ROCKWOOL Foundation Berlin
Institute for the Economy and the Future of Work (RFBerlin)

DISCUSSION PAPER SERIES 48/25

Errors in survey and administrative data on employment earnings: Austria and the United Kingdom compared

Christopher R. Bollinger, Stephen P. Jenkins, Fernando Rios-Avila, Iva V. Tasseva

www.rfberlin.com

AUGUST 2025

Errors in survey and administrative data on employment earnings: Austria and the United Kingdom compared

Authors

Christopher R. Bollinger, Stephen P. Jenkins, Fernando Rios-Avila, Iva V. Tasseva

Reference

JEL Codes: C81, C83, D31

Keywords: measurement error, linkage error, earnings, linked survey data, administrative data, finite mixture models

Recommended Citation: Christopher R. Bollinger, Stephen P. Jenkins, Fernando Rios-Avila, Iva V. Tasseva (2025): Errors in survey and administrative data on employment earnings: Austria and the United Kingdom compared. RFBerlin Discussion Paper No. 48/25

Access

Disclaimer

Imprint

RFBerlin
ROCKWOOL Foundation Berlin –
Institute for the Economy
and the Future of Work

Gormannstrasse 22, 10119 Berlin
Tel: +49 (0) 151 143 444 67
E-mail: info@rfberlin.com
Web: www.rfberlin.com

Errors in survey and administrative data on employment earnings: Austria and the United Kingdom compared

Christopher R. Bollinger
(University of Kentucky, and IZA)

Stephen P. Jenkins
(LSE, RF Berlin, IZA, and EUI)

Fernando Rios-Avila
(Universidad Privada Boliviana)

Iva V. Tasseva
(LSE)

28 July 2025

Abstract

We contribute new cross-national evidence about the nature of measurement errors in employment earnings, fitting the same error components model to harmonised earnings data for Austria and the UK. The model allows for measurement error in the administrative data and linkage error as well as survey measurement error. We find several cross-national similarities in error structure but also intriguing differences in error component probabilities, means, and dispersions.

Keywords: measurement error, linkage error, earnings, linked survey data, administrative data, finite mixture models

JEL codes: C81, C83, D31

Acknowledgements. Thanks to Statistics Austria and the UK Department for Work and Pensions for making their confidential linked datasets available to us. This paper was drafted when Jenkins was a Visiting Fellow at the Department of Economics, European University Institute.

Correspondence

Jenkins: Department of Social Policy, London School of Economics and Political Science, Houghton Street, London, WC2A 2AE, United Kingdom. Email: s.jenkins@lse.ac.uk

1. Introduction

We contribute new cross-national evidence about the nature of measurement errors in employment earnings, fitting the same error components models to harmonised earnings data for Austria and the United Kingdom (UK).

Cross-national comparative research is valuable because it documents similarities and differences across contexts, stimulating questions about their causes. Harmonised methods and data are an essential foundation for such research. Otherwise, apparent similarities and differences in findings may be artefacts rather than reflecting genuine institutional and societal variation.

Cross-national studies of earnings measurement error are rare, with most research focusing on the USA (see Jenkins and Rios-Avila 2023, Table 1). Although there are estimates of one specific model, Kapteyn and Ypma's (2007) 'full' model, for three countries, they are derived from datasets with different sample and income definitions. Kapteyn and Ypma (2007), hereafter 'KY', use Swedish survey data on 2002 employment earnings for individuals aged 50+ years. Jenkins and Rios-Avila (2020) use UK survey data on employment earnings in financial year 2011/12 for individuals aged 16+ years. Cavicchioli and Lala (2022) use survey data on individual taxable income in 2001 for individuals from the Italian city of Modena, without an age restriction. The administrative data sources differ in nature across the studies too.

KY were the first to allow for errors in the linkage of survey and administrative data (mismatch) in addition to survey data measurement errors. Thus, the difference between a linked survey and administrative report no longer reflects a single error type, and KY show that the impact of measurement error on inequality and regression coefficients depends on error magnitude and direction. Moreover, previous research treating linked data as error-free invariably found mean-reverting survey errors (over-reporting at the bottom, under-reporting at the top of the earnings distribution). In contrast, KY found negligible mean reversion, "thereby overturning a conventional wisdom" (Jenkins and Rios-Avila, 2021, p. 474). Jenkins and Rios-Avila (2020) and Cavicchioli and Lala (2022) also report negligible mean reversion in survey responses when they fit the KY full model. But researchers may be reluctant to claim the negligible mean reversion finding is general given the differences in the samples and income variables.

