J Econ Inequal (2013) 11:343–372
DOI 10.1007/s10888-012-9216-5

Identification of the covariance structure of earnings using the GMM estimator

Aedín Doris · Donal O'Neill · Olive Sweetman

Received: 12 May 2010 / Accepted: 16 January 2012 / Published online: 3 March 2012

Abstract In recent years there has been a rapid growth in the number of studies that have used the GMM estimator to decompose the earnings covariance structure into its permanent and transitory parts. Using a heterogeneous growth model of earnings, we consider the performance of the estimator in this context. We use Monte Carlo simulations to examine the sensitivity of parameter identification to key features such as panel length, sample size, the degree of persistence of earnings shocks and the specification of the earnings model. We show that long panels allow the identification of the model, even when persistence in transitory shocks is high. Short panels, on the other hand, are insufficient to identify individual parameters of the model even with moderate levels of persistence.

Keywords Identification · GMM · Covariance structure of earnings

JEL Classification J31 · D31

1 Introduction

In recent years there has been a rapid growth in the number of studies that have used the Generalised Method of Moments (GMM) estimator to estimate the covariance structure of earnings. In these models, earnings are written as the sum

A. Doris · D. O'Neill (✉) · O. Sweetman
Department of Economics, Finance and Accounting, Rhetoric House,
National University of Ireland Maynooth,
Maynooth, Co. Kildare, Ireland
e-mail: Donal.Oneill@nuim.ie

D. O'Neill
IZA, Bonn, Germany

of permanent and transitory components and the estimated parameter values are then used to construct measures of permanent and transitory inequality and to trace their evolution over time. Distinguishing between these two components is important because they have different policy implications; moreover the distinction can provide insight into the functioning of the labour market.

The GMM estimator uses panel data to estimate these models by matching the sample variances and covariances of earnings to their population counterparts. The model is identified from the long covariances. In these latter moments, the contribution of the transitory shock is negligible, which in turn allows researchers to recover the parameters associated with the permanent component. However long panels are not always available to researchers and as a result a number of recent studies, for example Ramos [34], Doris et al. [13, 14], Cervini and Ramos [9] and Sologon and O'Donoghue [36], have been constrained to use relatively short panels, with eight or nine years of data. It is unclear whether panel lengths of this order are sufficient to identify these models.

Although the performance of the GMM estimator has been evaluated elsewhere (e.g. Tauchen [40], Kocherlakota [26], Hansen et al. [23], Altonji and Segal [1], Clark [10], Stock and Wright [39], Blundell and Bond [4] and Pozzi [33]), as yet there has been no detailed study of the estimator for the type of earnings covariance models or the data structures often found in the empirical literature. In this paper, we use Monte Carlo techniques to consider identification of these models and discuss the consequences for estimation and inference.

2 The GMM approach to estimating earnings covariance structures

The GMM approach to parameter estimation is now well established in the econometric literature, having been introduced by Hansen [22]. Hall [20] and Cameron and Trivedi [5] provide comprehensive discussions of the approach. GMM estimation entails minimizing a criterion function which measures the distance between population and sample moment expressions. The GMM estimator will identify the model if the probability limit of the GMM criterion function is uniquely minimised at the true parameter vector, φ_0. The order condition for identification requires that the number of moment conditions, k, exceeds the number of parameters, p. The rank condition requires that the information provided by the k moment conditions, $E\left[m(yi;\varphi)\right]=0$, must differ; that is, as the p components of φ vary in the neighbourhood of φ_0, the k components of $E[m(y_i;\varphi)]$ vary in p independent directions (see for example [20], Chapter 3). Equivalently, the matrix $E\left[\left.\frac{\delta m}{\delta\varphi'}\right|_{\varphi_0}\right]$ must be of full column rank. If a model is not identified, there exist at least two distinct data generating processes (DGPs), characterised by different parameter vectors, which cannot be distinguished by any function of the data, even with infinitely large samples.

