Earnings dynamics and measurement error in matched survey and administrative data

Dean Hyslop* Wilbur Townsend†

2018-06-12

Abstract: This paper analyzes earnings dynamics and measurement error using a matched longitudinal sample of individuals' survey and administrative earnings. We reject that survey reports are measured with classical error and the administrative earnings are error free. The reported differences are negatively correlated with average administrative earnings, and with annual deviations, and characterized by both persistent and transitory factors. We formulate models for individuals' true earnings and measurement errors in each report. Assuming no measurement error in the administrative reports, we estimate mean-reverting errors in the survey report, but this finding is not robust to relaxing this assumption. The results imply measurement errors dominate the observed changes in earnings.

Keywords: Earnings dynamics, measurement error, panel data, validation study.

JEL classification codes: C33.

Acknowledgments: This research was supported in part by the Royal Society of New Zealand Marsden Fund Grant MEP1301. We thank seminar participants at the Asian and Australasian Society of Labour Economists Canberra conference, University of Technology Sydney, Victoria University of Wellington, and Dave Maré for comments and discussions, and Corey Allan for research assistance in putting together the data extract for the analysis.

*Address: Motu Economic and Public Policy Research, PO Box 24390, Wellington 6142, New Zealand. E-mail: dean.hyslop@motu.org.nz.

†Address: Institute for Economic and Policy Research, Stanford University, USA. E-mail: wilb@stanford.edu.

1 Introduction

There is an extensive literature characterising and estimating the dynamic properties of individuals' earnings (for a summary, see Meghir and Pistaferri, 2011). These models typically include a permanent random walk component of earnings, a low-order autoregressive moving-average (ARMA) transitory component, and a purely transitory component of variation to explicitly account for measurement error in reported earnings (e.g. MaCurdy, 1982; Abowd and Card, 1989; Meghir and Pistaferri, 2011).

A large literature on measurement error in survey data has also developed (for a summary see, Bound et al., 2001), and particularly in reported earnings (e.g. Bound and Krueger, 1991; Pischke, 1995; Kapteyn and Ypma, 2007; Abowd and Stinson, 2013). Using data from validation studies, which collect a second measure of earnings, the assumption of random classical measurement errors in survey reported earnings is strongly rejected when the second "validated" measure is considered true (e.g. Bound and Krueger, 1991; Pischke, 1995). These studies find that measurement error is negatively correlated with the validated true earnings, and conclude that there is mean-reversion in reported earnings. More recent studies that compare survey and administratively-collected earnings recognize that the administrative earnings measure may also contain errors (Kapteyn and Ypma, 2007; Abowd and Stinson, 2013). In fact, Kapteyn and Ypma (2007) conclude that the established finding of mean-reverting measurement error in survey earnings is not robust to allowing for error in the administrative earnings.

In this paper we exploit rich longitudinal data on individuals earnings from two sources to characterize the measurement error in each, and use this characterisation to identify and estimate the dynamics of the underlying true earnings. The data comes from Statistics New Zealand Integrated Data Infrastructure (IDI), which provides an extremely rich repository of linked data from various sources. Our primary source of data is Statistics New Zealand's Survey of Family, Income and Employment (SoFIE), which was a household panel survey over 8 annual waves from 2002/3–2009/10, and provides a panel sample of individuals' annual earnings. Our second source of earnings data is administrative linked employer employee data (LEED) within the IDI, which covers individuals' monthly earnings for the population of earners who have tax withheld at source. The SoFIE sample of individuals is matched to the LEED data to provide an 8-year panel sample of annual earnings from these two sources.

The paper makes two contributions to the literature. First, the rich longitudinal data facilitates a detailed anaysis of the measurement errors in the alternative measures of individuals' earnings, and contributes to the methodological literature on measurement error in earnings. We find that SoFIE reported earnings are 2-4% lower than LEED earnings on average, and slightly more variable. Although the average difference between the two reported earnings is small, there is substantial variability, with differences accounting for 25-30% of the variance in either source of reported earnings. We reject the joint hypthesis that SoFIE reported earnings contain classical measurement error and LEED earnings are recorded without error. A cross sectional comparison of the reported earnings implies the statistical reliability of earnings measured in LEED (0.87–0.91) is higher than in SoFIE (0.83–0.85). Consistent with the earnings validation literature we find that the differences between SoFIE and LEED earnings are negatively correlated both with individuals average (LEED) earnings over the sample and their annual transitory deviations. Longitudinally, the differences are characterized by both persistent and transitory factors; the latter consistent with a low-order autoregressive moving average (ARMA) model.

Second, by accounting for the measurment error in reported earnings, the paper provides a contribution to the substantive literature on earnings dynamics. We formulate and estimate a model for SoFIE and LEED earnings, which includes dynamics for true earnings together with specifications for the measurement errors in SoFIE and LEED. Under the conventional assumption that LEED earnings are measured without error, the estimated model is consistent with the standard result of mean-reversion in SoFIE errors. However, when we allow for measurement error in LEED earnings, this result is overturned. We estimate that measurement errors account for about 70% of the variation in male reported earnings changes, and one half of female earnings changes. Female earnings are substantially more variable than male earnings, both in levels and in changes. Permanent and transitory earnings shocks are both substantially larger for women than for men, although permanent shocks are relatively more important for men.

The remainder of the paper is organized as follows. The next section reviews the measurement error literature pertaining to earnings. Section 3 discusses the data and sample used in the empirical analysis, and the possible sources of error associated with each earnings measure. Section 4 describes both the cross-sectional and longitudinal properties of the differences between SoFIE and LEED reported earnings. In section 5, we develop and estimate a model of earnings dynamics in the presence of possible measurement errors in both sources of data. The paper concludes with a discussion of the main findings in section 6.

2 Background literature

In this section we briefly discuss the literature on the dynamics of individual earnings, and the validation study literature on survey reported earnings.

2.1 Earnings dynamics

Friedman's permanent income hypothesis (PIH, Friedman, 1957) provides the conceptual motivation for much of the modelling of income and earnings processes, with its distinction between permanent and transitory components of income. See Meghir and Pistaferri (2011) for a detailed review of modeling incomes and earnings processes.

Estimates of individuals' income or earnings processes usually specify a regression to estimate the contribution of observable factors such as education and age or experience, together with an error components model to estimate the contributions of unobserved factors. Early statistical analyses typically specified the error as consisting of a time-constant individual-specific permanent component, and a low-order stationary auto-regressive moving average

(ARMA) transitory component (e.g. Lillard and Weiss, 1979, and Hause, 1980).

Seminal empirical analyses of the longitudinal structure of US male earnings by MaCurdy (1982) and Abowd and Card (1989) have resulted in the, now standard, characterization of earnings that consists of an individual-specific non-stationary random walk permanent component, a low-order stationary ARMA transitory component, and a purely transitory component to capture (classical) measurement errors (Meghir and Pistaferri, 2011). Such non-stationary representations of permanent components are preferred both conceptually, as they capture the PIH notion of adjustment to individuals' permanent income via shocks, and statistically, because the variance of individuals' earnings and income tend to increase over the life cycle, at least in the US and UK (Meghir and Pistaferri, 2011).

In our analysis, we will follow this literature and adopt such a specification for the true earnings process. The model will then be embellished to allow for reporting errors in the observed SoFIE and LEED earnings measures.

2.2 Validation studies of measurement error in earnings

Validation studies of reported earnings generally assume that the reported earnings are potentially measured with error, and that the second "validation" report accurately measures true earnings. Two recent exceptions are Kapteyn and Ypma (2007), and Abowd and Stinson (2013), who match individuals' survey earnings to admininistrative earnings records in Sweden and the US respectively. In each case, the authors recognize that the administrative earnings may also include errors both due to the matching process and the administrative recording process, and develop methodologies that allow alternative assumptions on the reliability of the survey and administrative earnings.

Bound and Krueger (1991) analyze the properties of measurement error in reported earnings, using a sample of two matched years from the March Current Population Survey (CPS) linked to the Social Security Administration (SSA) employer-reported Social Security earnings records. They consider alternative characterisations of measurement error, including the assumption of classical and mean-reverting measurement error. Treating the SSA reports as true, Bound and Krueger show that, for males, the measurement error in CPS reported log(earnings) accounts for about 25% of earnings variation, is negatively correlated with the true earnings (correlation about -0.4) indicating mean-reversion in reported earnings, and is positively serially correlated (0.4). They also find that the measurement error is weakly correlated with covariates commonly included in earnings regression, which would cause bias in earnings regression coefficients. They conclude that the standard classical measurement error model is not appropriate for male earnings. In fact, the positive serial correlation implies that the effects of measurement error on differenced regression estimates are not as great as implied by classical measurement error.