We address this issue by using harmonised datasets for Austria and the UK. To each, we fit the 'Extended KY' model of Jenkins and Rios-Avila (2023) which allows for measurement error in the administrative data in addition to linkage error. We also fit the KY full model, nested within Extended KY model. We discuss expected differences in measurement error structure between Austria and the UK in the Data section. Given previous research, one parameter of specific interest is the degree of mean-reversion in survey earnings.

The measurement error models, data, and estimates, appear in the following sections. We find several cross-national similarities in error structure but also intriguing differences in error component probabilities, means, and dispersions.

2. The 'Extended KY' measurement error model

Our exposition draws on Jenkins and Rios-Avila (2023), who also discuss identification and estimation details.

Latent variable ξ_i represents the true variable of interest, log earnings, for each individual $i = 1, \ldots, N$. There are two measures of ξ_i, each potentially error-ridden: one from administrative data, r_i, and one from survey data, s_i.

The distribution of s is a mixture of three observation types: see (1). For $S1$, s_i equals true earnings with probability π_s. For $S2$, s_i contains mean-reverting error with probability $(1-\pi_s)(1-\pi_\omega)$, and ρ_s is the mean-reversion parameter (correlation between the error and true log earnings). S3 observations are subject to contamination error (ω_i) in addition to survey measurement error, with probability $(1-\pi_s)\pi_\omega$.

$$s_i = \begin{cases} \xi_i \quad \text{with probability } \pi_s & (S1) \\ \xi_i + \rho_s(\xi_i - \mu_\xi) + \eta_i \quad \text{with probability } (1-\pi_s)(1-\pi_\omega) & (S2) \\ \xi_i + \rho_s(\xi_i - \mu_\xi) + \eta_i + \omega_i \quad \text{with probability } (1-\pi_s)\pi_\omega. & (S3) \end{cases} \qquad (1)$$

The distribution of r is also a mixture of three observation types: see (2). A key distinction is between individuals for whom the linkage is correct (probability π_r) and individuals who are incorrectly linked (probability $1-\pi_r$). Among the former, r_i is either equal to true log earnings with probability π_υ (type $R1$) or measured with error with probability $1-\pi_\upsilon$ (type $R2$). Among $R2$ observations, ρ_r summarizes mean reversion in r_i. R3 observations include linkage error: the linked administrative data represent the log earnings not of the survey respondent as intended but of someone else (mismatch), denoted ζ_i. Their distribution represents an unknown subset of observations in the administrative database.

$$r_i = \begin{cases} \xi_i \quad \text{with probability } \pi_r \pi_\upsilon & (R1) \\ \xi_i + \rho_r(\xi_i - \mu_\xi) + \upsilon_i \ \text{ with probability } \pi_r(1-\pi_\upsilon) & (R2) \\ \zeta_i \quad \text{with probability } (1-\pi_r) & (R3) \end{cases} \qquad (2)$$

There are nine observation types (latent classes) in the linked dataset corresponding to the combinations of $\{S1, S2, S3\}$ and $\{R1, R2, R3\}$. Table 1 lists these and their probabilities. The KY full model arises when $1-\pi_\upsilon = 0$, resulting in only six latent classes.

<Table 1 near here>

We assume true log earnings, incorrectly-linked log earnings, and errors are independently normally distributed with mean μ_θ and standard deviation (SD) σ_θ, where $\theta \in \{\xi, \zeta, \upsilon, \eta, \omega\}$. We fit the Extended KY and KY full models by maximum likelihood treating observations with $|r_i - s_i| < 0.005$ as 'equal' and hence error-free.

3. Data

Our data were created by linking individual earnings records from nationally representative household survey respondents to their tax records. The Austrian survey is the EU Statistics on Income and Living Conditions; the UK survey is the Family Resources Survey. For both countries, the administrative data arise from employer reporting used in national systems for social insurance contributions and earnings withholding. For further details, see Bollinger and Tasseva (2025) and Jenkins and Rios-Avila (2023).

The datasets refer to similar time periods (survey year 2011 for Austria, 2011/12 financial year for the UK), and total gross (pre-tax) employment earnings from all employments. Each analysis dataset contains individuals aged 16+ years with positive earnings, excluding observations with imputed survey earnings.