In recent times there has been a growing interest in problems of weak identification in econometric models [38]. Weak identification occurs when the moment condition is not zero but still very small at parameter values other than φ_0. This gives rise to criterion functions with ridges or near flat spots in the region of the true parameter vector, φ_0. Stock and Wright [39] show that the asymptotic theory devised for the GMM estimator when applied to identified models is not valid when the model is weakly identified, even for very large but finite sample sizes. They illustrate their findings using both a simultaneous equation model and a consumption based asset pricing model. GMM estimation of the consumption based asset pricing model was also analysed in studies by Tauchen [40] and Kocherlakota [26]. Blundell and Bond [4] consider GMM estimation of a production function with persistent data and short panels. They argue that the poor performance of the GMM estimator in this context reflects a weak instrument problem. In our paper, we also consider GMM identification with panel data but in a different context, where the objective is to estimate the covariance structure of earnings.

The standard approach to estimating the covariance structure of earnings is to write earnings as the sum of two components. The first is a permanent component, which, once acquired, is maintained throughout an individual's working career. This may reflect pre-labour market characteristics such as the level of education and/or shocks that have permanent effects, such as involuntary job loss. The second component is a transitory one, reflecting temporary shocks that are mean-reverting. The objective is to measure the separate roles played by the permanent and transitory shocks in determining inequality and to examine how this may have changed over time. Table 1 provides an overview of the range of models used to date in the empirical literature. These papers vary in the specification of both the permanent and transitory components of the earnings process. When modelling the permanent component, it is typical to use either a Random Walk (RW) model [12], where individuals are subject to persistent shocks while facing similar life-cycle profiles, or a Random Growth (RG) model [2, 11, 18, 19], where the experience-earnings profile is individual-specific. A small number of papers have estimated covariance models that combine both the Random Walk model and the Random Growth model (RG+RW) [24, 31, 34]. When modelling the transitory component of earnings, researchers have recognised the importance of allowing for persistence in this component. This is generally modelled by specifying either an AR or ARMA process for the error term. Many papers also allow for the possibility that the earnings process is affected by the age/experience cohort to which the individual belongs.

In this paper, we consider earnings models that are representative of the variety of models shown in Table 1. Initially we consider an RG model and an ARMA process in the transitory component, although later we extend the model to include both a random walk and a random growth element in the permanent component, and cohort effects.

Earnings for individual i, with x years of experience at time t, y_{ixt}, are given by

$$y_{ixt} = p_t \alpha_{ix} + \lambda_t v_{it} \tag{1}$$

where $E(\alpha_{ix}) = E(v_{it}) = 0$. The factor loadings, p_t and λ_t, allow variances to change over time in a way that is common across individuals. The inclusion of factor loadings [29] was an important innovation in the modelling of earnings dynamics in that they allow for time trends in the earnings process. For the RG model, the permanent component is

$$\alpha_{ix} = \alpha_{i(x-1)} + \beta_i \tag{2a}$$

where $E(\beta_i) = 0$. The random growth terms α_{i0} and β_i have variances σ_α^2 and σ_β^2 respectively and covariance $\sigma_{\alpha\beta}$. Thus each individual may have a different permanent life-cycle growth rate of earnings, β_i, which may be correlated with

