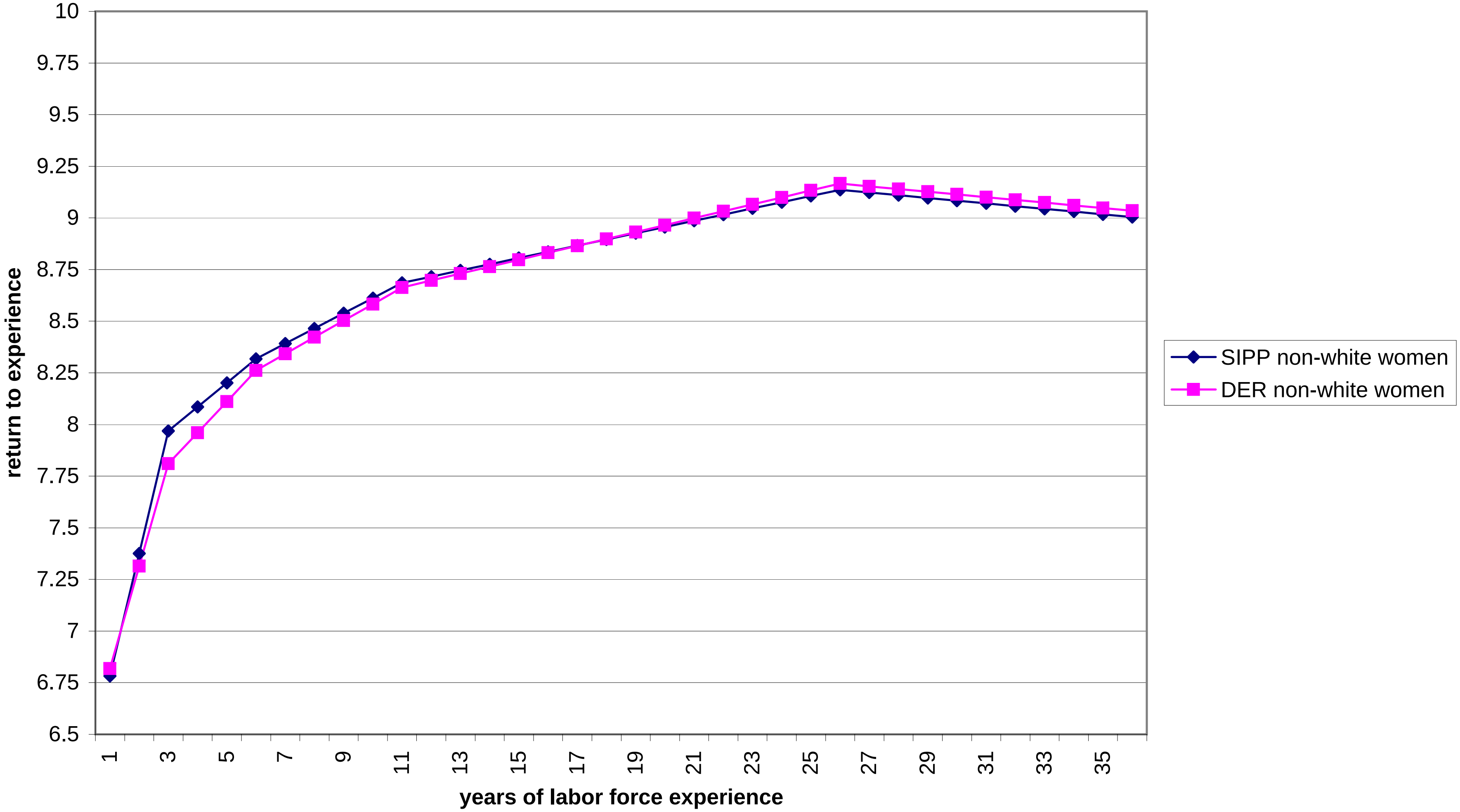


Table 17: Reliability Ratios by Demographic Group								
	White Males		Non-white Males		White Females		Non-white Females	
	Education level		Education level		Education level		Education level	
	<=12 Yrs	>12 Yrs	<=12 Yrs	>12 Yrs	<=12 Yrs	>12 Yrs	<=12 Yrs	>12 Yrs
κ_{sipp}	0.8885	0.8937	0.8695	0.8776	0.8937	0.8909	0.8766	0.8818
κ_{der}	0.8354	0.8438	0.8091	0.8170	0.8497	0.8476	0.8225	0.8264