The resulting samples contain 5,971 (UK) and 5,662 (Austria) individuals, of whom 3.4% and 3.5% have error-free earnings, respectively. Summary statistics for (s, r) are: mean (9.97, 9.98) and SD (0.98, 1.04) for Austria; mean (9.77, 9.75) and SD (0.81, 0.84) for the UK. The mean and median differences of $(s_i - r_i)$ are zero for both countries.

Data collection processes differ between Austria and the UK in three ways likely to impact measurement error estimates. First, administrative record linkage in the UK required survey respondents' consent (65% of employees provided it), and used first and last names, postcode, sex, and date of birth, as match keys (71% linkage success rate). Austria does not require consent and based linkage on pseudonymised personal identifiers already used nationally, with a very high linkage success rate (99%, Statistics Austria 2014). Given this, we would expect lower linkage error rates ($1-\pi_r$) for Austria, unless consenting UK respondents are sufficiently more likely to be correctly linked conditional on a link being made.

Second, although administrative earnings refer to a 12-month year in both countries, Austrian survey earnings refer to a 12-month year (calendar year 2010) and UK survey earnings typically refer to a shorter reference period. UK respondents were asked about the amount last received prior to the interview and its reference period (most reported 'calendar month'). The UK data producers converted earnings amounts to pounds per week pro rata, which we annualised. Jenkins and Rios-Avila (2023) view the survey-tax data reference period mismatch as a contamination error, the importance of which depends on individuals' earnings stability over the year. Thus, the probability of contamination error (π_ω) may be larger for the UK, but its shorter and hence more salient reference period may reduce error in survey reporting (i.e., μ_η and σ_η smaller in magnitude in the UK).

Third, how employers reported earnings to the tax authorities differed. Austria uses electronic reporting. The UK in 2011/12 allowed paper-based or electronic reporting, combined with a year-end return, our tax data source. Compared to Austria, and assuming digitisation enhances accuracy, we expect the UK system to lead to measurement error in r with greater probability ($1-\pi_\upsilon$), and μ_υ and σ_υ larger in magnitude.

4. Results

Estimates of the Extended KY and KY full models appear in Table 2 (model parameters) and Table 3 (latent class probabilities). Starting with the Extended KY model, for both Austria and the UK, mean r and s are similar and less than estimated mean true log earnings. The SDs of r and s are also similar but larger than the estimated SD of true log earnings, consistent with the Classical measurement error model (with only survey measurement error). Mean reversion in survey errors (ρ_s) is close to zero and not statistically significant.

<Table 2 near here>

Contrary to expectation, the linkage error probability is similar in Austria and the UK, around 3%. Also similar is the nature of measurement error in r. Estimated SDs are around

0.36, but the mean error and mean-reversion parameters are close to zero (not statistically significant). However, the probability of measurement error in *r* is larger in Austria (43%) than the UK (32%), a surprising finding.

For survey measurement error there are cross-national similarities and differences. Although the probability of measurement error ($1-\pi_s$) is similar in both countries (94%), the probability of contamination error is around 2½ times larger in the UK (11% compared to 4%), the direction we expected. The distributions of measurement error also have broadly similar means and SDs. But there are more distinct differences in the distribution of the contamination component. Austria has a substantially lower estimated mean (–1.22 compared to –0.27), i.e. Austrian survey responses under-estimate true log earnings by a greater extent. Also, the estimated SD of the contamination error is smaller for the UK than Austria (1.01 compared to 1.62). These findings are consistent with our earlier argument that the nature of the UK survey question – asking for the most recent amount received rather than annual earnings – boosts accuracy.

Latent class probability estimates are similar, with the sum of the three largest probabilities 89% for both countries. Class 2, with no measurement or linkage error in *r* but measurement error in *s*, has the largest probability (50% for Austria, 56% for the UK) followed by class 5, which is as class 2 but with measurement error in *r* (37% for Austria, 26% for the UK). The next largest class is small by comparison. Class 3 is as class 2 but also has contamination error (3% for Austria, 2% for the UK).

<Table 3 near here>

Table 2's bottom rows show that, for both countries, the linked administrative data are less reliable than the survey data. (Reliability is measured by the squared correlation between true log earnings and observed earnings (*r* or *s*): see Meijer et al., 2012.)

Tables 2 and 3 also report KY full model estimates. The goodness of fit statistics in Table 2 show that this model is rejected in favour of the Extended KY model. However, most qualitative differences across countries are replicated with the simpler model, e.g., the larger probability of contamination error in the UK survey data, differences in contamination error means, and similar linkage error probabilities. One difference is that the KY full model estimates of the survey mean-reversion parameter (ρ_s) are larger in magnitude and statistically significant – but they remain negligible.