Bound et al. (1994) and Pischke (1995) each use data from a validation study conducted using the US Panel Study of Income Dynamics (PSID) questionnaire administered to workers at a single firm and matched to the firm's payroll records (PSIDVS), to analyze the dynamic properties of the measurement error in reported earnings and the implications for modeling earnings dynamics. The PSIDVS has payroll earnings over six years, 1981–1986, and worker reported earnings for 1982 and 1986 (as well as recall reports of earnings in the intervening years). Similarly to Bound and Krueger, Bound et al. and Pischke treat the firm payroll reports as true earnings, and each conclude that errors are negatively correlated with true (i.e. payroll) earnings. For example, Pischke estimates that measurement error accounts for between 15% and 20% of reported log(earnings), and is negatively correlated with true earnings in 1982 (correlation -0.18) though weakly positively correlated in 1986. Pischke also estimates the correlation between measurement errors in 1982 and 1986 is 0.094, which is too high to be consistent with an AR(1) process with correlation coefficient 0.4.

Pischke then estimates a simple statistical model for true earnings and measurement error. The true earnings model consists of two components: a random-walk permanent component, and a pure-noise transitory component. The measurement error process consists of three components: first, a time-invariant person-specific component (unrelated to earnings);

second, a component correlated with the transitory earnings shock; and third, a pure-noise classical measurement error component. Estimating this model, he finds, consistent with Bound and Krueger's results, that the second component is negatively correlated with transitory earnings. Pischke estimates that classical measurement error accounts for about 80% of total measurement error, and the other two components each account for about 10%.

In contrast to these earlier studies, both Kapteyn and Ypma (2007) and Abowd and Stinson (2013) use survey data on individuals matched to administrative earnings records, and allow for errors in the administrative earnings. Kapteyn and Ypma (2007) analyze data from the Swedish administrative longitudinal database LINDA (Longitudinal Individual Data for Sweden), from which a sample was drawn for a validation survey in 2003. They focus on the survey and administrative reports of earnings, pensions and taxes. As well as measurement error in the validation survey responses, they allow for the possibility of measurement error in the administrative reports, assuming that any error will only be due to mismatch: with some probability an individual's record is mismatched, and drawn randomly from the LINDA population. Consistent with the studies above, when assuming no measurement error in the LINDA administrative earnings, Kapteyn and Ypma (2007) find mean-reversion in the validation survey earnings responses. However, when they allow for errors due to mismatch, this finding largely disappears, and they conclude the survey reported measurement error is almost entirely classical.

Abowd and Stinson (2013) match individuals and their jobs from the US Survey of Income Program Participation (SIPP) to the Detailed Earnings Records (DER) data from the Social Security Administration. They recognize that the administrative earnings may also include errors due to both the matching and administrative recording processes, and develop a methodology that allows alternative assumptions about the relative reliability of the survey and administrative earnings. (Gottschalk and Huynh (2010) also analyzed the SIPP/DER matched data: although they recognize that administrative data may not be error-free, their analysis implicitly assumes that the DER reported earnings are true.)

Abowd and Stinson (2013) matched both job- and person-level reported earnings to DER records, and discuss three reasons the DER earnings may be measured with error: definitional differences with survey reports; recording errors in the administrative data; and errors in matching jobs and inidividuals between the SIPP and DER data sources. They develop a methodology to characterize the measurement error properties of reported earnings under different beliefs about the reliability of each report. Assuming that the DER (respectively SIPP) reports are true, they estimate the reliability ratio of SIPP (DER) earnings is 0.78 (0.80); while giving equal weight to the alternative sources, the estimated reliability ratios of SIPP and DER earnings are 0.94 and 0.95 respectively.

The data structure in Abowd and Stinson (2013) is similar to our study. In particular, the earnings information collected in SIPP pertains to individuals' jobs, which can then be aggregated to obtain their (calendar year) annual earnings, and either job or individual-level earnings can then be matched to the administrative DER data. However, the DER job-earnings are only reported annually, compared to monthly LEED earnings reports. Thus, manipulation of both the survey collected information and administrative earnings is required to obtain comparable annual earnings from the SIPP and the DER. Analogous manipulations are required with the SoFIE earnings, however the monthly frequency of the LEED data means it can be readily matched to the annual reference periods in SoFIE.

3 Data description and selection

The data used in our analysis consists of matched data from two sources within Statistics New Zealand's Integrated Data Infrastructure. Individuals are matched across the various data sources in the IDI by the date of birth, name and sex recorded in each data source. The survey data we use is Statistics New Zealand's household panel Survey of Family, Income and Employment (SoFIE), which was conducted for 8-waves from 2002/3 until 2009/10: the first wave between October 2002 and September 2003, and the final wave in the year to September 2010. The administrative earnings data we use comes from linked employer-employee earnings data (LEED). The raw LEED data is reported in the Employer Monthly Schedules (EMS) filed by employers to Inland Revenue (IRD).

Each individual in SoFIE is allocated an annual reference period that typically corresponds to the 12-months to the end of the calendar month prior to their first-wave's interview month. Annual earnings in SoFIE are derived from spell-level information, in contrast to many households panel surveys that ask respondents directly about their annual earnings. In particular, the SoFIE survey instrument collected information on employment and gross (pre-tax) earnings for each "job spell" since the last survey. If there was a pay rate change during a job spell, the spell was split so that each sub-spell had a constant pay rate. The annual earnings contribution from each job-spell is derived using its pay rate and the length of spell within the annual period, and aggregated to give the individual's annual earnings.

The EMS covers the population of wage and salary employment earnings, as well as self-employment earnings and other monthly incomes which has had tax withheld at source. Because self-employment earnings in EMS is a self-selected subset of all self-employment, we exclude self-employment earnings, and individuals reporting self-employment in SoFIE will be dropped from the analysis below. Throughout the paper, we use SoFIE and LEED nominal annual earnings values.

We first construct an unbalanced panel over all 8 waves of individuals aged 20-64 from the SoFIE sample. We exclude 51 annual observations (across 45 individuals) with missing earnings data, and 48 observations (12 individuals) whose annual reference period changed. We retain only those individuals who could be matched to the IDI spine (95% of individuals and 97% of annual observations in our age range). This SoFIE panel is then matched to the LEED data, and each individual's annual LEED earnings are constructed by summing their monthly EMS wage and salary earnings over their SoFIE annual reference periods.

We sequentially subsample the full unbalanced panel of 20-64 year-olds using three criteria. First, we exclude annual observations of individuals reporting any self-employment activity in that wave, regardless of their other employment. Second, we select the balanced panel of individuals aged 20-64 throughout the 8 waves, who never report self-employment activity. Third, we select the balanced panel of individuals with both SoFIE and LEED reported earnings in each year.

Table 1 summarizes the characteristics of the various data samples, as well as the unmatched SoFIE subsample. Unmatched individuals are disproportionately female, Asian or Pasific Islander and less likely European, are more likely to have only high school qualifications, are less likely to have vocational qualifications, and have lower employment rates and earnings levels. Within the matched sample, the successive subsample selection criteria result in the exclusion of 9% individuals and 16% observations due to self-employment; 68% of individuals and 42% of observations due to the balanced panel requirement; and 35% of individuals and observations associated with both SoFIE and LEED earnings in each year. The final sample, which is used for our analysis of earnings dynamics, accounts for 21% of individuals and 37% of annual observations in the sample excluding self-employed.

The sample characteristics across the matched samples follow broadly similar patterns. Women are less likely to be employed, and are more likely to be observed in each wave of the SoFIE panel. Europeans are more likely to be self-employed, appear in the balanced panel, and work in each wave than other ethnic groups. Also, SoFIE employment, the number of months with earnings in LEED, and measured earnings from both sources, tend to be higher in the more selective samples. Finally, the average differences between SoFIE and LEED reported earnings are greater in the balanced than the unbalanced panel samples but less variable, reflected in lower standard deviations and average absolute differences.

