Working Paper Series
Congressional Budget Office
Washington, D.C.

MEASUREMENT ERROR IN THE SIPP: EVIDENCE FROM MATCHED ADMINISTRATIVE RECORDS

Julian Cristia (E-mail: julian.cristia@cbo.gov)
Jonathan A. Schwabish (E-mail: jonathan.schwabish@cbo.gov)
Congressional Budget Office
Washington, D.C.

January 2007
2007-03

The authors wish to thank participants of the 76th Annual Meeting of the Southern Economic Association. Working papers in this series are preliminary and are circulated to stimulate discussion and critical comment. These papers are not subject to CBO's formal review and editing processes. The analysis and conclusions expressed in them are those of the author and should not be interpreted as those of the Congressional Budget Office. References in publications should be cleared with the authors. Papers in this series can be obtained at www.cbo.gov/publications.

Abstract

Validation studies that compare survey-reported earnings to administrative-recorded earnings are useful to assess the extent and implications of measurement error in labor market data. While previous work typically used small restrictive samples with topcoded earnings, this paper uses data from the 1996 Survey of Income and Program Participation (SIPP) Panel matched to Social Security administrative records. This large representative sample contains uncapped administrative earnings and allows us to provide more definitive evidence on measurement error. Results show that SIPP respondents, on average, underreport earnings by a significant amount ($2,800). Consistent with previous studies we find that measurement error is negatively correlated with true earnings. Finally, unlike previous work, we find a consistent pattern between measurement error and covariates: factors positively associated with earnings are negatively correlated with measurement error.

I. Introduction

Validation studies that compare survey-reported earnings to administrative-recorded earnings, which are assumed to be error-free, are useful to assess the extent and implications of measurement error in labor market data. Previous work (Duncan and Hill, 1985; Bound and Krueger, 1991; Pedace and Bates, 2001) typically used small restrictive samples with topcoded earnings. This paper uses data from the 1996 Survey of Income and Program Participation (SIPP) matched to the Social Security Administration's Detailed Earnings Record (DER) file. Using this large representative sample containing uncapped administrative earnings allow us to provide more definitive evidence on measurement error.

Information about characteristics of measurement error is particularly useful to researchers analyzing labor market data. Such error may occur for a variety of reasons: issues related to timing, either of the survey itself or the aggregation of pay to an annual or monthly concept; the propensity of respondents to round their earnings; differences between pre- and posttax earnings; or simple respondent error. Measurement error can affect summary measures of economic status, such as the poverty rate, as well as regression results. For example, if the error is negatively correlated with true earnings (i.e., low earners overreport whereas high earners underreport), then estimates of poverty will be downward biased and inequality will be understated.

In regression models, if measurement error is "classical" (uncorrelated with true earnings and regressors), then estimates from earnings regressions will be consistent; but if earnings is an explanatory variable, its coefficient will be biased toward zero (attenuation bias). However, if measurement error is negatively correlated with true earnings, then coefficients from earnings regressions will be downward biased and attenuation bias will be partially (or even more than completely) offset.

Results show that SIPP respondents, on average, underreport earnings by a significant amount ($2,800).[1] Consistent with previous studies, we find that measurement error is negatively correlated with true earnings—that is, low earners tend to overreport their earnings while high earners tend to underreport. Finally, unlike previous work, we find a consistent pattern between measurement error and covariates: factors positively associated with earnings are negatively correlated with measurement error.

The paper begins with a review of the earnings validation literature and is followed by a detailed discussion of the data used in the analysis. Estimates of raw differences in the two earnings measures and decomposed by respondent characteristics are then presented; a number of regression models are estimated to statistically measure the correlation between standard demographic and economic regressors and measurement error. The paper concludes with a summary and a suggestion for future research.

II. Related Literature

In the last 20 years, a number of validation studies have analyzed the extent and consequences of earnings measurement error in survey data. Strong interest in these studies was fueled by the notion that basic empirical relationships in labor markets may be obscured due to measurement error.[2] Typically, these studies use survey data matched to either administrative or employer records, which are assumed to be error-free, and then

[1] All amounts in the paper are expressed in 2005 dollars using the Consumer Price Index for Urban Wage Earners and Clerical Workers (CPI-W).

[2] The small number of validation studies can be explained by the scarcity of data sets that match survey results with administrative or employer earnings records, coupled with the fact that many of these data sets are not publicly accessible.

analyze the properties of measurement error, which is identified as the difference between reported and administrative or employer earnings.

For example, Duncan and Hill (1985) report results from a validation study performed in a large unnamed manufacturing company where workers were interviewed using a Panel of Survey and Income Dynamics (PSID) instrument; survey answers were then compared to company records. The authors validate responses for different labor measures such as annual earnings, unemployment, fringe benefits, and hourly wages. In regard to employee-reported annual earnings, the authors find that responses are quite precise.

Bound and Krueger (1991) use Current Population Survey (CPS) data matched to Social Security records and find that the resulting measurement error is not "classical." In particular, errors are serially correlated and, for men, are negatively correlated with true earnings. Their analysis of longitudinal earnings data indicates less measurement error than in previous research, though the ratio of the signal to the total variance is lower when the data are first-differenced. Bollinger (1998), using the CPS data in a nonparametric framework, also finds that measurement error is higher in cross-sectional samples than in panel data sets. In addition, he finds significant overreporting among low earners, a conclusion similar to that found in this study.

Our analysis is most closely related to a number of studies that have used SIPP data merged to administrative earnings to analyze the extent of measurement error in the survey. Pedace and Bates (2001) use the 1992 SIPP longitudinal file matched to the Social Security payroll tax records (capped at the taxable maximum) and find that SIPP respondents tend to overreport their earnings. Consistent with previously mentioned studies, Pedace and Bates (2001) report that error is negatively correlated with administrative earnings. Finally, a substantial fraction of the variation in the error variable is explained by certain demographic characteristics.