In sum, there are similarities between the structure of measurement errors in Austrian and UK data on employment earnings – more than we expected given the differences in the data collection process – but also notable differences. For example, although the probability of contamination error is larger in the UK, Austrian survey earnings underestimate true earnings by more than the UK survey earnings on average. And the probability of measurement error in linked administrative earnings is larger for Austria than for the UK, suggesting the quality of earnings reporting by Austrian employers warrants further investigation. The UK has now changed to real-time electronic reporting by employers (Office for National Statistics, 2019), so our cross-national comparison should be re-run when new linked earnings data become available.

References

Bollinger, C. R. and Tasseva, I. V. (2023). Income source confusion using the SILC. *Public Opinion Quarterly*, 87, 542–574. https://doi.org/10.1093/poq/nfad025

Cavicchioli, L., and Lalla, M. (2022). Evidences from survey data and fiscal data: nonresponse and measurement errors in annual incomes. *Statistical Methods & Applications*, 31, 587–615. https://doi.org/10.1007/s10260-021-00593-3

Jenkins, S. P. and Rios-Avila, F. (2020). Modelling errors in survey and administrative data on labour earnings: sensitivity to the fraction assumed to have error-free earnings. *Economics Letters*, 192, 109253. https://doi.org/10.1016/j.econlet.2020.109253

Jenkins, S. P. and Rios-Avila, F. (2021). Measurement error in earnings data: replication of Meijer, Rohwedder, and Wansbeek's mixture model approach to combining survey and register data. *Journal of Applied Econometrics*, 36, 474–483. https://doi.org/10.1002/jae.2811

Jenkins, S. P. and Rios-Avila, F. (2023). Reconciling reports: modelling employment earnings and measurement errors using linked survey and administrative data. *Journal of the Royal Statistical Society (Series A: Statistics in Society)*, 186, 110–136. https://doi.org/10.1093/jrsssa/qnac003

Kapteyn, A. and Ypma, J. Y. (2007). Measurement error and misclassification: a comparison of survey and administrative data. *Journal of Labor Economics*, 25, 513–551. https://doi.org/10.1086/513298

Meijer, E., Rohwedder, S. and Wansbeek, T. (2012). Measurement error in earnings data: using a mixture model approach to combine survey and register data. *Journal of Business and Economic Statistics*, 30, 191–201. https://doi.org/10.1198/jbes.2011.08166

Office for National Statistics (2019). New methods for monthly earnings and employment estimates from Pay As You Earn Real Time Information (PAYE RTI) data: December 2019. London: Office for National Statistics.

Statistics Austria (2014). *Methodenbericht zur Rückrechnung von EU-SILC 2008-2011 auf Basis von Verwaltungsdaten*. Vienna: Statistics Austria.

Table 1. Latent classes in Extended KY model of survey and administrative earnings

Class	Description	Types
1	No error in *r*; no error in *s*	*R*1, *S*1
2	No error in *r*; measurement error in *s*	*R*1, *S*2
3	No error in *r*; measurement and contamination error in *s*	*R*1, *S*3
4	Measurement error in *r*; no error in *s*	*R*2, *S*1
5	Measurement error in *r*; measurement error in *s*	*R*2, *S*2
6	Measurement error in *r*; measurement error and contamination error in *s*	*R*2, *S*3
7	Mismatched *r*; no error in *s*	*R*3, *S*1
8	Mismatched *r*; measurement error in *s*	*R*3, *S*2
9	Mismatched *r*; measurement error and contamination error in *s*	*R*3, *S*3

Notes. The KY full model has 6 latent classes (no type *R*2).

Table 2. Estimates of earnings measurement error model parameters, Austria and UK