3.1 Sources of error in reported earnings

The two sources of earnings data we use are expected to have alternative and various sources of measurement error. Errors in the derived annual earnings recorded in SoFIE may be due to three factors. First, from errors in respondents' reporting the existence or duration of job spells; second, from errors in their reporting of the wage or salary rate associated with a job spell; and third, from the derivation of annual earnings based on the spell-level information collected. The first and second reflect errors in respondents' recall or accuracy, while the third reflects any errors in the derivation algorithm used to estimate annual earnings from job-spell information that the survey instrument collected.

The sources of errors in the annual earnings recorded in LEED most likely are quite different to those recorded in SoFIE, and we assume the errors are independent. First, there may be differences in coverage, such as self-employment earnings, or informal sector (under-the-table) earnings, that are not included in the EMS forms. Second, there may be errors in the matching of data from different sources in the IDI, which will result in persistent errors in earnings. Third, there may be recording errors in the administrative data, associated with either the amount or the timing of earnings recorded through the EMS returns.

4 SoFIE versus LEED earnings comparisons

4.1 Cross-sectional earnings comparison

We begin by summarizing some descriptive patterns of the earnings reported in each of SoFIE and LEED. In this paper we focus on the intensive margin of earnings, conditional on non-zero earnings being reported (Hyslop and Townsend, 2017, analyze the misclassification errors in these employment reports). Figure 1 presents the distributions of SoFIE and LEED reported annual earnings, as well as the log-difference between these measures, for sample 3. (The distributions for the more inclusive samples are broadly similar, although have more mass near to zero earnings and also greater variability in the differences.) For presentation, we have top-coded earnings at \$200,000, and censored the log-differences at +/-1, in these figures. SoFIE earnings are more prevalent at salient levels (e.g. \$10K values) than LEED earnings, perhaps reflecting a tendency for individuals to report rounded earnings values. In addition, the distribution of log-differences is comparatively bell-shaped about zero, although with thicker tails than associated with a normal distribution.

To examine other correlates of the difference in the two reported earnings, we regress the log-difference on individuals' demographic and other characteristics. The results, reported in Table 2, show that although many of the coefficients are statistically significant, they are generally modest in size, and the socio-demographic variables explain almost none of the variation in the differences. The vast majority of the R^2 in these regression is due to the final three variables listed: weekly hours of work reported in SoFIE is positively correlated with the log(earnings) differences, while the number of jobs and number of months observed in LEED is negatively correlated.

4.2 Classical measurement error

The simplest and most restrictive assumption on the errors between the SoFIE and LEED reported earnings, is that the LEED earnings accurately measure true annual earnings (Y^*), and the SoFIE earnings are reported with purely random errors. That is, denoting Y_{Lit} as log(LEED earnings) of individual-i in year-t and Y_{Sit} as log(SoFIE earnings), the classical measurement error model is

$$\begin{aligned} &Y_{Lit} = Y^*_{it}, \\ &Y_{Sit} = Y^*_{it} + \epsilon_{it}, \epsilon_{it} \sim iid(0, \sigma^2_\epsilon), \end{aligned} \tag{1}$$

and the difference between the SoFIE and LEED reports is simply the measurement error in SoFIE earnings: $\epsilon_{it} = Y_{Sit} - Y_{Lit}$.

The assumption of classical errors has several strong implications. First, it implies that differences between SoFIE and LEED reports should be serially uncorrelated, and uncorrelated with the LEED reports and any observed covariates, and we explore each of these implications later. Second, it implies that in a regression of Y_S on Y_L, the coefficient on Y_L will be close to 1, while in the reverse regression, the coefficient on Y_S will be attenuated towards zero. To see this, consider the two regressions:

$$
\begin{aligned}
&Y_{Sit} = \alpha_0 + \alpha_1.Y_{Lit} + u_{Sit}, \\
&Y_{Lit} = \delta_0 + \delta_1.Y_{Sit} + u_{Lit}.
\end{aligned} \tag{2}
$$

In the first regression, $plim(\hat{\alpha_1}) = Cov(Y_{Lit}, Y_{Sit})/Var(Y_{Lit}) = \sigma^2_{Y^*}/\sigma^2_{Y^*} = 1$, while in the second regression $plim(\hat{\delta_1}) = Cov(Y_{Lit}, Y_{Sit})/Var(Y_{Sit}) = \sigma^2_{Y^*}/(\sigma^2_{Y^*} + \sigma^2_\epsilon) < 1$, where $\sigma^2_{Y^*}$ and σ^2_ϵ are the variances of true log(earnings) and measurement error respectively. In this context, the signal-to-total variability $(\sigma^2_{Y^*}/(\sigma^2_{Y^*} + \sigma^2_\epsilon))$ is also the reliability ratio of Y_S.

To evaluate this hypothesis, we present the estimates for each of these regressions in Table 3. The coefficients in the regressions of Y_{Sit} on Y_{Lit} range from 0.87 to 0.91 across the three samples, and are each statistically significantly lower then 1. In a classical error context, this rejects the assumption of no measurement error in the LEED reported earnings. The coefficients in the reverse regression of Y_{Lit} on Y_{Sit} range from 0.82 to 0.85, and also reject the hypothesis of no measurement error in log(SoFIE earnings). If we relax the assumption of no error in LEED reports to allow classical measurement error (and maintain classical errors in SoFIE), these estimates imply the statistical reliability of LEED earnings is 87–91%, and is higher than that of SoFIE earnings (82–85%).

4.3 Longitudinal earnings comparison

We next examine the relationship between the errors between the two reports and LEED earnings. In the third panel of Table 3 we report estimates from regressions of the difference in log(earnings) ($DY_{it} = Y_{Sit} - Y_{Lit}$) on individuals' average LEED log(earnings) over the period (Y_{Li}) and the year-specific deviation from this average ($Y_{Lit} - Y_{Li}$):

$$DY_{it} = \alpha + \beta.Y_{Li} + \delta.(Y_{Lit} - Y_{Li}) + u_{it}. \tag{3}$$

The differences in individuals reported log(earnings) are negatively correlated both with their average LEED earnings and their annual deviations, suggesting measurement errors are mean reverting with respect to both permanent earnings differences across individuals and also their transitory earnings. The coefficient on annual deviations is roughly constant across the three samples, meaning individuals under report about 20% of transitory earnings; while the coefficient on average log(earnings) is higher in the unbalanced panels (-0.1) than in the balanced panel of earners (-0.05). These results are consistent with the notion that persistent differences are less under-reported than transitory differences, and also less by those with persistent employment.

Table 4 analyzes the longitudinal properties of differences between SoFIE and LEED earnings further, documenting the auto-covariance matrix of the differences in SoFIE and LEED reported log(earnings) over the period. In order to abstract from possible selection effects associated with unbalanced panel samples, for this analysis we focus on the balanced panel of earners (the results based on the unbalanced samples are substantively similar, although with greater variability). To allow for possible variation over time in the average difference in reported earnings, we calculate the variances and covariances relative to year-specific mean differences, and present these means at the bottom of each panel of the table.

The variances of the differences between SoFIE and LEED log(earnings) range from 0.11 to 0.20, and are much higher in the first and last waves. The latter effect is a result of the balanced panel selection criteria: the end years have a larger fraction of part-year earners who are either entering or leaving employment, with greater associated variability in the difference in reported earnings. That is, individuals entering employment in wave-1 or leaving employment in wave-8 would be excluded if the panel was extended back or forward one year respectively. Similar end-year differences are apparent in the variance of earnings changes that we discuss in the next section.

A further implication of the classical measurement error model outlined in section 4.2 is that the autocovariance structure of the T-vector of errors ($\hat{\epsilon}_i$) will be diagonal with σ^2_ϵ on the diagonal. The auto-correlation patterns in Table 4 are clearly inconsistent with this prediction, and suggest that the errors between the SoFIE and LEED reported earnings include both persistent and transitory components. The first-order correlations are on the order of 0.25–0.35, while the higher order correlations are close to 0.1. In fact, the declining auto-correlation patterns across the two samples suggest the errors include transitory components that persist for 1-2 lags, as well as permanent components that persist longer, accounting for about 10% of the variance of the difference in reported earnings. We will draw on these patterns, as well as results in the literature, to help inform the nature of measurement errors that we allow in the next subsection.

5 Modeling earning dynamics

We now consider modelling the dynamics of individuals' earnings in the context of mis-measured reported earnings. In order to abstract from the additional difficulty of modeling individuals' employment decisions in the presence of measurement error, we focus on the balanced panel of individuals with both SoFIE and LEED earnings reported in each year.