Table 1 Summary of literature

Author, Year	T	Mean N per year	Permanent component	Transitory component	Cohort effects	ρ	θ	σ_{α}^{2}	σ_{β}^{2}	$\sigma_{\alpha\beta}$	σ_{w}^{2}	σ_{v1}^{2}	σ_{ε}^{2}
Moffitt & Gottschalk [31]	18	1,400	RW	ARMA(1,1)	N	0.641	−0.367	0.056	–	–	0.001	–	0.117
Baker [2]	20	534	RG	ARMA(1,2)	N	0.674	−0.187	0.139	0.0004	−0.004	–	n.r.	n.r.
	20	534	RG	ARMA(1,1)	N	0.745	−0.260	0.102	0.0003	−0.003	–	n.r.	n.r.
Dickens [12]	21	62,637	RW	ARMA(1,1)	Y	0.956	−0.569	–	–	–	0.001–0.007	–	0.024
Cappellari [6]	15	10,605	RG	AR(1)	N	0.801	–	0.202	0.0003	−0.009	–	0.097	0.155
Haider [19]	25	1,400	RG	ARMA(1,1)	N	0.639	−0.264	0.295	0.0004	−0.008	–	n.r.	n.r.
Cappellari [7]													
Public sector	13	49,541	RG	AR(1)	Y	0.536	–	0.005	0.0000	0.000	–	0.004	0.014
Private sector	13	14,031	RG	AR(1)	Y	0.841	–	0.008	0.0002	0.000	–	0.028	0.011
Ramos [34]	9	2,518	RG+RW	AR(1)	Y	0.30	–	0.27	0.0003	−0.01	0.005	0.01–0.14	0.10
Baker & Solon [3]	17	25,230	RG+RW	AR(1)	Y	0.540	–	0.134	0.0001	−0.003	0.007	n.r.	–
Cappellari [8]	17	55,020	RG	AR(1)	Y	0.941	–	0.021	0.0002	0.002	–	0.0452	–
Kalwij & Alessie [25]	27	66,000	RW	ARMA (1,4)	Y	0.980	−0.500	0.001	–	–	n.r.	n.r.	n.r.
Gustavsson [16]	9	37,167	RW	AR(1)	Y	0.555	–	0.047	–	–	0.005, 0.001	0.031–0.369	–
Santos & Souza [35]	9	342,450	RG	ARMA(1,1)	N	0.404	−10.40	0.240	0.0009	−0.003	–	–	0.001
Moffitt & Gottschalk [29]	30	1,035	RG+RW	ARMA(1,1)	N	0.847	−0.574	0.0901	0.0000	−0.002	0.003	–	–
Cervini-Pla & Ramos [9]	8	2,576	RG	AR(1)	Y	0.583	–	0.524	0.0002	−0.01	–	0.292–1.061	0.068
Lilla & Staffolani [27]													
White collar	15	19,658	RG	AR(1)	Y	0.186	–	0.397	0.0004	−0.012	–	0.439	0.345
Blue collar	15	37,981	RG	AR(1)	Y	0.118	–	1.499	0.0005	−0.027	–	0.558	0.226
Myck, Ochmann & Qari [32]	13	952	RE	ARMA(1,1)	Y	0.839	−0.431	0.098	–	–	–	0.054	0.033
Gustavsson [17]	31	36,925	RW	ARMA	Y	0.573, 0.819	−0.258, −0.505	0.085	–	–	0.010–0.017	–	0.053
Doris, O'Neill & Sweetman [13]	8		RE	ARMA(1,1)	N	0.635	−.230	0.138	–	–	–	0.112	0.045

Guvenen [18]	26	n.r.	RG	AR(1)	N	0.821	–	0.022	0.0004	−0.23	–	0.047	0.029
Hoffmann [24]													
Low education	30	13,980	RG+RW	AR(1)	Y	0.887	–	0.018	0.0000	0.0000	0.0007	0.033	
Medium education	29	164,369	RG+RW	AR(1)	Y	0.862	–	0.030	0.0000	0.0002	0.0011	0.009	
High education	29	17,292	RG+RW	AR(1)	Y	0.847	–	0.029	0.0000	0.0002	0.0005	0.025	n.r.
Sologon & O'Donoghue [36] [a]													
Germany	8	n.r.	RG	AR(1)	Y	0.358	–	7.261	0.0024	−0.131	–	0.004–0.083	0.258
Denmark	8	n.r.	RW	AR(1)	Y	0.547	–	0.0097	–	–	0.0014	0.025–0.037	0.131
Netherlands	8	n.r.	RG	AR(1)	Y	0.329	–	0.1913	0.0002	−0.005	–	0.011–0.041	0.126
Belgium	8	n.r.	RE	AR	Y	0.628	–	0.0698	–	–	–	0.036–0.064	0.244
France	8	n.r.	RE	AR	Y	0.399	–	0.1653	–	–	–	0.049–0.104	0.797
Luxembourg	7	n.r.	RE	AR	Y	0.239	–	0.1071	–	–	–	0.022–0.106	0.019
UK	8	n.r.	RG	AR	Y	0.451	–	0.0467	0.0001	−0.002	–	0.031–0.079	0.070
Ireland	8	n.r.	RG	AR	Y	0.291	–	0.0564	0.0002	−0.003	–	0.069–0.094	0.028
Italy	8	n.r.	RG	ARMA(1,1)	Y	0.644	−0.251	0.0325	0.0001	−0.001	–	0.028–0.052	0.058
Greece	8	n.r.	RG	ARMA(1,1)	Y	0.599	0.149	0.0779	0.0002	−0.003	–	0.057–0.101	0.118
Spain	8	n.r.	RG	ARMA(1,1)	Y	0.849	−0.364	0.2940	0.0000	−0.006	–	0.052	0.099
Portugal	8	n.r.	RE	AR	Y	0.778	–	0.2561	–	–	–	0.043	0.258
Austria	7	n.r.	RE	AR	Y	0.701	–	0.0811	–	–	–	0.075	0.483
Finland	6	n.r.	RG	AR	Y	0.290	–	0.0616	0.0001	−0.002	–	0.046–0.071	0.055
Doris, O'Neill & Sweetman [14]													
Germany	8	2750	RE	ARMA(1,1)	N	0.630	−0.340	0.100	–	–	–	0.08	0.050
Denmark	8	1180	RE	AR	N	0.530	–	0.060	–	–	–	0.060	0.030
Netherlands	8	2327	RE	AR	N	0.410	–	0.090	–	–	–	0.060	0.090
Belgium	8	1218	RE	AR	N	0.530	–	0.070	–	–	–	0.050	0.060
France	8	2447	RE	ARMA(1,1)	N	0.550	−0.240	0.090	–	–	–	0.080	0.070
UK	8	1798	RE	ARMA(1,1)	N	0.830	−0.310	0.100	–	–	–	0.110	0.070
Ireland	8	1174	RE	AR	N	0.300	–	0.160	–	–	–	0.150	0.020
Italy	8	2614	RE	AR	N	0.440	–	0.060	–	–	–	0.050	0.030
Spain	8	2174	RE	ARMA(1,1)	N	0.480	−0.360	0.160	–	–	–	0.080	0.060