Abowd and Stinson (2003) depart from the usual assumption that administrative data are error-free and use job data from the SIPP matched to the DER to investigate the reliability of both data files. Gottschalk and Huynh (2005) use the same data set as in this study but mainly focus on differences in patterns of measurement error across age groups. They conclude that older workers are not more likely to have measurement error than their younger counterparts. Consistent with previous studies, their results provide evidence of mean reversion (i.e., negative correlation between error and administrative earnings).

In our study, we assume that the DER corresponds to true earnings in order to investigate the empirical patterns of measurement error in the 1996 SIPP. We add to the growing literature on validation studies by using this large representative matched sample to analyze the distributional characteristics of measurement error in this setting. The table below shows that the data set used in this study is a significant improvement over data sets used in previous studies.

Study	**Survey data**	**Employer/Administrative data**	**Number of observations**	**% topcoded observations**
Duncan and Hill (1985)	Employee survey using the PSID instrument	Large anonymous manufacturing company	357	0%
Bound and Krueger (1991)	1977-1978 CPS	SSA Summary Earnings Record (Social Security payroll data)	3,389	40%
Pedace and Bates (2001)	1992 SIPP	SSA Summary Earnings Record (Social Security payroll data)	32,183	6%[3]
This study	1996 SIPP	SSA Detailed Earnings Record (IRS data)	140,269	0.7%

III. Matching Survey Data to Administrative Data

The analysis in this paper matches twelve waves of the Census Bureau's 1996 Survey of Income and Program Participation (SIPP) to the Detailed Earnings Record (DER) provided by the Social Security Administration (SSA). Each data set is discussed in turn and characteristics of the match follow.

The SIPP provides information about income and program participation of individuals and households in the United States. It contains information about cash and noncash income, taxes, assets, liabilities, demographics, labor force status and participation in government transfer programs. The survey is a continuous series of panels with sample size ranging from 14,000 to 36,700 households and was conducted annually from 1984 through 1993, and then in 1996 and 2001; the surveys range in duration from 2.5 to 4 years.[4]

The SSA DER data file includes a variety of earnings measures but only limited demographic information; hence, the SIPP is necessary to decompose differences across

[3] Not reported in Pedace and Bates (2001). Computed from the *2002 Annual Statistical Supplement to the Social Security Bulletin* (Table 4.B.1).
[4] For more details on the SIPP, see www.bls.census.gov/sipp/index.html.

groups. Earnings measures are available from 1951 to 2003 and are taken from the worker's W-2 and include IRS taxable income from wages and tips (box 1) and deferred wages (box 13).[5] The sum of these two measures generates the measure of wages and salaries used in the analysis below. Because SIPP earnings are topcoded, a similar topcode adjustment is applied to DER earnings.[6]

Matching SIPP and DER Data

Three main sets of sample restrictions are applied to the matched data. First, SIPP respondents who are missing in any month of the calendar year are dropped from the sample.[7] Second, respondents are restricted to those at least 16 years old and, third, they must be matched to the SIPP survey data. The resulting data set has 140,269 person-year observations, more than four times the number of observations in matched 1992 SIPP data.

The characteristics of the SIPP sample in each step of the merge and sample-restriction process are presented in Table 1. The raw SIPP data—for which individuals appear in any wave—have nearly 84,000 individuals, over 80 percent of whom are white and more than half married. Recalculating these averages by observation (person-year as opposed to person in column 2 of Table 1) yields only slight differences.[8] Under this

[5] SIPP self-employment earnings may not be well defined and may include income received before deductions, while administrative earnings are net of deductions (see Pedace and Bates, 2001); thus, we do not include self-employment earnings in the analysis. All the qualitative results in this study are robust to the inclusion of self-employment earnings as part of total earnings.
[6] The SIPP topcodes the earnings in each wave (a wave is four months) at $50,000, yielding an annual topcode of $150,000 (see SIPP 2001 User Guide, Appendix B). Individuals in the DER with earnings that exceed $150,000 are assigned average earnings across six sex-race cells. This procedure affects only 990 observations, or less than 1 percent of the entire sample.
[7] We apply this restriction because administrative earnings are reported on an annual basis.
[8] For each person in the SIPP, there could exist several calendar-year observations; hence, we use each calendar year as an observation.

formulation there are slightly fewer 16-29-year-olds and slightly more married respondents. The person-year data format allows for the calculation of the share of beneficiaries in the disability insurance (DI), supplemental social insurance (SSI), Old-Age and Survivors Insurance (OASI), and food stamp programs. Over 12 percent of the sample receives OASI benefits, the largest of the four beneficiary categories, representative of the U.S. population.

In the third column of Table 1, person-year observations with incomplete SIPP earnings data are dropped from the sample. This eliminates over 140,000 person-year observations, or about 22,500 individuals from the sample.[9] This sample selection has little effect on the overall demographic makeup of the sample, although average earnings increase from \$19,741 to \$20,267. The estimates in the last column are taken from the matched sample and contain 140,269 observations, or 52,297 individuals. The match rate of 85 percent is comparable to previous studies using this data set and demonstrates similar differences between the matched and unmatched samples.[10] Relative to the restricted SIPP data set in the third column, the matched data set has marginally fewer Hispanics, widows, and people receiving SSI. There is also some shifting toward younger ages as the share of people under age 50 rises by about three percentage points and the portion of people age 70 and over falls by two percentage points.

Table 2 shows differences in earnings in the two data sets by distinguishing between the share of observations with zero or positive earnings records. The table reports the cross-tabulations across the four cells for the 1992, 1996 and 2001 matched

---

[9] The roughly 140,000 dropped observations correspond not only to the 22,000 individuals completely eliminated from the sample but also those who remained in the sample and had incomplete earnings data in any particular year.

[10] The match rate is less than 100 percent because some people may not have supplied their Social Security Number or may have reported an incorrect Social Security Number.