5.1 A model of earnings dynamics and measurement error

In this section, we discuss the structure of measurement error that we will assess. We do this within the context of a dynamic model of earnings, by characterizing the measurement error in SoFIE to be consistent with results in the literature, and relax the assumption that the administrative LEED reported earnings are measured without error.

Following the literature (e.g. Abowd and Card, 1989, Meghir and Pistaferri, 2004), we assume individual-i's true log(earnings) in year-t consist of a permanent random walk component plus a transitory MA(1) component:

$$
\begin{aligned}
&Y_{it}^* = \alpha_{it} + u_{it}, \\
&\alpha_{it} = \alpha_{it-1} + \eta_{it}, \qquad\qquad (4) \\
&u_{it} = \omega_{it} + \theta\omega_{it-1},
\end{aligned}
$$

where $\eta_{it} \sim iid(0, \sigma_\eta^2)$, and $\omega_{it} \sim iid(0, \sigma_\omega^2)$.

The literature on validation studies of survey reported earnings concludes that the classical measurement error model is too restrictive. First, both Bound and Krueger (1991) and Pischke (1995) found evidence of serial correlation in the measurement error in PSID reported earnings, which is also consistent with the pattern of autocorrelations in Table 4, and implies $Corr(\epsilon_{it}, \epsilon_{is}) \neq 0, s \neq t$. Second, they also found evidence of mean-reversion in the measurement error, which implies that $Corr(\epsilon_{it}, Y_{it}^*) < 0$, and is consistent with the results in Table 3. The interpretation of such non-classical measurement errors is usually that respondents under-report transitory earnings shocks due to memory lapses.

Broadly consistent with Bound and Krueger (1991), Pischke (1995), and the patterns in Table 4, we assume that the measurement error in individual-i's SoFIE reported earnings consists of a person-specific component, components related to each of their permanent earnings shock and their transitory earnings, and a classical error component:

$$\epsilon_{Sit} = \lambda_{Si} + \delta_{1S}\eta_{it} + \delta_{2S}u_{it} + \nu_{Sit}, \tag{5}$$

where $\lambda_{Si} \sim iid(0, \sigma_{S\lambda}^2)$ and $\nu_{Sit} \sim iid(0, \sigma_{S\nu}^2)$. Mean-reverting errors imply δ_{1S} and δ_{2S} are negative. Combining equations (4) and (5) implies the individual reports SoFIE earnings:

$$Y_{Sit} = Y_{it}^* + \epsilon_{Sit} = \alpha_{it} + \lambda_{Si} + \delta_{1S}\eta_{it} + (1 + \delta_{2S})u_{it} + \nu_{Sit}. \tag{6}$$

As discussed in section 3.1, errors in administrative earnings reports consist of a combination of coverage, mis-matching of individuals, or possibly their jobs, and random mis-coding of their earnings. Coverage issues may result in both persistent and transitory errors, while matching errors will likely be persistent over time, and coding errors should be transitory. (To the extent that mis-matches are job-specific, this may generate a person-job component of error: if so, the measurement error may be serially correlated, but is unlikely to be mean reverting.) We adopt a simple specification for the combined measurement error in individual-i's LEED reported earnings, and assume that it consists of a person-specific component and a classical error component:

$$\epsilon_{Lit} = \lambda_{Li} + \nu_{Lit}, \tag{7}$$

where $\lambda_{Li} \sim iid(0, \sigma^2_{L\lambda})$ and $\nu_{Lit} \sim iid(0, \sigma^2_{L\nu})$. Combining equation(7) with (4), gives i's reported LEED earnings:

$$Y_{Lit} = Y^*_{it} + \epsilon_{Lit} = \alpha_{it} + \lambda_{Li} + u_{it} + \nu_{Lit}. \tag{8}$$

In order to abstract from the initial conditions associated with individual i's permanent earnings component (α_{i0}), we estimate the model using the first differences of SoFIE and LEED earnings. This also eliminates the person-specific components of error (λ_{Si} and λ_{Li}), which are not identified in our estimation. In particular, equations (6) and (8) imply:

$$\begin{aligned} \Delta Y_{Sit} &= (1+\delta_{1S})\eta_{it} - \delta_{1S}\eta_{it-1} + (1+\delta_{2S})\Delta u_{it} + \Delta\nu_{Sit}, \\ \Delta Y_{Lit} &= \eta_{it} + \Delta u_{it} + \Delta\nu_{Lit}. \end{aligned} \tag{9}$$

The model parameters of interest are $(\sigma^2_\eta, \sigma^2_\omega, \theta, \delta_{1S}, \delta_{2S}, \sigma^2_{S\nu}, \sigma^2_{L\nu})$. These are identified from the auto-covariances and cross-covariances of ΔY_{Sit} and ΔY_{Lit}: the model variances and covariances are presented in the appendix. One implication of the model is that all auto- and cross-covariances of ΔY_{Sit} and ΔY_{Lit} beyond second-order are zero.

We estimate the model using minimum distance estimation methods (Abowd and Card, 1989; Chamberlain, 1984). This involves choosing the vector of parameter estimates to minimize the weighted sum of squared differences between the empirical and model-predicted second moments. Because of finite sample bias associated with the second and fourth moments being correlated (Altonji and Segal, 1996), we weight using the diagnonal matrix with inverse sampling variances of the empirical moments on the diagonal instead of the optimal weight matrix which also includes the off-diagonal sampling covariances.

5.2 Estimation results

Table 5 summarizes the empirical covariance structures of $(\Delta Y_{Sit}, \Delta Y_{Lit})$, separately for males and females. First, the variances of SoFIE earnings changes are substantially larger than for LEED changes, which suggests a greater degree of random measurement error in SoFIE earnings. As discussed in section 4.3, the variances of each measure's change in earnings is much greater in the end years than the intermediate years of the panel. As well as greater variance of earnings change in these years, the mean earnings growth is generally stronger between waves 1 and 2 (0.09 for SoFIE, and 0.12 for LEED male earnings, and 0.13 for female earnings), and much weaker between waves 7 and 8 (about -0.01 or -0.02 for males, and 0.0 for females), consistent with end-year entry and exit patterns respectively. Female earnings changes are substantially more variable than male changes, with female variances nearly double those of males for each measure.

Second, consistent with the patterns of autocorrelations in the log differences between SoFIE and LEED reports and much of the literature on earnings dynamics, and in line with the model predictions in section (5.1), the first- and second-lagged auto-covariances of SoFIE and LEED earnings changes are generally statistically significantly different from zero, while all the higher order covariances are small and individually statistically insignficantly different from zero. The first-order auto-correlations are typically between -0.2 and -0.3, and the second order auto-correlations are also negative and generally smaller than 0.1 in magnitude.

Third, the contemporaneous cross-covariances between the changes in SoFIE and LEED earnings are statistically significant, with implied correlations of 0.40 on average for males and 0.57 for females. In contrast to the auto-covariances, most of the first-order, and all of the higher-order, cross-covariances are statistically insignificant. Interestingly, the first-order correlations associated with LEED leading SoFIE, $Cov(\Delta Y_{Sit}, \Delta Y_{Lit-1})$, tend to be greater than for SoFIE leading LEED, $Cov(\Delta Y_{Sit-1}, \Delta Y_{Lit})$: typically on the order of (-0.1,-0.05) for the former, and (-0.05,0) for the latter.

We next present estimates of alternative specifications of the model in equations (4)–(9). The results are contained in Table 6, for models estimated separately for males and females. To provide a baseline comparison with literature that assumes that the validated administrative earnings are reported without error, we first estimate models that has this condition. The estimates, presented in the first column for males and females, are similar to findings in the literature. In particular, there is evidence of positively correlated transitory components of earnings, with the measurement error in SoFIE earnings being insignificantly correlated with permanent shocks and more strongly mean-reverting with respect to the transitory component of earnings. Consistent with Bound and Krueger (1991), we estimate stronger mean reversion in male than in female earnings, but more transitory and classical measurement error in female earnings.

In column (2) we present estimates of the model allowing for measurement error in the LEED earnings. First, there is evidence of significant measurement error in LEED earnings in this model: the estimated variance of the measurement error is almost the same for males and females (0.023 and 0.025 respectively). The results imply substantially less measurement error in LEED than in SoFIE earnings: the estimated variance of LEED errors is about one-half the variance of classical measurement error in SoFIE earnings for males, and about one-third for females.