Table 18: Variances of Random Effects

		Spec1	Spec2	Spec 3	Spec6
person effect	$\sigma^2_{\theta 1}$ (avg. intercept)	0.5781	0.5719	0.4995	0.4956
	$\sigma^2_{\theta 1Error}$ (dev. from avg.)			0.1861	0.1850
	$\sigma^2_{\theta 2}$ (avg. slope)	0.3176	0.3161	0.2951	0.2944
	$\sigma^2_{\theta 2Error}$ (dev. from avg.)			0.0380	0.0376
	$\sigma_{\theta 1\theta 2}$ (covar of avgs.)	-0.3793	-0.3766	-0.3435	-0.3421
	$\sigma_{\theta 1E\theta 2E}$ (covar of devs.)			-0.0806	-0.0800
total true var. person eff.		0.3239	0.3227	0.2977	0.2967
total error var. person eff.				0.0197	0.0196
common error	σ^2_{η}	0.4867	0.4882	0.5340	0.5353
	ρ	0.6257	0.6267	0.6607	0.6614
SIPP error	σ^2_{ω}	0.2021	0.2019	0.2122	0.2123
	ρ_{SIPP}	0.1328	0.1318	0.1793	0.1792
DER error	σ^2_{ν}	0.2770	0.2760	0.1854	0.1852
	ρ_{DER}	0.6063	0.6051	0.4091	0.4087
SIPP Reliability Ratio		0.7998	0.8001	0.7815	0.7816
St err of RR		0.0013	0.0013	0.0017	0.0017
DER Reliability Ratio		0.7446	0.7453	0.8017	0.8021
St err of RR		0.0013	0.0013	0.0022	0.0022
SIPP First Diff. Rel. Ratio		0.6321	0.6315	0.6176	0.6172
DER First Diff. Rel. Ratio		0.7341	0.7337	0.71967	0.7196

Table19: Fixed Effects: Intercepts, Panels, and Time

		Spec 1	Spec 2	Spec 3	Spec 6
SIPP Intercept	β_{0SIPP}	6.5521	6.5544	6.5368	6.5417
		0.0196	0.0209	0.0195	0.0213
DER Intercept	β_{0DER}	6.7392	6.72751	6.7241	6.7120
		0.0195	0.0208	0.0193	0.0219
time effect	β_2 (time)	-0.0018	-0.0021	-0.0014	-0.0017
		0.0008	0.0008	0.0008	0.0008
SIPP 1990 panel	β_{31990}	0.0478	0.0434	0.0500	0.0456
		0.0042	0.0042	0.0041	0.0042
SIPP 1991 panel	β_{31991}	0.0325	0.0288	0.0337	0.0300
		0.0049	0.0049	0.0049	0.0049
SIPP 1992 panel	β_{31992}	0.0449	0.0403	0.0458	0.0411
		0.0040	0.0040	0.0039	0.0039
SIPP 1993 panel	β_{31993}	0.0460	0.0420	0.0470	0.0430
		0.0043	0.0044	0.0043	0.0043

*coefficients in first row, standard errors in second row for each effect

Table 20: Fixed Effects: Race and Gender

	Spec 1	Spec2				Spec3	Spec6			
	Main	DER	SIPP	Average	abs value dev.		DER	SIPP	Average	abs value dev.
Non-white Males	**-0.2306**	**-0.1654**	**-0.2889**	**-0.2271**	**0.0618**	**-0.2289**	**-0.1690**	**-0.2768**	**-0.2229**	**0.0539**
	0.0528	0.0566	0.0563			0.0523	0.0596	0.0576	0.0523	0.0263
White Females	**-0.1380**	**-0.1150**	**-0.1536**	**-0.1343**	**0.0193**	**-0.1398**	**-0.1161**	**-0.1541**	**-0.1351**	**0.0190**
	0.0271	0.0291	0.0289			0.0269	0.0307	0.0295	0.0269	0.0135
Non-white Females	**-0.1572**	**-0.1240**	**-0.1796**	**-0.1518**	**0.0278**	**-0.1549**	**-0.1252**	**-0.1729**	**-0.1491**	**0.0239**
	0.0493	0.0528	0.0523			0.0487	0.0555	0.0535	0.0488	0.0243