SIPP-DER data sets.[11] For the 1996 sample, 27 percent of individuals have zero earnings in both sources; in these cases there is obviously no measurement error. In the majority of cases, both the DER and SIPP earnings records are positive while much smaller portions of the sample disagree on whether earnings are nonzero.

The percentages of each cell in Table 2 across the three SIPP panels are fairly stable. The top-left and bottom-right cells account for about 27 percent and 65 percent of the matched samples. The significantly fewer number of observations in the 1992 and 2001 panels shows one reason why the 1996 sample is preferred. Analysis of the 2001 matched SIPP sample (not reported) shows a lower match rate (about 50 percent) and appears to be correlated with observed covariates. This analysis builds on previous work by using the larger, and more representative, 1996 SIPP data file.

IV. Results

This section presents a number of different aspects of measurement error found in the matched SIPP-DER data. First, summary statistics by demographic and economic group are presented to show the general direction and magnitude of the discrepancy. Then, we check whether standard covariates and a measure of lifetime earnings volatility can predict measurement error. In order to illustrate the effect of measurement error, we estimate a basic earnings equation employing each earnings measure (SIPP and DER) separately in different regressions.

---

[11] The 1992 values are taken from Pedace and Bates (2001), who use Social Security payroll tax earnings (the Summary Earnings Record) matched to the SIPP.

Summary Statistics

The distribution of the two earnings measures is shown in Table 3; average earnings in the DER are about $2,800 larger than in the SIPP and the median is larger by a smaller amount of about $250.[12] The dispersion of DER earnings, as measured by the standard deviation, is significantly greater than SIPP earnings. The last two rows of Table 3 reflect these differences by showing summary statistics of the raw difference between the SIPP and DER and the absolute value of this difference. The latter variable shows that, on average, the absolute measurement error is about $7,000. Both measures demonstrate that there is significant measurement error in the matched data set and are generally consistent with those found elsewhere in the literature. The only exception is the work by Pedace and Bates (2001), who find that SIPP survey respondents, on average, tend to overreport their earnings by $634. This finding, however, is at least in part due to the fact that their administrative earnings measure does not include earnings from work not covered by Social Security and is topcoded at $55,000 (the OASDI taxable maximum). This topcoding problem biases the estimate of average measurement error because high earners tend to significantly underreport their earnings.

The differences between SIPP and DER earnings in our sample are explored further in Figures 1 and 2. In Figure 1, the distribution of the difference is graphed by $1,000 categories. Over one-third of SIPP respondents are within $1,000 of their DER-recorded earnings, a proportion that falls as the difference grows. Approximately 74 percent of this group, or 27 percent of the full sample, report zero earnings in both data sets. A majority of respondents (59 percent) are within $2,000 of their administrative

[12] The difference between DER and SIPP earnings drops to $2,100 if deferred wages are excluded from the DER earnings measure.

earnings record. There are spikes at either end of Figure 1 as the difference between SIPP and DER earnings exceeds $15,500. Notice also that 21 percent of respondents *over*report their earnings (between $1,000 and $15,000) while a larger portion (31.7 percent) *under*report their earnings. This difference is also reflected in Table 2, where the mean of DER earnings is larger than average SIPP earnings by $2,800.

In Figure 2, the distribution of the earnings difference is presented by $5,000 DER earnings categories. As demonstrated in previous work (Pedace and Bates, 2001 and Bollinger, 1998, for example), the average difference between SIPP and DER earnings declines (and becomes negative) as DER earnings rise, suggesting that workers with high earnings tend to underreport their earnings. (Similarly, the median value of the difference falls, although it is flatter in the lower earnings categories.) The various percentile points, also graphed in Figure 2, illustrate that *within* each earnings category, the absolute difference between the 25th percentile and the average is greater than the difference between the 75th percentile and the average. Further, the difference between the bottom and middle percentiles grows significantly faster than the difference between the top and middle, which suggests that the distribution of measurement error becomes more skewed to the right as DER earnings rise.

The evidence thus far suggests that high earners are more likely to have larger differences in their earnings. The estimates in Table 4 show the mean in the raw and absolute differences between SIPP and DER earnings by demographic (age, education) and economic (earnings quintile, OASDI) status. The table also shows differences across two measures of lifetime earnings volatility. In the first set of estimates, the difference in the two earnings measures is the largest (in both raw and absolute terms) for those near the peak of their lifetime earnings profile. As people age and work patterns change, the level of underreporting falls dramatically, from over $5,000 for 50- to 59-year-olds, to about $2,300 for 60- to 69-year-olds, and to a much smaller $550 for those 70 and older.

There are stark differences in measurement error by education status: college graduates underreport their earnings by more than $7,000 while the three other education groups underreport by much smaller amounts. In absolute terms, the difference in the two measures is nearly $15,000 for college graduates, more than twice the level for those with only some college experience. These differences persist by DER earnings quintile and are certainly correlated: workers with higher earnings, who are also more likely to have higher levels of education, are much more likely to underreport their earnings.

The last set of estimates in Table 4 shows the distribution of these differences by lifetime earnings volatility quintile.[13] Here, as earnings volatility increases (higher quintiles), the propensity to underreport earnings grows significantly. Those with the largest volatility in earnings over the past 5 years underreport their earnings by over $16,000; the average absolute value is about $24,000. These results are consistent with those seen above: people with higher earnings are more likely to have higher earnings volatility than those at the other end of the distribution.

Regression Models of Measurement Error

In this section, we explore whether standard demographic and economic

[13] For an example of an analysis on earnings variability, see Gottschalk et al. (1994). The measures of volatility used here are the variance in DER earnings and the average absolute deviation in DER earnings from the mean over the years preceding the year of observation. For example, the 5-year deviation for an observation in 1996 was estimated using DER earnings over the 5 years prior to 1996, that is, from 1991 to 1995. All earnings are used, regardless of whether the record was zero in any particular year. The sample for the earnings variance measures is restricted to those between the ages of 30 and 55.

covariates can predict measurement error in multivariate regression models. This finding is important in regressions where earnings is the left-hand-side variable; such regressions will yield biased estimates if the error in this variable is correlated with the set of regressors used.