Table 1 (continued)

Author, Year	T	Mean N per year	Permanent Component	Transitory Component	Cohort Effects	ρ	θ	σ^2_α	σ^2_β	$\sigma_{\alpha\beta}$	σ^2_w	σ^2_{v1}	σ^2_ε
Portugal	8	2058	RE	AR	N	0.590	–	0.160	–	–	–	0.090	0.070
Austria	7	1380	RE	ARMA(1,1)	N	0.820	−0.310	0.060	–	–	–	0.100	0.070
Finland	6	922	RE	AR	N	0.420	–	0.100	–	–	–	0.090	0.080
Sologon & O'Donoghue [37]	17	90,295	RW	ARMA(1,1)	Y	0.964	−0.206	0.0122	–	–	0.0004–0.0106	0.000–0.060	0.000

n.r. means 'not reported'. 'Mean N per year' refers to the average number of individuals in each wave of the panel data. RE is 'Random Effect' – a model with an individual-specific intercept, but no heterogeneity in the slope of the earning function; RG and RW as explained in text. 'Cohort Effects' indicates that parameters capturing cohort effects were also estimated in the model. Where authors report more than one specification, but identify a 'preferred model', only that model is reported. Where two estimates or a range of estimates are presented, this is because the model allowed for variation either in age or by cohort in the relevant parameter

[a] Weighted number of observations given; data is ECHP, which typically give 1,000 to 5,000 individuals, depending on the country

initial earnings, α_{i0}. Persistence in the transitory shock, v_{it}, is modelled using an ARMA(1,1) process with AR parameter ρ and MA parameter θ. Specifically,

$$v_{it} = \rho v_{i(t-1)} + \varepsilon_{it} + \theta\varepsilon_{it-1} \tag{2b}$$

where ε_{it} is a random variable with variance σ_ε^2. The recursive nature of the transitory process requires consideration of initial conditions. Since the assumption of an infinite history is untenable in this context, we follow the approach suggested by MaCurdy [28] and widely adopted in the literature. This approach treats the variance of the initial condition at the start of our sample period, σ_{v1}^2, as an additional parameter to be estimated.

The GMM estimator matches sample variances and covariances to their population counterparts. In the model specified by Eqs. 1, 2a and 2b the true variance-covariance matrix at time t has diagonal elements:

$$\begin{aligned}
\sigma_1^2 &= \left[p_1^2(\sigma_\alpha^2 + \sigma_\beta^2 X_1^2 + 2\sigma_{\alpha\beta}X_1)\right] + \left[\lambda_1^2\sigma_{v1}^2\right] \quad \text{for } t = 1 \\
\sigma_t^2 &= \left[p_t^2(\sigma_\alpha^2 + \sigma_\beta^2 X_t^2 + 2\sigma_{\alpha\beta}X_t)\right] + \left[\lambda_t^2\left(\rho^{2t-2}\sigma_{v1}^2 + \sigma_\varepsilon^2 K\sum_{w=0}^{t-2}\rho^{2w}\right)\right], \quad \text{for } t > 1
\end{aligned} \tag{3}$$