Table 21: Fixed Effects: Education and Experience

	Spec 1	Spec2				Spec3	Spec6			
	main	DER	SIPP	Average	abs value dev.	main	DER	SIPP	Average	abs value dev.
White Male										
high school	**0.3801**	**0.4042**	**0.3656**	**0.3849**	**0.0193**	**0.3788**	**0.4031**	**0.3653**	**0.3842**	**0.0189**
	0.0112	0.0121	0.0115			0.0112	0.0122	0.0115	0.0113	0.0037
some college	**0.5220**	**0.5333**	**0.5147**	**0.5240**	**0.0093**	**0.5209**	**0.5323**	**0.5142**	**0.5232**	**0.0091**
	0.0112	0.0121	0.0115			0.0112	0.0122	0.0115	0.0112	0.0037
college	**0.9694**	**0.9987**	**0.9524**	**0.9755**	**0.0232**	**0.9675**	**0.9968**	**0.9520**	**0.9744**	**0.0224**
	0.0134	0.0145	0.0138			0.0134	0.0146	0.0138	0.0135	0.0043
graduate school	**1.0689**	**1.0896**	**1.0569**	**1.0732**	**0.0164**	**1.0678**	**1.0862**	**1.0578**	**1.0720**	**0.0142**
	0.0140	0.0151	0.0144			0.0140	0.0152	0.0143	0.0141	0.0045
annual return	**0.0976**	**0.0959**	**0.0996**	**0.0977**	**0.0019**	**0.0975**	**0.0959**	**0.0995**	**0.0977**	**0.0018**
5-10 years exp	0.0030	0.0033	0.0032			0.0030	0.0033	0.0032	0.0030	0.0011
Non-white Male										
high school	**0.4031**	**0.4081**	**0.3989**	**0.4035**	**0.0046**	**0.4003**	**0.4043**	**0.3973**	**0.4008**	**0.0035**
	0.0269	0.0291	0.0278			0.0269	0.0292	0.0277	0.0270	0.0089
some college	**0.6196**	**0.6420**	**0.6061**	**0.6240**	**0.0179**	**0.6167**	**0.6374**	**0.6054**	**0.6214**	**0.0160**
	0.0275	0.0297	0.0284			0.0275	0.0299	0.0283	0.0276	0.0091
college	**1.0416**	**1.0406**	**1.0414**	**1.0410**	**0.0004**	**1.0386**	**1.0353**	**1.0402**	**1.0378**	**0.0025**
	0.0366	0.0395	0.0376			0.0366	0.0396	0.0375	0.0367	0.0117
graduate school	**1.2901**	**1.2769**	**1.2973**	**1.2871**	**0.0102**	**1.2879**	**1.2703**	**1.2971**	**1.2837**	**0.0134**
	0.0376	0.0407	0.0386			0.0376	0.0408	0.0385	0.0378	0.0121
annual return	**0.0937**	**0.0986**	**0.0916**	**0.0951**	**0.0035**	**0.0939**	**0.0989**	**0.0919**	**0.0954**	**0.0035**
5-10 years exp	0.0075	0.0081	0.0079			0.0075	0.0081	0.0078	0.0075	0.0026

	Spec 1	Spec2				Spec3	Spec6			
	main	DER	SIPP	Average	abs value dev.	main	DER	SIPP	Average	abs value dev.
White Female										
high school	**0.4387**	**0.4634**	**0.4228**	**0.4431**	**0.0203**	**0.4363**	**0.4581**	**0.4234**	**0.4407**	**0.0173**
	0.0125	0.0135	0.0129			0.0125	0.0135	0.0128	0.0125	0.0042
some college	**0.6285**	**0.6485**	**0.6153**	**0.6319**	**0.0166**	**0.6265**	**0.6437**	**0.6158**	**0.6297**	**0.0140**
	0.0123	0.0133	0.0126			0.0122	0.0133	0.0126	0.0123	0.0041
college	**1.0553**	**1.0735**	**1.0423**	**1.0579**	**0.0156**	**1.0529**	**1.0664**	**1.0431**	**1.0548**	**0.0117**
	0.0147	0.0159	0.0152			0.0147	0.0160	0.0151	0.0148	0.0048
graduate school	**1.2684**	**1.2875**	**1.2554**	**1.2714**	**0.0161**	**1.2661**	**1.2810**	**1.2563**	**1.2687**	**0.0123**
	0.0156	0.0169	0.0161			0.0156	0.0170	0.0160	0.0157	0.0051
annual return	**0.0584**	**0.0529**	**0.0620**	**0.0575**	**0.0046**	**0.0583**	**0.0528**	**0.0618**	**0.0573**	**0.0045**
5-10 years exp	0.0029	0.0032	0.0031			0.0029	0.0032	0.0031	0.0029	0.0010
Non-white Female										
high school	**0.3489**	**0.3727**	**0.3340**	**0.3534**	**0.0194**	**0.3475**	**0.3690**	**0.3351**	**0.3521**	**0.0170**
	0.0260	0.0281	0.0268			0.0260	0.0282	0.0267	0.0261	0.0085
some college	**0.6108**	**0.6258**	**0.6014**	**0.6136**	**0.0122**	**0.6098**	**0.6226**	**0.6025**	**0.6125**	**0.0101**
	0.0255	0.0276	0.0263			0.0255	0.0277	0.0262	0.0256	0.0084
college	**1.1235**	**1.1417**	**1.1124**	**1.1270**	**0.0146**	**1.1229**	**1.1381**	**1.1145**	**1.1263**	**0.0118**
	0.0331	0.0358	0.0340			0.0331	0.0359	0.0339	0.0333	0.0106
graduate school	**1.2898**	**1.3312**	**1.2662**	**1.2987**	**0.0325**	**1.2882**	**1.3267**	**1.2682**	**1.2974**	**0.0293**
	0.0370	0.0400	0.0380			0.0370	0.0401	0.0379	0.0372	0.0119
annual return	**0.0773**	**0.0805**	**0.0754**	**0.0779**	**0.0025**	**0.0773**	**0.0805**	**0.0754**	**0.0779**	**0.0025**
5-10 years exp	0.0067	0.0072	0.0070			0.0067	0.0072	0.0070	0.0067	0.0023