In column 1 of Table 5, we present results of regressing the SIPP-DER variable on sex, age, race, ethnicity, marital status, and educational dummies.[14] Though only around 2% of the variation in the dependent variable can be explained, we observe a consistent pattern in the regression results: factors that are correlated with higher earnings present negative (and typically significant) coefficients. For example, prime-age (30-55) earners, non-Hispanic whites, married people and highly educated individuals tend to underreport (compared to their corresponding omitted categories). Of particular importance is the finding that college graduates tend to underreport earnings by around $5,800 when compared to high school dropouts. In a sense, the pattern uncovered in this regression should not be surprising given that typically all validation studies have reported evidence of measurement error negatively correlated with true earnings. As a consequence, factors associated with higher earnings should be correlated with negative measurement error.

Comparing our results with Pedace and Bates (2001), we see that although they report much larger R-squared values, their results tend to follow the same pattern described above. However, many of their reported coefficients are not statistically significant and the magnitudes are smaller (for example, coefficients on educational attainment). Interestingly, Bound and Krueger (1991) do not find consistent patterns when regressing their error measure on covariates. There are two factors that may explain

[14] Because the error term in these regressions is nonnormal, we implement robust standard errors.

why we observe this strong pattern, which has not been recognized in other validation studies. First, as noted in Section II, our data set is much larger than those used in previous studies. Second, the variation in our measurement error variable is higher than in much of the existing literature because we are not forced to drop or adjust as large a portion of our data set for topcoded earnings records.

In column 2 of Table 5, we present estimates for the subsample of individuals aged 30 to 55. In general, we see that focusing on this group of working age individuals, the estimated coefficients become larger in absolute value. For example, married individuals underreport their earnings by $3,000 (compared to around $1,800 in the full sample) and college graduates tend to underreport by around $7,800 when compared to high school dropouts ($5,800 in the full sample). This is not surprising given that we should expect that misreporting for individuals at ages with low attachment to the labor market would be limited.

The next set of regressions, presented in column 3, investigates whether individuals with higher earnings volatility tend to have different average earnings error. For that, we add a measure of earnings volatility (the average absolute deviation in earnings over the prior 5 years) to the model and sample presented in column 2 of Table 5. The results show that individuals with higher earnings volatility tend to underreport, although as with the other covariates included in these regressions we cannot disentangle whether this result obeys the direct effect of increased earnings volatility or rather captures that individuals with higher earnings volatility tend to have higher earning levels (as shown in Table 4).

In the remaining three columns of Table 5 we explore how much of the variation in the *dispersion* of measurement error these demographic and economic measures are able to explain. We can see that a sizable fraction of the variation is explained (5-6 percent) and, not surprisingly, that factors associated with high earnings are also correlated with higher dispersion in measurement error. In particular, prime-age white, married, and highly educated workers tend to have higher dispersion of measurement error.[15]

Finally, in Table 6 we empirically explore the impact of differential measurement by running regressions of annual earnings on educational categories and some standard controls on the full sample.[16] In column 1, using SIPP earnings, we find that male working college graduates earn around $29,000 more than male high school dropouts. Using DER earnings we estimate this difference is around $38,000. That is, using the SIPP data biased this estimate down by around 25%. Quite differently, the gap between the corresponding estimates for females is about $1,000. Note that the estimated coefficients for the age and race/ethnicity controls are fairly similar across the two data sets for both men and women. Additionally, the regressions were also estimated using log earnings as the dependent variable and the resulting estimates are similar to those shown in the table. Under this logarithmic formulation, the earnings elasticity for men (women) college graduates increases from 1.042 (1.192) when SIPP earnings are used to 1.116 (1.254) when DER earnings are used.

[15] We also estimated all regressions in Table 5 using a log-linear specification; results were qualitatively similar.

[16] The sole purpose of this exercise is to illustrate the impact of differential measurement error on estimates from a standard regression. By no means do we attempt to estimate the casual effect of education on earnings.

The results in Table 6 are obtained on the full sample that includes observations with zero earnings. To deal with the issue of earnings censored at zero, we also estimated the model using separate Tobit regressions. Under the Tobit specification, the same patterns emerge—estimated coefficients for race/ethnicity and marital status obtained using SIPP earnings as the dependent variable are very close to those computed when DER earnings are used. The difference in the estimated coefficients for male college graduate indicators increases from $9,500 in the OLS (ordinary least squares) specification ($38,441-$28,969) in Table 6 to $12,500 in the Tobit specification. For women, this difference decreases from $1,100 to $600. Finally, the uncovered patterns are also robust to running OLS regressions but dropping all individuals with zero earnings in both data sources.

An explanation for the finding in Table 6 is that as high earners (who tend to be highly educated) are more likely to underreport their earnings, the estimated difference in earnings between college and high school graduates is downward biased. The fact that we find a larger downward bias for men than for women could be expected given that the earnings distribution for men is more dispersed than for women.[17]

V. Conclusion

This paper adds to the small but growing literature of earnings validation studies by analyzing the extent and characteristics of measurement error in the SIPP. To that end, we match the 1996 SIPP Panel with the Social Security Administration DER file, which contains IRS earnings data. Our particular contribution is to use this large representative

[17] These results are robust to including only individuals with positive earnings in the regressions.

data set, in which the problem of topcoded earnings is minimal, to investigate the patterns of measurement error.

Consistent with other studies we find evidence that measurement error is not “classical” in the sense that it is negatively correlated with true earnings. A new finding from our study is that, on average, earnings are underreported by a sizable amount ($2,800). This result is driven by very high earners underreporting large amounts. Another new result is that measurement error is consistently negatively correlated with covariates that are positively correlated with earnings.