		Extended KY model		KY full model	
Label	Parameter	Austria	UK	Austria	UK
True log earnings: mean	μ_ξ	10.057***	9.808***	10.025***	9.811***
		(0.014)	(0.011)	(0.017)	(0.010)
True log earnings: SD	σ_ξ	0.870***	0.724***	0.923***	0.757***
		(0.017)	(0.009)	(0.025)	(0.010)
Mismatched r: mean	μ_ζ	7.526***	8.094***	7.838***	8.621***
		(0.182)	(0.169)	(0.190)	(0.135)
Mismatched r: SD	σ_ζ	1.490***	1.230***	1.567***	1.288***
		(0.090)	(0.086)	(0.084)	(0.062)
Measurement error in s: mean	μ_η	–0.040***	–0.010***	–0.214***	–0.124***
		(0.005)	(0.003)	(0.165)	(0.024)
Measurement error in s: SD	σ_η	0.123***	0.094***	0.846***	0.637***
		(0.011)	(0.007)	(0.245)	(0.071)
Contamination error in s: mean	μ_ω	–1.222***	–0.266***	–0.024***	–0.009***
		(0.363)	(0.048)	(0.004)	(0.003)
Contamination error in s: SD	σ_ω	1.625***	1.008***	0.183***	0.114***
		(0.329)	(0.105)	(0.017)	(0.006)
Measurement error in r: mean	μ_υ	–0.051*	–0.035		
		(0.024)	(0.034)		
Measurement error in r: SD	σ_υ	0.362***	0.363***		
		(0.025)	(0.026)		
Mean reversion in s	ρ_s	–0.012	0.007	–0.055***	–0.019***
		(0.007)	(0.004)	(0.005)	(0.004)
Mean reversion in r	ρ_r	0.008	0.091		
		(0.030)	(0.061)		
Pr(measurement error in s)	$1-\pi_s$	0.937***	0.948***	0.963***	0.964***
		(0.007)	(0.005)	(0.003)	(0.003)
Pr(contamination error in s)	π_ω	0.044***	0.109***	0.159**	0.261***
		(0.008)	(0.021)	(0.059)	(0.023)
Pr(linkage error)	$1-\pi_r$	0.036***	0.029***	0.042***	0.064***
		(0.005)	(0.006)	(0.008)	(0.007)
Pr(measurement error in r)	$1-\pi_\upsilon$	0.426***	0.319***		
		(0.053)	(0.046)		
Log(pseudo-likelihood)		–10141.9	–8805.2	–10376.8	–9034.3
AIC		20315.8	17642.3	20777.7	18092.6
BIC		20422.1	17749.5	20857.4	18173.0
Number of households		3,708	4,874	3,708	4,874
Number of individuals		5,662	5,971	5,662	5,971
Reliability(r)		0.653	0.741	0.705	0.691
Reliability(s)		0.803	0.816	0.838	0.825

Notes. The Extended KY and KY full models are fitted by maximum likelihood, treating observations with $|r_i - s_i| < 0.005$ as ‘equal’ and hence error-free. Cluster-robust standard errors in parentheses (clusters are households). * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. Reliability(x) is the squared correlation between true earnings and observed earnings measure, $x \in \{r, s\}$.

Table 3. Estimates of latent class probabilities, Austria and UK

	Extended KY model		KY full model	
Latent class probabilities	Austria	UK	Austria	UK
$\pi_1 = \pi_r\, \pi_\upsilon\, \pi_s$	0.035***	0.034***	0.035***	0.034***
	(0.003)	(0.002)	(0.003)	(0.002)
$\pi_2 = \pi_r\, \pi_\upsilon\, (1-\pi_s)(1-\pi_\omega)$	0.496***	0.559***	0.776***	0.667***
	(0.050)	(0.034)	(0.049)	(0.024)
$\pi_3 = \pi_r\, \pi_\upsilon\, (1-\pi_s)\pi_\omega$	0.023***	0.069***	0.1468**	0.235***
	(0.005)	(0.017)	(0.056)	(0.020)
$\pi_4 = \pi_r\, (1-\pi_\upsilon)\pi_s$	0.026***	0.016***		
	(0.006)	(0.004)		
$\pi_5 = \pi_r\, (1-\pi_\upsilon)(1-\pi_s)(1-\pi_\omega)$	0.368***	0.262***		
	(0.043)	(0.041)		
$\pi_6 = \pi_r\, (1-\pi_\upsilon)(1-\pi_s)\pi_\omega$	0.017***	0.032***		
	(0.004)	(0.004)		
$\pi_7 = (1-\pi_r)\pi_s$	0.002***	0.002***	0.002***	0.002***
	(0.001)	(0.000)	(0.000)	(0.000)
$\pi_8 = (1-\pi_r)(1-\pi_s)(1-\pi_\omega)$	0.032***	0.025***	0.034***	0.046***
	(0.005)	(0.005)	(0.009)	(0.005)
$\pi_9 = (1-\pi_r)(1-\pi_s)\pi_\omega$	0.002***	0.003***	0.006***	0.016***
	(0.000)	(0.001)	(0.001)	(0.003)

Note. The latent classes are described in Table 1. Otherwise, as for Table 2.