Figure 5
Labor Force Experience Profiles, Person Level:
White Men

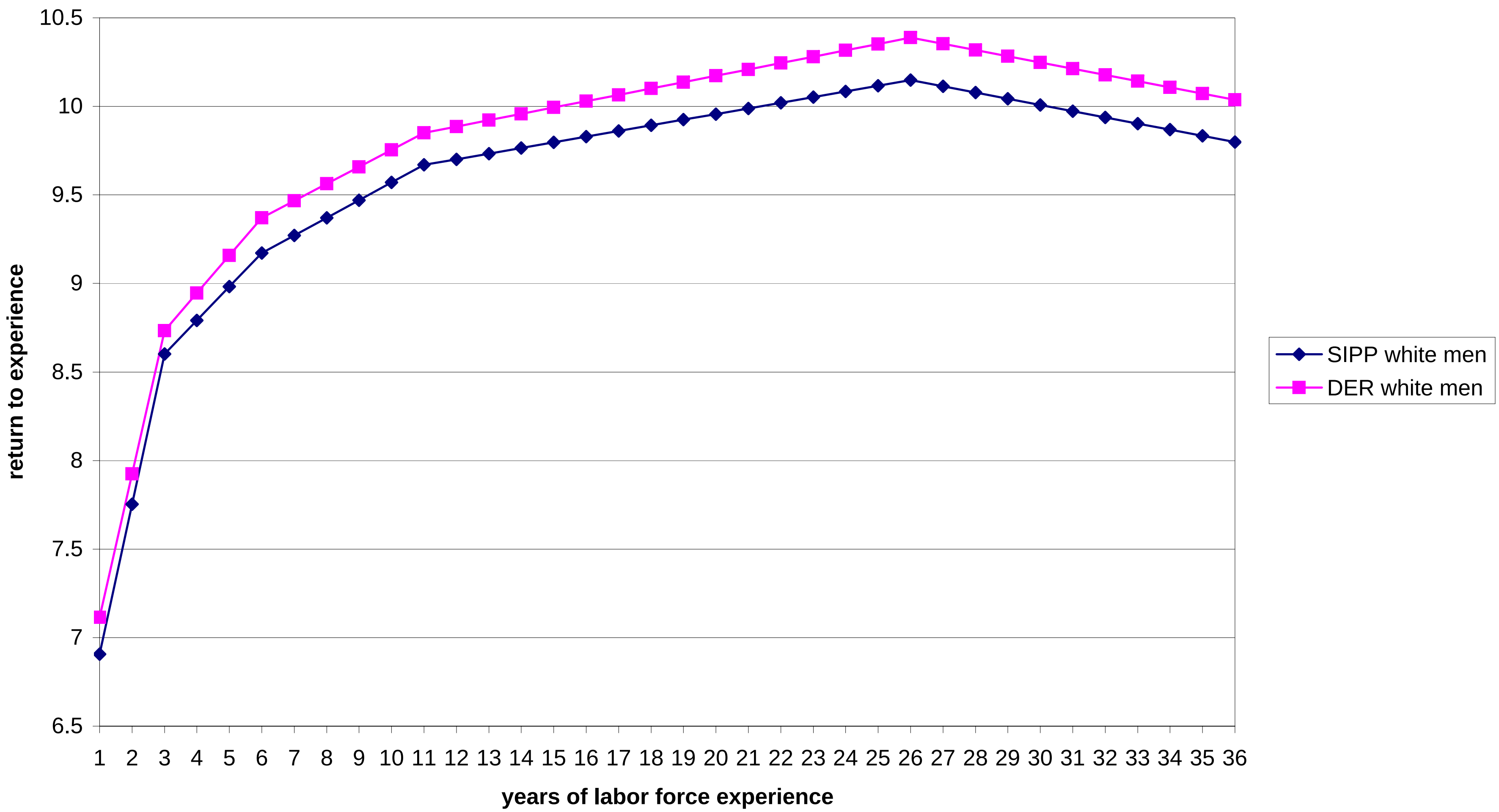


Figure 8