Second, allowing for measurement error in LEED earnings, the result that SoFIE measurement error is negatively correlated with transitory earnings is overturned, with each of the estimated δ_{2S} parameters being positive. The δ_{2S} parameters are imprecisely estimated, particularly for males, in this model. Inspection of the moments in the appendix implies δ_{2S} is identified by the ratio of alternative second-order moments with $Cov(\Delta Y_{Sit}, \Delta Y_{Lit-2})$ in the denominator, and by the ratio of first-order moments whose denominator is $Cov(\Delta Y_{Lit}, \Delta Y_{Lit-1}) + \sigma^2_{L\nu}$. The empirical second-order moments are generally close to zero, which suggests they provide relatively weak power to identify δ_{2S}, although they are somewhat larger for females than males. This suggests the main identification of δ_{2S} comes form the first-order moments which, as seen in column (1) is empirically strong for males if the LEED reports are assumed to be error free ($\sigma^2_{L\nu} = 0$), but weak given the estimated $\sigma^2_{L\nu}$ value. The result that mean reversion in survey earnings is not robust to allowing measurement error in administrative data is consistent with Kapteyn and Ypma (2007), based on a different approach. However, we do estimate small and insignificant negative correlations between SoFIE measurement error with individuals' permanent earnings shocks (δ_{1S}) for each sample.

The final row contains a formal goodness-of-fit (GOF) statistic for each model in Table 6. Although the second model fits salient aspects of the empirical covariance matrix, neither of the models provides an adequate statistical fit to the structure as judged by the GOF statistics. Both models fit relatively better for males than females, while excluding the zero-predicted moments has a larger effect for females. One possible issue is that both models restrict the earnings and error process to be stationary over the period, while there is evidence of time varying variances and covariances. As discussed above, the first and last year variances are noticeably higher, perhaps due to these end years including individuals with more variable earnings associated with moving in or out of employment, who would have been excluded if the sample period was extended in either direction.

To account for this source of non-stationarity, we next estimate a model that allows for separate end-year variances in the classical measurement error components of SoFIE earnings ($\sigma^2_{S\nu 0}$) and LEED earnings ($\sigma^2_{L\nu 0}$). The results of this model are in presented in column (3). The estimated end-year variances for the measurement errors in LEED are nearly three times the variances of the other years, again with similar magnitudes for males and females. For SoFIE errors, the end-year variance for males is also much larger (about 2.5 times) than the variance for other years, but for females these variances are of similar magnitude. The decrease in the GOF statistics for this specification imply a substantial improvement in the fit of the model, especially for females. However, it has almost no effect on the other parameter estimates in the model, especially those concerned with true earnings.

The final model, presented in column (4), extends this idea and allows wave-specific classical measurement error variances in each of SoFIE and LEED earnings, which will allow the variances and first-order covariances to vary over time. Although the model's statistical adequacy is still rejected by the formal GOF statistics, the large falls in the GOF statistics shown in the table indicate it provides a significantly better fit to the empirical moments. Perhaps more substantively, the other parameter estimates are almost unchanged as a result of this relaxation of the model.

We have also estimated baseline earnings dynamics models using SoFIE or LEED data separately and ignoring measurement errors, and also a joint model that allows for classical measurement errors in each measure of earnings (reported in Hyslop and Townsend, 2016). Using the separate earnings reports results in quite different estimates of the permanent shocks: the estimated variance is much larger based on SoFIE than on LEED reported earnings, and these estimates sandwich those in Table 6. The estimated variances of the transitory earnings shocks are also much larger, and the MA persistence parameter (θ) smaller, than in Table 6. This is because it is not possible to separately identify the transitory earnings effects from measurement error using a single source of earnings. As a result, random measurement errors will inflate the estimated transitory shocks, and reduce the estimated persistence. Finally, when both sources of earnings are used and allowed to have only classical measurement errors, the estimates are quite similar to those for the models that also allow SoFIE errors to have non-classical components. This isn't surprising as the estimated δ parameters in those models are statistically insignficant. However, the GOF statistics suggest this model provides a noticeably worse fit to the moments, especially for females.

Although the estimated models in Table 6 are rejected on the basis of the formal GOF criteria, there are no obvious patterns of misfit apparent in the predicted moments. Also, the simpler model (2) specification provides broadly comparable predictions to model (4), other than the latter providing a noticeably better fit to the variances. Given this, and that the core model parameters are largely unaffected by relaxing the stationarity of the classical error variances, we will use this simpler model in the next section for discussing the implications of this analysis for understanding individuals' earnings inequality.

5.3 Implications for earnings dynamics and inequality

We now discuss the implications of the earnings dynamics model for the extent and persistence of inequality in true earnings. Because the permanent components of measurement error are not separately identified in the model, the model does not directly inform the level of earnings inequality. However, we can address the source and impact of earnings shocks as measured by the variance of changes in log(earnings). A summary of the results are presented in Table 7, based on the estimates of model (2) in Table 6.

First, measurement errors in both SoFIE and LEED reported earnings account for substantial proportions of the variances of reported earnings changes. In particular, the predicted variance of true log(earnings) changes is 0.038 for males, which accounts for only 26% and 36% of the predicted variances of change in SoFIE (0.145) and LEED (0.105) reported earnings respectively. For females, the predicted variance of true log(earnings) changes is 0.129, which accounts for 42% and 57% of the predicted variances of SoFIE (0.305) and LEED (0.226) reported earnings changes respectively.

Second, female log(earnings) changes are substantially more variable than males. For example, the variances of observed log(earnings) changes are roughly twice for females compared to males: 0.305 versus 0.145 for SoFIE reports, and 0.226 versus 0.105 for LEED. In fact, the female variance of true earnings changes is more than three times larger than for males: 0.129 versus 0.038. This difference is due both to larger permanent shocks (0.085 versus 0.031), transitory shocks (0.029 versus 0.004) which are also (slightly) more persistent for females (θ=0.55 versus 0.51 for males).

Third, although the permanent components of measurement errors in the model are not identified, we are able to bound their effects. Equations (6) and (8) imply:

$$DY_{it} = Y_{Sit} - Y_{Lit} = \delta_{1S}\eta_{it} + \delta_{2S}u_{it} + (\lambda_{Si} - \lambda_{Li}) + (\nu_{Sit} - \nu_{Lit}). \tag{10}$$

This implies the auto-covariance in DY_{it} equals $Var(\lambda_{Si} - \lambda_{Li})$ beyond first-order. As discussed above, the auto-correlations in Table 4 suggest that this variance accounts for about 10% of the $Var(DY_{it})$. (Although $Var(DY_{it})$ differs for males (0.12 on average) and females (0.17 average), the correlations are quite similar.) Assuming the permanent components of reporting errors (λ_{Si} and λ_{Li}) are independent (or at least not negatively correlated), then the combined and separate effects of $\sigma^2_{\lambda Si}$ and $\sigma^2_{\lambda Li}$ are less than 10% of $Var(DY_{it})$. This provides upper bounds for $\sigma^2_{\lambda Si}$ and $\sigma^2_{\lambda Li}$ of about 0.012 for males and 0.017 for females.

Together with the empirical variances of SoFIE and LEED log(earnings), these bounds can help inform the discussion of the extent of measurement error in the level of earnings, and their effects on inequality. For example, using the empirical variances of observed SoFIE and LEED log(earnings) of 0.164 and 0.135 respectively for males, the estimates above imply the maximal measurement error contributions are 0.072 and 0.037, giving lower bound estimates of the variances of true log(earnings) of 0.09–0.10 across the SoFIE and LEED reports. Similarly, the variances of SoFIE and LEED log(earnings) for females are 0.424 and 0.352, with estimated maximal measurement error contributions of 0.121 and 0.046, implying estimated variances of true log(earnings) of 0.30.

Finally, using these estimates of the variance of true earnings and the estimates of the components of earnings, suggests that transitory variation accounts for a relatively trivial fraction of the variance of male earnings (about 6%), and about 13% of female earnings.

6 Concluding discussion

We have used a rich longitudinal sample of survey earnings matched to administrative earnings to analyze the measurement error and dynamic properties of individuals' earnings. The analysis provides several conclusions. First, we reject that survey earnings are reported with classical measurement error and the administrative earnings are recorded without error. Under the assumption that LEED earnings are correct, we estimate similar mean-reverting patterns in SoFIE earnings to that found in the validation literature. In particular, differences between the reported earnings are negatively correlated with both persistent differences and transitory changes in individuals' earnings. Allowing classical measurement error in each implies LEED earnings are more reliably reported than in SoFIE, with reliability ratios about 90% and 85% respectively.