These results have important implications for labor market empirical studies. First, estimates in regressions of earnings on usual regressors (sex, race, marital status, education) will be biased due to the correlation of these regressors with measurement error. Second, as noted in previous studies, in regressions where earnings is an explanatory variable, attenuation bias due to noise is partially offset by the fact that earnings are negatively correlated with noise.

It is important to advise caution in the interpretation of these results, as well as those from other validation studies. Typically, measurement error in these studies is identified by assuming that administrative/employer earnings correspond to true earnings. If this assumption is false, then the conclusions of these analyses can be overturned. In particular, note that some results obtained in this study are also consistent with the possibility that administrative earnings contain a significant amount of noise.

For future work, it would be interesting to know, assuming that reporting error is truly random, how much noise in administrative earnings would be needed to generate the patterns observed in our study.

References

Abowd, John M. and Martha H. Stinson. 2003. "Estimating Measurement Error in SIPP Annual Job Earnings: A Comparison of Census Survey and SSA Administrative Data," mimeo (September).

Bollinger, Christopher R. 1998. "Measurement Error in the Current Population Survey: A Nonparametric Look," *Journal of Labor Economics,* Vol. 16, No. 3 (July): 576-594.

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,* Vol. 9, No. 1 (January): 1-24.

Bound, John, Charles Brown, Greg J. Duncan and Willard L. Rogers. 1994. "Evidence on the Validity of Cross-sectional and Longitudinal Labor Market Data," *Journal of Labor Economics,* Vol. 12, No. 3 (July): 345-368.

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,* Vol. 3, No. 4 (October): 508-532.

Gottschalk, Peter, Robert Moffitt, Lawrence F. Katz and William T. Dickens. 1994. "The Growth of Earnings Instability in the U.S. Labor Market," *Brookings Papers on Economic Activity,* Vol. 1994, No. 2: 217-272.

Gottschalk, Peter and Minh Huynh. 2005. "Validation Study of Earnings Data in the SIPP—Do Older Workers Have Larger Measurement Error?" *Center for Retirement Research at Boston College Working Paper 2005-07.*

Pedace, Roberto and Nancy Bates. 2001. "Using Administrative Records to Assess Earnings Reporting Error in the Survey of Income and Program Participation," *Journal of Economic and Social Measurement,* Vol. 26, No. 3-4: 173-192.

Social Security Administration. 2002. "Annual Statistical Supplement to the Social Security Bulletin," Washington, DC: U.S. Government Printing Office.

Survey of Income and Program Participation. 2001. "User's Guide," www.sipp.census.gov/sipp/usrguide/sipp2001.pdf.