Labor Force Experience Profiles, Person Level:
White Women

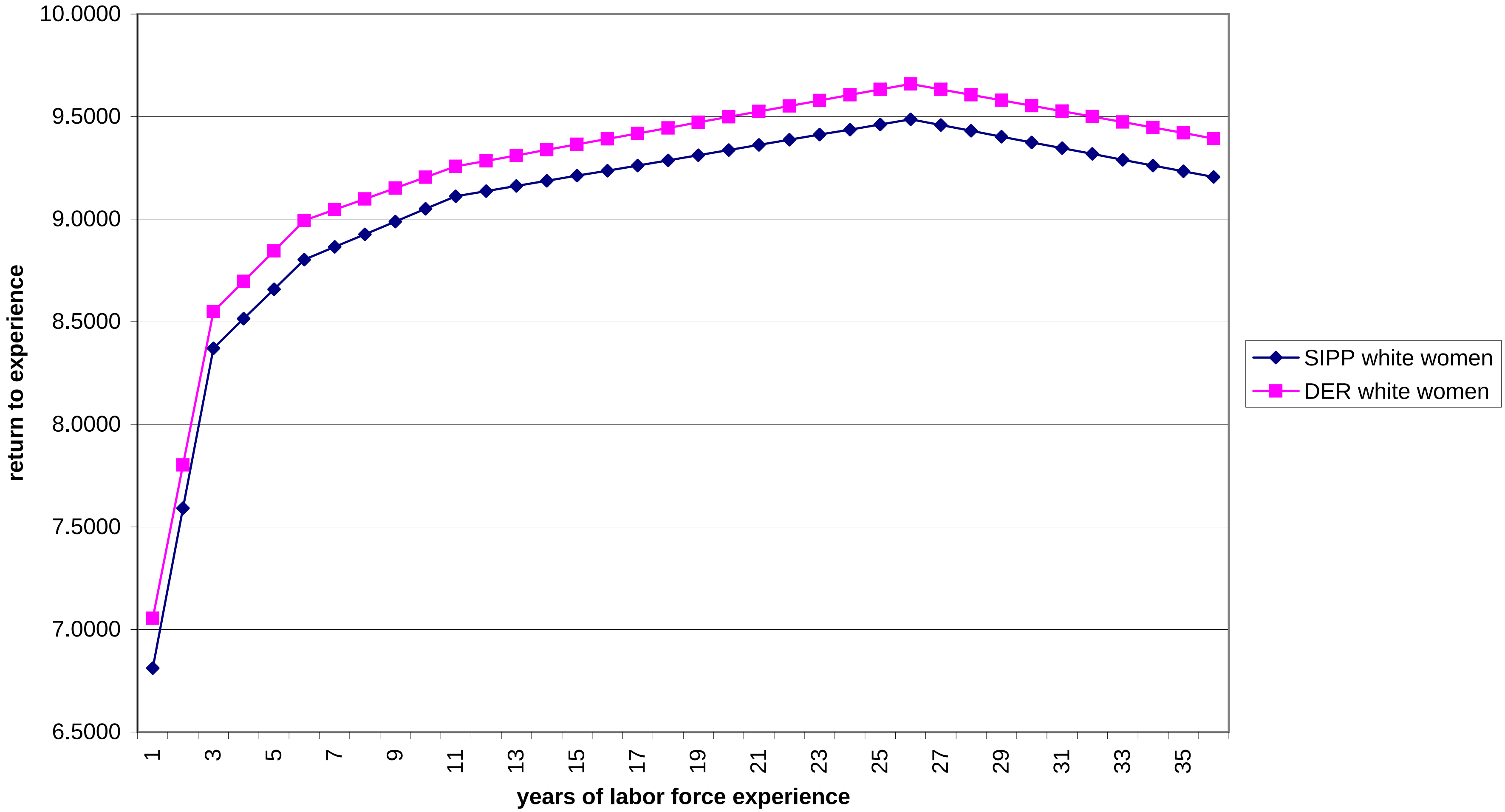


Figure 7
Labor Force Experience