Second, the longitudinal properties of the difference between SoFIE and LEED earnings for individuals further confirms the non-classical nature of the errors. In particular, the covariance structure of the differences is characterized by both persistent and serially correlated transitory components, consistent with a simple model including a person specific permanent component plus a low-order ARMA component of error.

Third, we have formulated a model for the SoFIE and LEED reported earnings based on the empirical characteristics, together with findings from the literature on individuals' true earnings dynamics. We find that the measurement error in SoFIE earnings is mean-reverting within this formulation when the LEED earnings are assumed to be true. However, as with Kapteyn and Ypma (2007), this result is not robust to allowing errors in LEED earnings. In fact, we conclude that each source of earnings is largely characterized by classical measurement error and possibly a person-specific permanent component of error. Due to both greater permanent and transitory shocks, true Female earnings, as well as earnings changes, are about three times more variable than for males. Finally, measurement errors account for about 70% of the variance of change in male earnings, and about half of the variance of female earnings.

References

Abowd, J. M. and D. Card (1989). On the covariance structure of earnings and hours changes. *Econometrica 57*(2), 411–445.

Abowd, J. M. and M. H. Stinson (2013). Estimating measurement error in annual job earnings: A comparison of survey and administrative data. *Review of Economics and Statistics 95*(5), 1451–1467.

Altonji, J. G. and L. M. Segal (1996). Small-sample bias in gmm estimation of covariance structures. *Journal of Business & Economic Statistics 14*(3), 353–366.

Bound, J., C. Brown, G. J. Duncan, and W. L. Rodgers (1994). Evidence on the validity of cross-sectional and longitudinal labor market data. *Journal of Labor Economics*, 345–368.

Bound, J., C. Brown, and N. Mathiowetz (2001). Measurement error in survey data. *Handbook of econometrics 5*, 3705–3843.

Bound, J. and A. B. Krueger (1991). The extent of measurement error in longitudinal earnings data: Do two wrongs make a right? *Journal of Labor Economics*, 1–24.

Chamberlain, G. (1984). Panel data. *Handbook of econometrics 2*, 1247–1318.

Friedman, M. (1957). The permanent income hypothesis. In *A theory of the consumption function*, pp. 20–37. Princeton University Press.

Gottschalk, P. and M. Huynh (2010). Are earnings inequality and mobility overstated? the impact of nonclassical measurement error. *The Review of Economics and Statistics 92*(2), 302–315.

Hause, J. C. (1980). The fine structure of earnings and the on-the-job training hypothesis. *Econometrica: Journal of the Econometric Society*, 1013–1029.

Hyslop, D. R. (2001). Rising us earnings inequality and family labor supply: The covariance structure of intrafamily earnings. *American Economic Review*, 755–777.

Hyslop, D. R. and W. Townsend (2016). Earnings dynamics and measurement error in matched survey and administrative data. *Motu Working Paper 16-18*.

Hyslop, D. R. and W. Townsend (2017). Employment misclassification in survey and administrative reports. *Economics Letters 155*, 19–23.

Kapteyn, A. and J. Y. Ypma (2007). Measurement error and misclassification: A comparison of survey and administrative data. *Journal of Labor Economics 25*(3), 513–551.

Lillard, L. A. and Y. Weiss (1979). Components of variation in panel earnings data: American scientists 1960-70. *Econometrica: Journal of the Econometric Society*, 437–454.

MaCurdy, T. E. (1982). The use of time series processes to model the error structure of earnings in a longitudinal data analysis. *Journal of econometrics 18*(1), 83–114.

Meghir, C. and L. Pistaferri (2004). Income variance dynamics and heterogeneity. *Econometrica 72*(1), 1–32.

Meghir, C. and L. Pistaferri (2011). Earnings, consumption and life cycle choices. *Handbook of Labor Economics 4*, 773–854.

Pischke, J.-S. (1995). Measurement error and earnings dynamics: Some estimates from the psid validation study. *Journal of Business & Economic Statistics 13*(3), 305–314.

Appendix A Model predicted covariances

In this appendix we present the variances and covariances of the changes in SoFIE and LEED log(earnings) in equations (9), from the earnings model described by equations (4), (6) and (8) in section 5.1. In particular, the variances and covariances are:

$$Var(\Delta Y_{Sit}) = (\delta_{1S}^2 + (1+\delta_{1S})^2)\sigma_\eta^2 + (1+\delta_{2S})^2\sigma_{\Delta u}^2 + 2\sigma_{S\nu}^2,$$

$$Cov(\Delta Y_{Sit}, \Delta Y_{Sit-1}) = (1+\delta_{2S})^2 Cov(\Delta u_{it}, \Delta u_{it-1}) - (1+\delta_{1S})\delta_{1S}\sigma_\eta^2 - \sigma_{S\nu}^2,$$

$$Cov(\Delta Y_{Sit}, \Delta Y_{Sit-2}) = (1+\delta_{2S})^2 Cov(\Delta u_{it}, \Delta u_{it-2}),$$

$$Cov(\Delta Y_{Sit}, \Delta Y_{Sit-k}) = 0, k > 2,$$

$$Var(\Delta Y_{Lit}) = \sigma_\eta^2 + \sigma_{\Delta u}^2 + 2\sigma_{L\nu}^2,$$

$$Cov(\Delta Y_{Lit}, \Delta Y_{Lit-1}) = Cov(\Delta u_{it}, \Delta u_{it-1}) - \sigma_{L\nu}^2,$$

$$Cov(\Delta Y_{Lit}, \Delta Y_{Lit-2}) = Cov(\Delta u_{it}, \Delta u_{it-2}),$$

$$Cov(\Delta Y_{Lit}, \Delta Y_{Lit-k}) = 0, k > 2,$$

$$Cov(\Delta Y_{Sit}, \Delta Y_{Lit}) = (1+\delta_{1S})\sigma_\eta^2 + (1+\delta_{2S})\sigma_{\Delta u}^2,$$

$$Cov(\Delta Y_{Sit}, \Delta Y_{Lit-1}) = (1+\delta_{2S})Cov(\Delta u_{it}, \Delta u_{it-1}) - \delta_{1S}\sigma_\eta^2,$$

$$Cov(\Delta Y_{Sit}, \Delta Y_{Lit-2}) = (1+\delta_{2S})Cov(\Delta u_{it}, \Delta u_{it-2}),$$

$$Cov(\Delta Y_{Sit}, \Delta Y_{Lit-k}) = 0, k > 2,$$

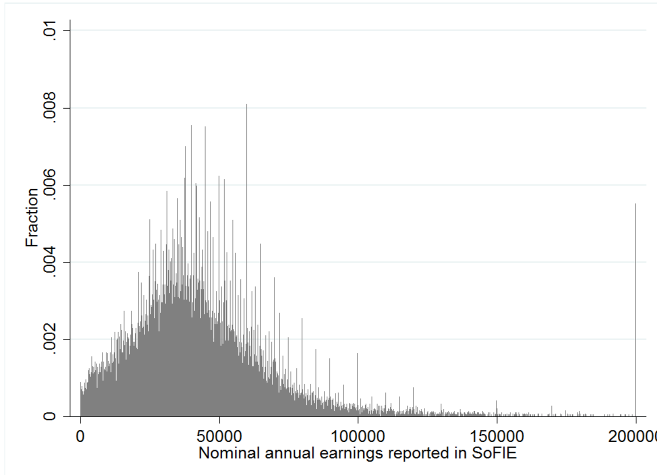
$$Cov(\Delta Y_{Sit-1}, \Delta Y_{Lit}) = (1+\delta_{2S})Cov(\Delta u_{it}, \Delta u_{it-1}),$$

$$Cov(\Delta Y_{Sit-k}, \Delta Y_{Lit}) = Cov(\Delta Y_{Sit}, \Delta Y_{Lit-k}), k > 1,$$

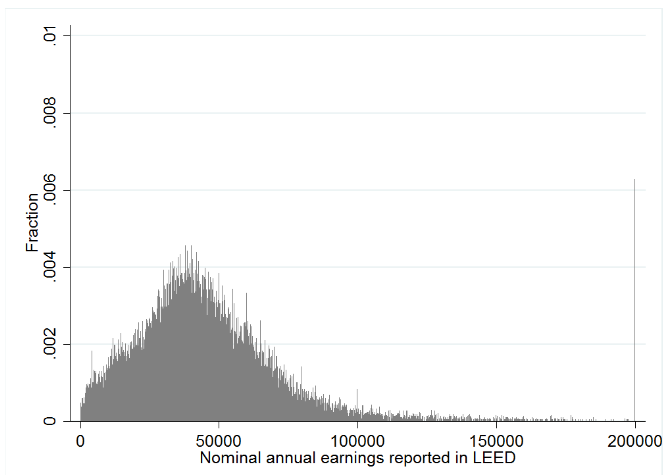
where $\sigma_{\Delta u}^2 = Var(\Delta u_{it}) = (1 + (1-\theta)^2 + \theta^2)\sigma_\omega^2$, $Cov(\Delta u_{it}, \Delta u_{it-1}) = ((\theta-1) + \theta(1-\theta))\sigma_\omega^2$, and $Cov(\Delta u_{it}, \Delta u_{it-2}) = -\theta\sigma_\omega^2$.