Figure 1. Distribution of SIPP Earnings - DER Earnings

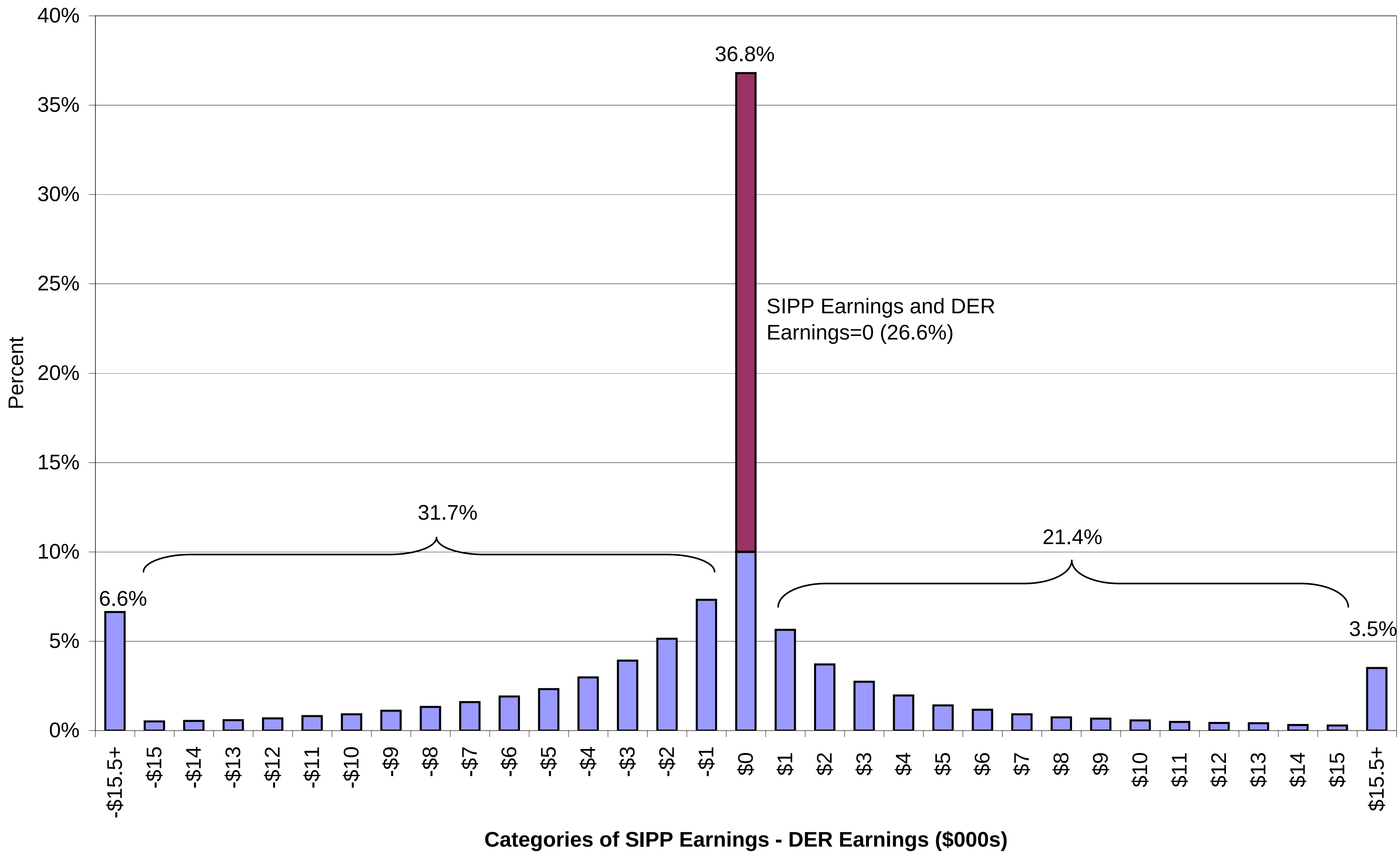

Figure 2. Distribution of SIPP Earnings - DER Earnings, by DER Earnings Category

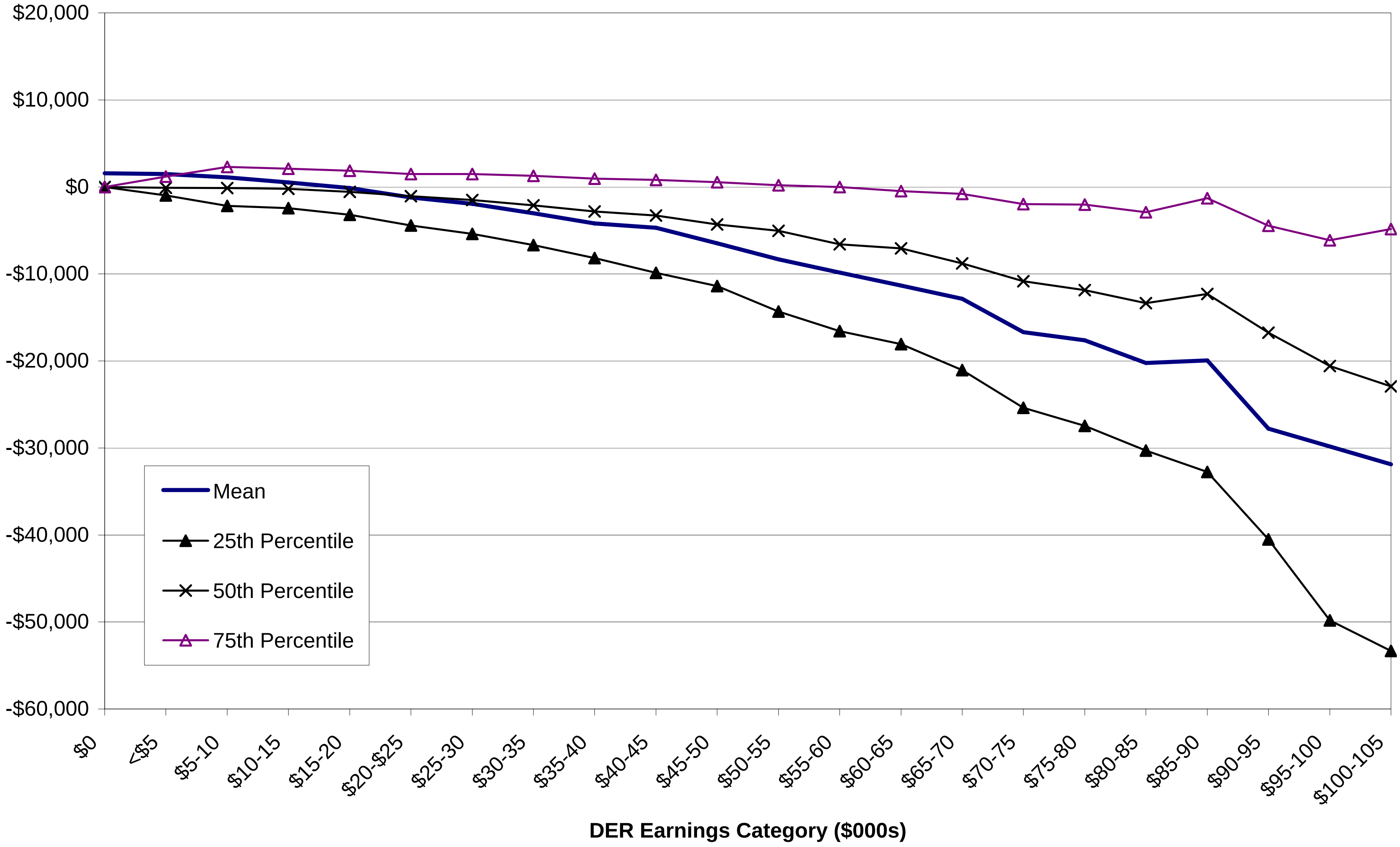

Table 1. Sample Descriptive Statistics

Sample	Individuals in the 1996 SIPP in any wave	Individuals in the 1996 SIPP that reported information in any month of a year	Individuals in the 1996 SIPP that have 12 months of earnings in the SIPP in a calendar year	Individuals in the 1996 SIPP that have 12 months of earnings in the SIPP in a calendar year and were matched to SSA administrative records
Unit of observation	Person	Person-year	Person-year	Person-year
% Male	47.3%	46.7%	46.0%	47.6%
% Aged 16-29	26.7%	21.6%	18.9%	23.2%
% Aged 30-39	19.2%	19.1%	19.3%	21.8%
% Aged 40-49	17.1%	18.3%	18.9%	21.0%
% Aged 50-59	10.9%	12.5%	13.0%	14.1%
% Aged 60-69	8.2%	8.9%	9.4%	9.9%
% Aged 70+	9.9%	11.6%	12.3%	10.0%
Average age	42.2	44.3	45.4	43.7
% White	82.9%	83.3%	84.0%	84.9%
% Black	12.5%	12.2%	11.6%	11.1%
% Nonwhite Hispanic	10.2%	9.8%	9.4%	8.3%
% Other race	4.6%	4.5%	4.4%	3.9%
% Never married	27.1%	23.1%	20.7%	22.7%
% Married	52.4%	56.1%	58.8%	58.3%
% Widowed	7.1%	7.6%	8.0%	6.5%
% Divorced or separated	13.4%	13.1%	12.5%	12.5%
% Less than high school	19.7%	18.4%	18.3%	18.6%
% High school	35.4%	33.7%	32.2%	31.1%
% Less than college	26.5%	27.8%	28.4%	28.9%
% College	18.4%	20.1%	21.0%	21.4%
% Wage/salary earnings only	--	63.3%	64.4%	67.4%
Average SIPP earnings	--	$19,741	$20,267	$20,965
Average SSA earnings	--	--	--	$23,759
% Receiving DI	--	4.6%	4.9%	4.4%
% Receiving SSI	--	3.1%	3.2%	2.4%
% Receiving food stamps	--	4.4%	4.4%	4.0%
% Receiving OASI	--	12.2%	13.0%	11.8%
Number of observations	83,957	306,501	164,942	140,269
Number of individuals	83,957	83,957	61,444	52,297