Profiles, Person Level:
Non-white Men

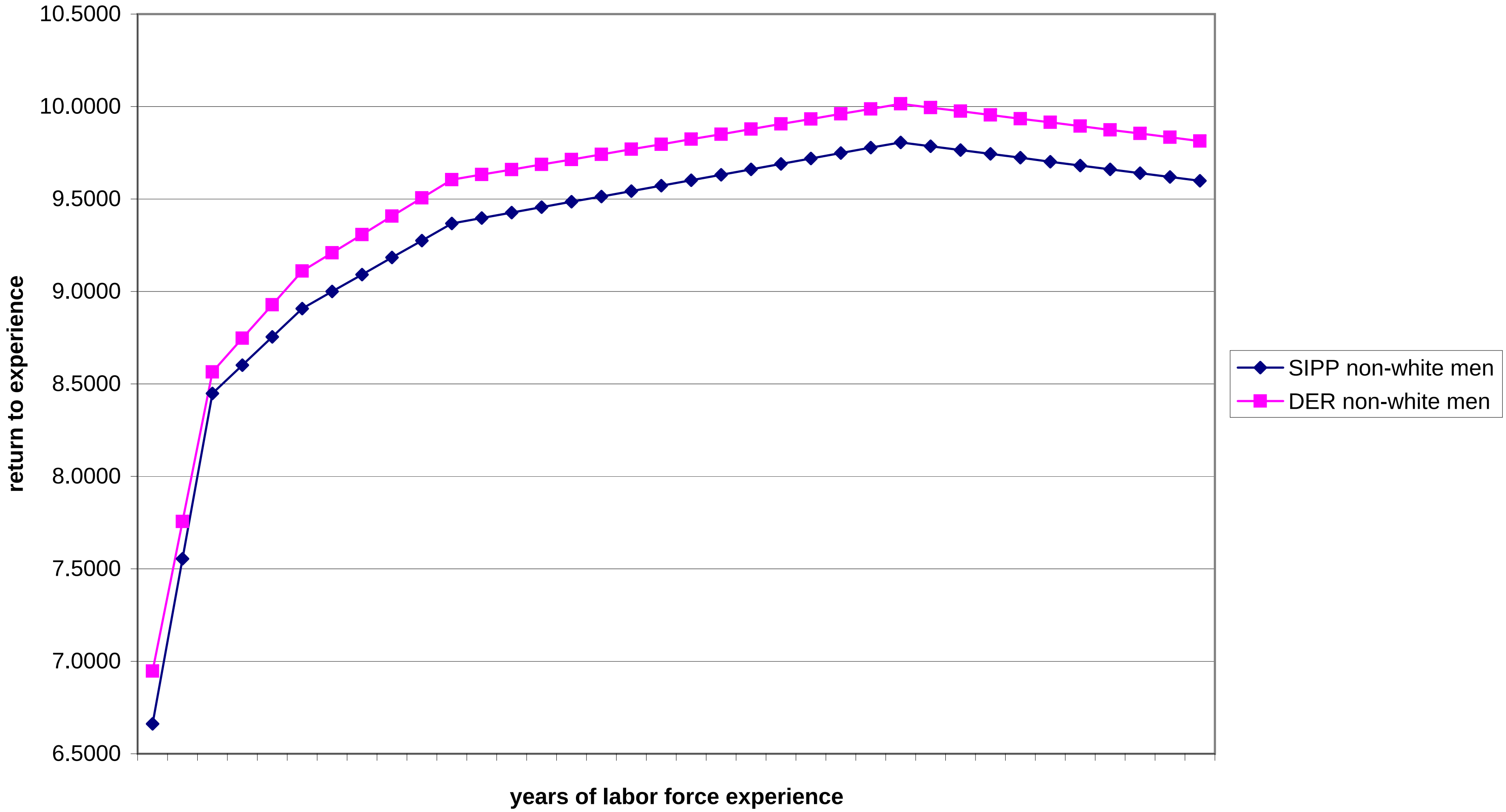
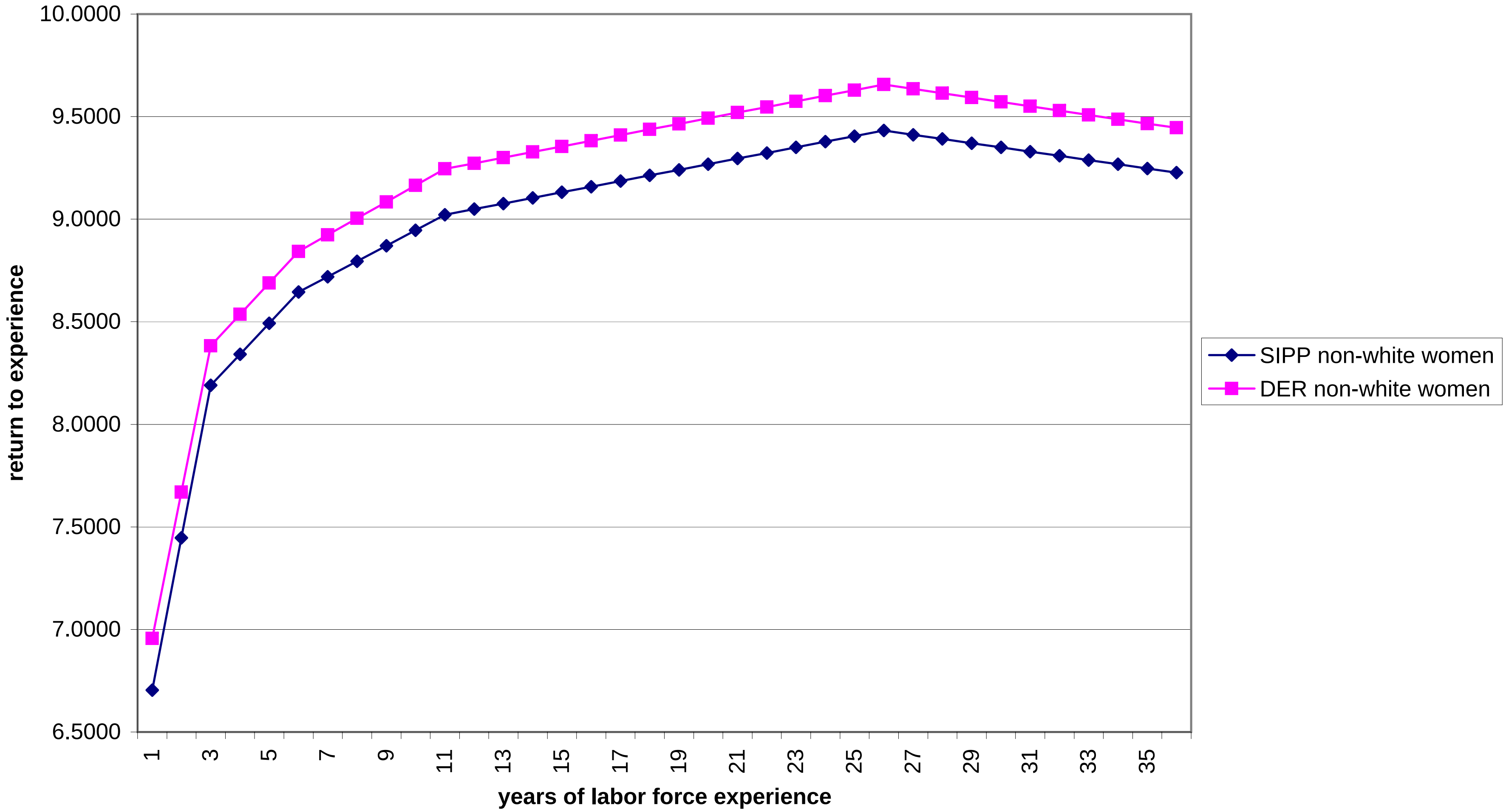