The vector of parameters of interest in this model is $(\sigma_\eta^2, \sigma_\omega^2, \theta, \delta_{1S}, \delta_{2S}, \sigma_{S\nu}^2, \sigma_{L\nu}^2)$ are identified by comparing the theoretical moments in the above equations with the empirical second moments (variances and covariances) of the changes in SoFIE and LEED log(earnings). Minimum distance estimation chooses the vector of parameter estimates to minimize the weighted sum of squared differences between the vectors of empirical and model-predicted second moments. Optimal minimum distance (OMD) estimation involves using as the weight matrix the inverse of the sampling variance-covariance matrix of the empirical moments. However, this involves the fourth moments of the data which, as Altonji and Segal (1996) show, results in substantial finite sample estimation bias due to correlation between the second and fourth moments. For this reason, instead of the OMD, we use as weight matrix the diagonal matrix which has the inverse of the sampling variances on the diagonal. This weighting approach, which takes account of the different variability across the second moments being fit by the model but not the correlations between the moments, has also been used by Hyslop (2001) and Pischke (1995).

Figure 1: Distributions of positive earnings – Sample (3)
(A) SoFIE earnings



(B) LEED earnings



(C) log(SoFIE earnings) - log(LEED earnings)

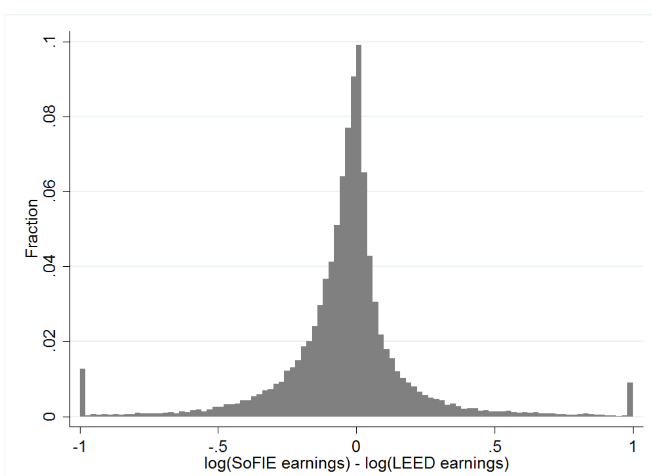


Table 1: Descriptive statistics – SoFIE and LEED matched samples

	Unmatched	Full	Analysis samples		
	Sample	Matched	(1)	(2)	(3)
Age	41.81	41.77	40.84	43.04	43.32
	(13.51)	(12.46)	(12.63)	(10.4)	(10.11)
Female	0.59	0.53	0.57	0.59	0.52
	(0.49)	(0.50)	(0.50)	(0.49)	(0.50)
European	0.57	0.76	0.74	0.77	0.81
	(0.50)	(0.43)	(0.44)	(0.42)	(0.39)
Maori	0.12	0.12	0.13	0.12	0.10
	(0.33)	(0.32)	(0.34)	(0.32)	(0.30)
Pacific Islander	0.12	0.05	0.05	0.05	0.04
	(0.32)	(0.21)	(0.23)	(0.22)	(0.20)
Asian	0.16	0.06	0.06	0.05	0.04
	(0.37)	(0.23)	(0.23)	(0.22)	(0.19)
High school	0.33	0.27	0.28	0.26	0.26
	(0.47)	(0.44)	(0.45)	(0.44)	(0.44)
Vocational	0.27	0.36	0.35	0.36	0.38
	(0.44)	(0.48)	(0.48)	(0.48)	(0.49)
Bachelor's degree	0.11	0.12	0.12	0.12	0.13
	(0.32)	(0.33)	(0.32)	(0.32)	(0.33)
Higher degree	0.06	0.07	0.06	0.07	0.08
	(0.23)	(0.25)	(0.24)	(0.25)	(0.27)
SoFIE: Employed	0.60	0.72	0.81	0.84	1
	(0.49)	(0.45)	(0.39)	(0.37)	
Hours worked/week	38.36	38.88	38.17	38.31	39.96
	(16.65)	(14.91)	(13.73)	(13.47)	(12.17)
Earnings ($)	20,579	27,798	31,525	35,010	45,974
	(31,268)	(34,578)	(35,106)	(37,502)	(32,752)
LEED: Employed	—	0.74	0.81	0.82	1
		(0.44)	(0.40)	(0.38)	
No. jobs	—	1.14	1.24	1.21	1.45
		(1.12)	(1.11)	(0.99)	(0.87)
No. months earnings	—	7.81	8.65	9.12	11.61
		(5.27)	(4.9)	(4.73)	(1.4)
Earnings ($)	—	28,081	30,992	34,882	47,644
		(32,812)	(32,806)	(34,486)	(33,679)
log(Sofie/LEED earn)		-0.02	-0.02	-0.04	-0.04
	—	(0.58)	(0.56)	(0.49)	(0.38)
abs[log(Sofie/LEED earn)]		0.26	0.25	0.21	0.17
	—	(0.52)	(0.5)	(0.45)	(0.35)
No. Individuals	1,383	24,138	22,017	7,104	4,572
No. Observations	3,417	116,643	97,962	56,832	36,576

Notes: Analysis sample (1) excludes annual observations with reported self-employment; (2) is the balanced panel of persons; and (3) is the balanced panel with SoFIE and LEED earnings in each year. Standard deviations are in parentheses. All earnings are in nominal $-values. Sample sizes throughout the paper are randomly rounded to base 3 to maintain confidentiality.

Table 2: Correlates of differences between SoFIE and LEED earnings

Dependent variable:	log(earn) difference			abs(log(earn) difference)		
Sample:	(1)	(2)	(3)	(1)	(2)	(3)
Age 25-54	0.02	0.01	0.01	-0.01	-0.02	-0.01
	(0.01)	(0.01)	(0.01)	(0.00)	(0.01)	(0.00)
Age<25	0.02	0.01	0.02	0.00	-0.03	-0.02
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Female	0.01	0.02	0.01	-0.01	-0.02	-0.01
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Maori	-0.02	-0.01	-0.01	0.03	0.02	0.02
	(0.01)	(0.01)	(0.01)	(0.00)	(0.01)	(0.01)
Pacific Islander	-0.08	-0.08	-0.07	0.05	0.04	0.04
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Asian	-0.06	-0.05	-0.05	0.02	0.03	0.02
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Other ethnicity	-0.07	-0.03	-0.03	0.02	-0.01	0.00
	(0.01)	(0.02)	(0.02)	(0.01)	(0.02)	(0.01)
High school	0.01	0.02	0.02	-0.03	-0.02	-0.02
	(0.01)	(0.01)	(0.01)	(0.00)	(0.01)	(0.01)
Vocational	0.02	0.02	0.02	-0.04	-0.03	-0.03
	(0.01)	(0.01)	(0.01)	(0.00)	(0.01)	(0.00)
Bachelor degree	0.02	0.03	0.02	-0.06	-0.05	-0.04
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Higher degree	0.02	0.03	0.03	-0.05	-0.04	-0.03
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Weekly hours (x10)	0.04	0.03	0.02	-0.03	-0.02	-0.02
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
No. LEED jobs	-0.02	-0.02	-0.02	0.03	0.03	0.03
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
No. LEED months	-0.09	-0.08	-0.07	-0.09	-0.09	-0.09
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
R^2	0.117	0.096	0.064	0.201	0.186	0.136
No. individuals	16,578	6,237	4,566	16,578	6,237	4,563
No. observations	67,905	41,949	34,968	67,905	41,949	34,968

Notes: Each column reports estimates from the regression of either the difference between log(SoFIE earnings) and log(LEED earnings), or the absolute value of this difference.