Note:
DI = Disability Insurance; SSI = Supplemental Security Income; OASI = Old-Age and Survivors Insurance.

Table 2. SIPP-DER Cross-Tabulations:
Number of Observations and Percent of Total

2001*

		DER Earnings =0	DER Earnings >0	Total
SIPP Earnings	=0	17,189	3,419	20,608
		26.9%	5.4%	32.3%
	>0	1,992	41,231	43,223
		3.1%	64.6%	67.7%
	Total	19,181	44,650	63,831
		30.0%	70.0%	100.0%

*Source: Authors' calculations.

1996*

		DER Earnings =0	DER Earnings >0	Total
SIPP Earnings	=0	37,656	8,057	45,713
		26.8%	5.7%	32.6%
	>0	4,194	90,362	94,556
		3.0%	64.4%	67.4%
	Total	41,850	98,419	140,269
		29.8%	70.2%	100.0%

*Source: Authors' calculations.

1992**

		SER Earnings =0	SER Earnings >0	Total
SIPP Earnings	=0	9,346	1,150	10,496
		29.0%	3.6%	32.6%
	>0	2,070	19,617	21,687
		6.4%	61.0%	67.4%
	Total	11,416	20,767	32,183
		35.5%	64.5%	100.0%

**Source: Pedace and Bates (2001).

Table 3. Summary Statistics: Earnings and Differences in Matched SSA-SIPP Records

Measure	Mean	Standard Deviation	25th Percentile	Median	75th Percentile
SIPP	20,965	29,581	0	12,100	32,429
DER	23,759	38,603	0	12,343	34,931
SIPP-DER	-2,795	25,406	-2,696	0	486
|SIPP-DER|	7,149	24,540	0	1,523	5,830

Table 4. Measurement Error across Different Demographic Groups

		SIPP-DER					|SIPP - DER|	
	DER Earnings	Average	Percent Positive	Average Positive	Percent Negative	Average Negative	Average	Divided by DER
Age Category								
Ages 17-29	$15,313	-$221	40.3%	$5,587	49.0%	-$5,055	$4,729	1.83
Ages 30-39	31,384	-2,921	34.5	7,930	51.5	-10,983	8,387	0.74
Ages 40-49	36,431	-5,189	31.8	8,700	51.9	-15,321	10,725	0.63
Ages 50-59	32,044	-5,195	27.1	8,945	48.0	-15,866	10,047	2.58
Ages 60-69	10,490	-2,317	13.9	7,912	27.6	-12,355	4,511	0.76
Ages 70+	1,570	-553	4.4	5,967	8.5	-9,608	1,080	0.82
Education Category								
Less than high school	6,863	-582	18.9	6,188	28.4	-6,167	2,927	0.86
High school graduate	17,930	-1,730	26.9	6,762	42.5	-8,355	5,368	0.98
Some college	23,922	-2,474	29.8	7,118	47.6	-9,653	6,723	1.00
College graduate	46,664	-7,385	32.8	10,907	48.8	-22,478	14,542	1.44
OASDI Status								
No	28,476	-3,296	33.9	7,581	50.3	-11,668	8,442	1.31
Yes	1,859	-470	7.1	4,722	13.9	-5,811	1,563	1.41
DER Earnings Quintile								
1	0	1,552	10.0	15,490	---	---	1,552	---
2	2,050	1,486	41.8	5,105	57.9	-1,121	2,784	7.23
3	12,453	727	46.4	5,720	53.5	-3,602	4,577	0.43
4	30,089	-1,552	37.3	6,316	62.5	-6,255	6,269	0.21
5	75,215	-16,221	26.1	9,492	73.9	-25,311	21,169	0.22
Earnings Variance Measures (Quintiles)								
Variance (5-year)*								
1	12,919	581	29.1	7,503	32.4	-4,940	3,785	0.73
2	26,916	-1,156	36.8	6,149	53.6	-6,375	5,681	1.01
3	30,156	-1,903	34.6	7,166	55.6	-7,883	6,860	1.33
4	34,494	-2,746	32.9	8,209	55.9	-9,738	8,147	0.41
5	65,375	-16,121	27.7	14,484	59.7	-33,724	24,135	1.02
Average Absolute Deviation (5-year)*								
1	11,914	591	28.2	7,488	30.7	-4,955	3,629	0.99
2	26,604	-1,083	36.7	6,325	53.7	-6,335	5,722	0.80
3	30,345	-1,811	35.1	7,064	55.6	-7,719	6,775	0.59
4	35,324	-2,737	33.4	8,273	56.7	-9,695	8,265	0.80
5	65,674	-16,306	27.6	14,313	60.5	-33,469	24,217	2.46

* The measures of volatility used here are the variance in DER earnings and the average absolute deviation in DER earnings from the mean over the years preceding the year of observation. For example, the 5-year deviation for an observation in 1996 was estimated using DER earnings over the 5 years prior to 1996, that is, from 1991 to 1995. All earnings are used, regardless of whether the record was zero in any particular year. The sample for the earnings variance measures is restricted to those between the ages of 30 and 55.