Figure 6
Labor Force Experience Profiles, Person Level:
Non-white Women



Appendix Table A1:
First Round of Job Name Matching, SIPP Panels 1990-1993
Description of Person-Job-Wave observation matching by pass

	Blocking Variables	Matching Variables	m, u prob.	cutoffs
Pass 1	Person ID	Full employer name*	.9, .02	2, .3
Pass 2	Person ID	Fields from employer name*:		2, .3
		word one	.9, .15	
		word two	.95, .6	
		word three	.95, .6	
		word four	.95, .6	
		qualifier word one	.95, .6	
		qualifier word two	.95, .6	
		type word one	.95, .6	
		type word two	.95, .6	
		SIPP original job id number	.6, .5	
Pass 3	Person ID	Fields from employer name*:		5, .05
		Array: first 4 words	.95, .6	
		word one	.9, .15	
		qualifier word one	.95, .6	
		qualifier word two	.95, .6	
		type word one	.95, .6	
		type word two	.95, .6	

*Jobs with missing names were excluded from this round of name matching

Appendix Table A1 (continued):
Second Round of Job Name Matching, SIPP Panels 1990-1993
Description of Person-Job-Wave observation matching by pass

	Blocking Variables	Matching Variables	m, u prob.	cutoffs
Pass 1	Person ID	Fields from employer name:		.1, .1
		full name**		
		array: first 4 words	.9, .1	
		Array: first 2 qualifier words	.9, .15	
		Array: first 2 type words	.9, .15	
		Geo word	.7, .5	
Pass 2	Person ID	Full employer name**	.95, .1	.1, .1

Job Name Matching, SIPP Panel 1996
Description of Person-Job record matching by pass***

	Blocking Variables	Matching Variables	m, u prob.	cutoffs
Pass 1	Person ID	Full employer name**		2, .15
Pass 2	Person ID	Fields from employer name:		2, .15
		full name**		
		word one	.9, .15	
		word two	.95, .6	
		word three	.95, .6	
		word four	.95, .6	
		qualifier word one	.95, .6	
		qualifier word two	.95, .6	
		type word one	.95, .6	
		type word two	.95, .6	
		geo word	.7, .5	
Pass 3	Person ID	Fields from employer name:		2, .15
		full name**		
		Array: first 4 words	.95, .6	
		word one	.9, .15	
		qualifier word one	.95, .6	
		qualifier word two	.95, .6	
		type word one	.95, .6	
		type word two	.95, .6	
		geo word	.7, .5	

**When no weights were assigned, complete employer name was included but given zero weight unless it was blank and then the full disagreement weight was assigned. This was used to prevent jobs with blank names from matching. If weights were assigned, full disagreement weight was also assigned if name was missing.