Table 3: Measurement error regressions

Covariate:	Sample (1)	(2)	(3)
(A) Dependent variable: Y_{Sit}= log(SoFIE earnings)			
Y_{Lit} (α_1)	0.873	0.893	0.911
	(.002)	(.002)	(.003)
Intercept (α_0)	1.280	1.068	0.899
	(.020)	(.025)	(.028)
(B) Dependent variable: Y_{Lit}= log(LEED earnings)			
Y_{Sit} (δ_1)	0.832	0.848	0.837
	(.002)	(.002)	(.002)
Intercept (δ_0)	1.742	1.606	1.753
	(.019)	(.023)	(.026)
R^2	0.726	0.758	0.763
(C) Dependent variable: $DY_{it} = Y_{Sit} - Y_{Lit}$			
Y_{Li} (β)	-0.100	-0.074	-0.047
	(.005)	(.007)	(.008)
$(Y_{Lit} - Y_{Li})$ (δ)	-0.230	-0.213	-0.206
	(.010)	(.013)	(.014)
Intercept (α)	1.008	0.736	0.455
	(.054)	(.073)	(.085)
R^2	0.065	0.060	0.048
No. individuals	17,814	6,351	4,572
No. observations	75,225	45,336	36,576

Notes: Panels (A) and (B) report results of equation (2) regressions; panel (C) reports results of equation (3) regression.

Table 4: The covariance structure of SoFIE and LEED log(earnings) differences

	log(SoFIE earn) - log(LEED earn) in Wave:							
Wave:	1	2	3	4	5	6	7	8
1	**0.193**	0.352	0.121	0.095	0.070	0.084	0.079	0.095
	(.025)							
2	0.064	**0.168**	0.247	0.108	0.110	0.098	0.109	0.105
	(.016)	(.024)						
3	0.019	0.037	**0.131**	0.287	0.194	0.146	0.120	0.101
	(.004)	(.005)	(.014)					
4	0.015	0.016	0.037	**0.124**	0.295	0.188	0.146	0.109
	(.003)	(.004)	(.005)	(.017)				
5	0.010	0.015	0.023	0.035	**0.110**	0.272	0.177	0.128
	(.003)	(.003)	(.004)	(.005)	(.011)			
6	0.013	0.014	0.019	0.023	0.032	**0.125**	0.274	0.158
	(.003)	(.004)	(.003)	(.004)	(.004)	(.014)		
7	0.012	0.016	0.015	0.018	0.021	0.034	**0.122**	0.240
	(.003)	(.005)	(.003)	(.003)	(.003)	(.006)	(.020)	
8	0.019	0.019	0.016	0.017	0.019	0.025	0.037	**0.198**
	(.005)	(.006)	(.004)	(.004)	(.004)	(.005)	(.006)	(.030)
Mean	-0.026	-0.040	-0.050	-0.050	-0.040	-0.045	-0.046	-0.050
	(.007)	(.006)	(.005)	(.005)	(.005)	(.005)	(.005)	(.007)

Notes: Estimated using the sample (3) balanced panel of earnings. Variances are in **bold** on the diagonal, covariances are below the diagonal, and correlations above. Standard errors are in parentheses below the variances and covariances. Means and standard errors are in the final two rows.

Table 5: Summary of SoFIE and LEED log(earnings) changes

lag	Average variance or correlation			
k	$(\Delta Y_{Sit},\Delta Y_{Sit-k})$	$(\Delta Y_{Lit},\Delta Y_{Lit-k})$	$(\Delta Y_{Sit-k},\Delta Y_{Lit})$	$(\Delta Y_{Sit},\Delta Y_{Lit-k})$
		(A) Males		
Variances:	0.172	0.107	—	—
Correlations:				
0	1	1	0.400	0.400
1	-0.304	-0.257	-0.030	-0.053
2	-0.051	-0.042	-0.019	-0.029
3	-0.031	0.008	-0.011	0.003
4	0.038	-0.018	0.011	0.023
5	-0.023	0.014	0.000	0.010
6	-0.025	0.005	0.000	0.001
		(B) Females		
Variances:	0.321	0.205	—	—
Correlations:				
0	1	1	0.572	0.572
1	-0.270	-0.206	-0.035	-0.100
2	-0.090	-0.100	-0.091	-0.051
3	-0.014	-0.014	-0.018	-0.016
4	-0.007	-0.009	-0.009	0.018
5	0.007	-0.006	0.002	-0.013
6	-0.033	-0.020	-0.024	-0.030

Notes: Each panel summarizes the average variances, auto- and cross-correlations of annual earnings changes over the sample period.

Table 6: Estimated models of earnings dynamics and measurement errors

	Males				Females			
Parameter	(1)	(2)	(3)	(4)	(1)	(2)	(3)	(4)
σ^2_η	0.034	0.031	0.031	0.031	0.083	0.085	0.085	0.083
	(.003)	(.003)	(.003)	(.003)	(.007)	(.007)	(.007)	(.007)
σ^2_ω	0.032	0.004	0.005	0.004	0.066	0.029	0.030	0.030
	(.002)	(.002)	(.002)	(.002)	(.005)	(.007)	(.007)	(.007)
θ	0.149	0.513	0.502	0.511	0.350	0.554	0.552	0.545
	(.028)	(.230)	(.217)	(.224)	(.025)	(.111)	(.110)	(.106)
δ_{1S}	-0.010	0.082	0.083	0.082	-0.068	0.115	0.115	0.117
	(.094)	(.124)	(.126)	(.124)	(.058)	(.093)	(.093)	(.095)
δ_{2S}	-0.797	0.643	0.617	0.634	-0.274	0.171	0.170	0.172
	(.080)	(.409)	(.401)	(.407)	(.082)	(.162)	(.161)	(.163)
$\sigma^2_{S\nu}$	0.053	0.045	0.040	—	0.085	0.069	0.069	—
	(.006)	(.007)	(.007)		(.009)	(.009)	(.010)	
$\sigma^2_{L\nu}$	—	0.025	0.023	—	—	0.029	0.027	—
		(.002)	(.003)			(.006)	(.007)	
$\sigma^2_{S\nu 0}$	—	—	0.101	—	—	—	0.065	—
			(.021)				(.017)	
$\sigma^2_{L\nu 0}$	—	—	0.060	—	—	—	0.070	—
			(.011)				(.012)	
GOF	934.4	813.5	794.7	427.2	3747.9	3730.7	3651.2	727.9
(df)	(99)	(98)	(96)	(84)	(99)	(98)	(96)	(84)

Notes: Standard errors are in parentheses. The models are fit to the 105 distinct second moments of ΔY_{Sit} and ΔY_{Lit}. Model (1) assumes LEED earnings are measured without error, and has 6 parameters; model (2) allows classical measurement error in LEED earnings, and has 7 parameters; model (3) has 9 parameters, including separate classical measurement error variances in SoFIE and LEED earnings for the end-years; and model (4) has 21 parameters, including year-specific classical measurement error variances in SoFIE and LEED earnings. The GOF-statistics are based on all 105 second moments.

Table 7: Predictions for earnings inequality

	Males		Females	
Variance:	SoFIE	LEED	SoFIE	LEED
1. True log(earnings) changes:				
σ^2_η	0.031		0.085	
Var(Δu_{it})	0.007		0.044	
Total	0.038		0.129	
2. Observed log(earnings) changes:				
$((1+\delta_{1S})^2+\delta^2_{1S}).\sigma^2_{S\eta}$	0.037		0.107	
$(1+\delta_{2S})^2.Var(\Delta u_{it)}$	0.018		0.061	
$2.\sigma^2_{.\nu}$	0.091	0.050	0.137	0.059
Total	0.145	0.105	0.305	0.226
3. Observed log(earnings) levels:				
Avg var(log(earnings))	0.164	0.135	0.424	0.352
Max $\sigma^2_{\lambda.i}$	0.012	0.012	0.017	0.017
Max var(meas. error)	0.072	0.037	0.121	0.046
Min var(true log(earnings))	0.092	0.098	0.303	0.320
Var(u_{it})	0.006		0.039	
Fraction of true var	0.061	0.057	0.127	0.126

Notes: Predictions based on model (2) in Table 6.