Table 5. Measurement Error Regressions: Earnings Variance

Dependent Variable:	SIPP-DER			|SIPP-DER|		
Sample (Age range):	**16+**	**30-55**	**30-55**	**16+**	**30-55**	**30-55**
	(1)	(2)	(3)	(4)	(5)	(6)
Sex (Men=1)	-3,661.81	-5,485.49	-5,440.47	5,657.31	8,011.98	7,949.21
	(138.00)**	(234.69)**	(233.44)**	(130.99)**	(221.87)**	(220.99)**
Age Groups						
Age (16-29) (0,1)	1,439.38	---	---	-1,960.17	---	---
	(165.81)**			(153.84)**		
Age (40-49) (0,1)	-2,077.97	-1,976.43	-1,964.58	2,079.83	1,960.03	1,943.50
	(242.61)**	(242.35)**	(242.09)**	(228.92)**	(228.45)**	(227.99)**
Age (50-59) (0,1)	-2,301.23	-2,316.64	-2,242.25	1,814.14	2,117.04	2,013.30
	(269.68)**	(323.05)**	(319.52)**	(254.92)**	(304.74)**	(300.83)**
Age (60-69) (0,1)	118.52	---	---	-2,916.33	---	---
	(243.64)			(234.00)**		
Age (70+) (0,1)	1,540.29	---	---	-5,840.98	---	---
	(200.88)**			(194.68)**		
Race, Ethnic Groups						
Black (0,1)	216.32	589.03	573.59	658.31	956.23	977.75
	(204.33)	(351.95)	(351.67)	(195.54)**	(335.72)**	(335.33)**
Nonwhite Hispanic (0,1)	382.80	668.23	653.02	-222.58	-502.25	-481.04
	(135.83)**	(233.79)**	(233.57)**	(122.01)	(211.67)*	(211.40)*
Other race (0,1)	-414.60	-25.41	-51.16	603.81	837.60	873.51
	(360.15)	(575.57)	(575.01)	(337.95)	(537.65)	(536.91)
Marital Status						
Married (0,1)	-1,791.05	-3,027.29	-3,011.46	1,785.03	3,034.37	3,012.30
	(147.45)**	(275.16)**	(274.54)**	(137.26)**	(257.43)**	(256.52)**
Widowed (0,1)	-1,816.41	-989.38	-994.29	2,413.09	1,768.30	1,775.15
	(200.60)**	(443.98)*	(443.55)*	(192.10)**	(405.56)**	(404.69)**
Separated/divorced (0,1)	-449.11	-1,606.85	-1,604.23	462.84	1,678.60	1,674.94
	(178.27)*	(281.89)**	(281.41)**	(164.94)**	(261.07)**	(260.24)**
Educational Attainment						
High school graduate (0,1)	-619.79	-1,151.91	-1,150.43	1,766.47	2,138.10	2,136.04
	(109.81)**	(178.70)**	(178.52)**	(104.48)**	(164.41)**	(164.10)**
Some college (0,1)	-1,341.50	-2,268.76	-2,236.59	2,907.79	3,616.25	3,571.39
	(115.71)**	(194.24)**	(194.20)**	(109.13)**	(179.15)**	(179.32)**
College graduate (0,1)	-5,850.35	-7,810.95	-7,740.90	10,220.91	12,625.94	12,528.26
	(250.63)**	(373.29)**	(370.69)**	(238.48)**	(352.24)**	(349.46)**
Proxy report (0,1)	-1,342.40	-1,525.64	-1,501.82	1,989.24	2,322.26	2,289.05
	(151.99)**	(246.06)**	(245.41)**	(144.25)**	(232.63)**	(231.76)**
5-year variance	---	---	-0.0723	---	---	0.1008
			(0.0272)**			(0.0401)*
Constant	2,899.96	5,761.19	5,705.51	-827.24	-4,124.76	-4,047.11
	(222.24)**	(360.95)**	(359.14)**	(209.35)**	(339.86)**	(338.03)**
Observations	140,269	73,161	73,161	140,269	73,161	73,161
R-squared	0.022	0.023	0.026	0.056	0.053	0.059

Robust standard errors in parentheses
* significant at 5%; ** significant at 1%

Table 6. Regressions of Earnings on Education Groups Dummies

Dependent Variable: Annual Earnings

	Men		Women	
	SIPP	DER	SIPP	DER
Age	2,624.32	3,389.40	1,327.13	1,463.96
	(38.18)**	(53.55)**	(20.77)**	(22.74)**
Age squared	-30.77	-38.44	-15.74	-17.13
	(0.40)**	(0.56)**	(0.22)**	(0.24)**
Race, Ethnic Groups				
Black (0,1)	-5,895.62	-7,259.63	107.93	122.04
	(422.83)**	(593.03)**	(207.43)	(227.14)
Nonwhite Hispanic (0,1)	-3,277.48	-5,085.27	-1,818.57	-1,095.30
	(455.63)**	(639.03)**	(252.88)**	(276.91)**
Other race (0,1)	-5,684.20	-6,219.69	-549.43	807.31
	(628.70)**	(881.77)**	(352.17)	(385.64)*
Educational Attainment				
High school graduate (0,1)	5,142.49	5,487.00	4,011.73	4,163.06
	(373.09)**	(523.27)**	(208.57)**	(228.39)**
Some college (0,1)	10,028.35	11,366.33	8,126.00	8,490.08
	(384.52)**	(539.30)**	(211.73)**	(231.85)**
College graduate (0,1)	28,969.40	38,441.17	20,585.49	21,644.09
	(409.43)**	(574.24)**	(235.39)**	(257.75)**
Constant	-29,918.22	-43,818.24	-15,268.29	-17,821.22
	(798.00)**	(1,119.21)**	(448.83)**	(491.47)**
Observations	66,741	66,741	73,528	73,528
R-squared	0.2224	0.1914	0.2317	0.2192

Robust standard errors in parentheses

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