***Job records with observations in the same wave were disqualified from matching to each other because the same job could not be reported on twice in the same wave.

Appendix Table B1: DER Match to the Business Register

SIPP Panel		DER Total	Match to					
			Business Register		Single-Unit File		Multi-Unit File	
1990	EINs	60,131	58,991	98.10%	58,255	96.88%	16,990	28.25%
	Jobs	96,086	93,520	97.33%	92,379	96.14%	37,807	39.35%
1991	EINs	38,628	38,096	98.62%	37,686	97.56%	12,497	32.35%
	Jobs	58,020	56,725	97.77%	56,118	96.72%	24,526	42.27%
1992	EINs	62,406	61,391	98.37%	60,777	97.39%	19,361	31.02%
	Jobs	99,524	96,982	97.45%	96,029	96.49%	43,197	43.40%
1993	EINs	51,880	50,839	97.99%	50,376	97.10%	17,029	32.82%
	Jobs	81,320	78,933	97.06%	78,198	96.16%	35,977	44.24%
1996	EINs	105,095	95,122	90.51%	94,438	89.86%	28,923	27.52%
	Jobs	192,720	172,832	89.68%	171,585	89.03%	82,546	42.83%

Appendix Table C1:
Description of SIPP Job to DER Job Matching Algorithm by Pass

	Blocking Variables	Matching Variables	m, u prob.	cutoffs
Pass 1	Person ID	Fields from SU name:		2, .3
		Array: first 4 words	.95, .1	
		Array: first 2 qualifier words	.9, .3	
		Array: first 2 type words	.9, .3	
		Geo word	.7, .5	
		year indicators*	.75, .3	
		Complete SU name**		
Pass 2	Person ID	Fields from MU name:		2, .3
		Array: first 4 words	.95, .1	
		Array: first 2 qualifier words	.9, .3	
		Array: first 2 type words	.9, .3	
		Geo word	.7, .5	
		year indicators*	.75, .3	
		Complete MU name**		
Pass 3	Person ID	year indicators*	.9, .3	2, .3
	3-digit SU Industry			
Pass 4	Person ID	year indicators*	.9, .3	2, .3
	3-digit MU Industry			
Pass 5	Person ID	start year***	.9, .3	2, .3
	3-digit SU Industry			
Pass 6	Person ID	year indicators*	.9, .3	2, .3
	1-digit SU Industry			
Pass 7	Person ID	year indicators*	.9, .1	2, .3
		3-digit SU Industry	.9, .1	2, .3

*Year Indicators by Panel
1990: 1990, 1991, 1992
1991: 1991, 1992, 1993
1992: 1992, 1993, 1994, 1995
1993: 1993, 1994, 1995
1996: 1996, 1997, 1998, 1999
**Complete employer name was included but given zero weight unless it was blank and then the full disagreement weight was assigned. This was used to prevent jobs with blank names from matching in the first 2 passes.
***Start year was first year during survey time frame when job was observed in the SIPP or DER.

Appendix Table C2: SIPP Jobs matched to DER Jobs

	1990		1991		1992		1993		1996	
	SIPP Jobs	DER Jobs	SIPP Jobs	DER Jobs	SIPP Jobs	DER Jobs	SIPP Jobs	DER Jobs	SIPP Jobs	DER Jobs
1 Master Match	37,207	37,207	22,513	22,513	34,593	34,593	32,289	32,289	66,366	66,366
2 Clerical Match	4,678	4,678	2,745	2,745	5,136	5,136	4,180	4,180	8,476	8,476
3 Duplicate Match on SIPP side	529		294		490		436		971	
4 Duplicate Match on DER side		907		609		1,007		771		1,920
5 Total Matches	42,414	42,792	25,552	25,867	40,219	40,736	36,905	37,240	75,813	76,762
6 Match Rate	76.99%	48.45%	78.75%	48.99%	77.87%	45.08%	77.33%	50.11%	78.04%	44.21%
7 Percent of Matches that are Master	87.72%	86.95%	88.11%	87.03%	86.01%	84.92%	87.49%	86.71%	87.54%	86.46%
8 Residual Job (non-match)	12,673	45,532	6,895	26,930	11,431	49,624	10,818	37,077	21,336	96,861
9 Total Jobs	55,087	88,324	32,447	52,797	51,650	90,360	47,723	74,317	97,149	173,623

Appendix Table C3: Example of Duplicate Match on SIPP Side

Type of Match	DER EIN	SIPP Jobnum
Master Match	A	1
Duplicate Match on SIPP side	A	2

Appendix Table C4: Example of Duplicate Match on DER Side

Type of Match	DER EIN	SIPP Jobnum
Master Match	A	1
Duplicate Match on DER side	B	1