and off-diagonal elements:

$$\begin{aligned}
&Cov(y_{t,}y_{(t+s)}) = p_t p_{t+s}(\sigma_\alpha^2 + \sigma_\beta^2 X_t X_{(t+s)} + \sigma_{\alpha\beta}(X_t + X_{(t+s)})) + \lambda_t\lambda_{t+s}(\rho^s\sigma_{v1}^2 + \rho^{s-1}\theta\sigma_\varepsilon^2), \\
&\text{for } t = 1,\ s > 0 \\
&Cov(y_{t,}y_{(t+s)}) = p_t p_{t+s}(\sigma_\alpha^2 + \sigma_\beta^2 X_t X_{(t+s)} + \sigma_{\alpha\beta}(X_t + X_{(t+s)})) \\
&\qquad + \lambda_t\lambda_{t+s}\left(\rho^{2t+s-2}\sigma_{v1}^2 + \rho^s\sigma_\varepsilon^2 K\sum_{w=0}^{t-2}\rho^{2w} + \rho^{s-1}\theta\sigma_\varepsilon^2\right),
\end{aligned} \tag{4}$$

for $t > 1$, $s > 0$, where $K = (1 + \theta^2 + 2\rho\theta)$, X_t is average experience at time t, and X_t^2 is the average value of experience-squared at time t.

The parameter vector to be estimated is given by $\varphi = \{\sigma_\alpha^2, \rho, \sigma_\varepsilon^2, \sigma_{v1}^2, p_1 ... p_T, \lambda_1 ... \lambda_T, \sigma_\beta^2, \sigma_{\alpha\beta}, \theta\}$. Identification requires a normalization of the factor loadings and in keeping with the literature we set p_1 and λ_1 equal to one. We then use the estimated parameter vector to recover the individual components of aggregate inequality. The permanent component at time t is given by the corresponding first term in square brackets in Eq. 3 while the second term in square brackets is the transitory component.

GMM estimation requires the selection of an appropriate weighting matrix; based on the work of Altonji and Segal [1] and Clark [10], it is standard practice in the earnings covariance literature to use the identity matrix as the weighting matrix, and we follow this practice here to ensure comparability of our results with the existing literature.

As noted earlier, identification using the GMM estimator requires that the matrix of derivatives of the moment conditions be of full column rank. Key columns of the matrix relevant to our model are presented in Table 2. It is well known that strong persistence in the transitory term, as measured by ρ, may cause problems

Table 2 Derivatives of moment conditions

	$\frac{\partial(\cdot)}{\partial\sigma_\alpha^2}$	$\frac{\partial(\cdot)}{\partial\sigma_\beta^2}$	$\frac{\partial(\cdot)}{\partial\sigma_{\alpha\beta}}$	$\frac{\partial(\cdot)}{\partial\sigma_\varepsilon^2}$	$\frac{\partial(\cdot)}{\partial\sigma_{v1}^2}$	$\frac{\partial(\cdot)}{\partial\rho}$	$\frac{\partial(\cdot)}{\partial\theta}$	$\frac{\partial(\cdot)}{\partial p_1}$	$\frac{\partial(\cdot)}{\partial\lambda_1}$
σ_1^2	p_1^2	$p_1^2X_1^2$	$p_1^2 2X_1$	0	λ_1^2	0	0	$2p_1\left(\sigma_\alpha^2+\sigma_\beta^2X_1^2 +2\sigma_{\alpha\beta}X_1\right)$	$\lambda_2(\sigma_{v1}^2)$
σ_2^2	p_2^2	$p_2^2X_2^2$	$p_2^2 2X_2$	λ_2^2K	$\lambda_2^2\rho^2$	$\lambda_2^2(2\rho\sigma_{v1}^2+\sigma_\varepsilon^2(2\theta))$	$\lambda_2^2(\sigma_\varepsilon^2(2\rho+2\theta))$	**0**	**0**
σ_3^2	p_3^2	$p_3^2X_3^2$	$p_3^2 2X_3$	$\lambda_3^2K(1+\rho^2)$	$\lambda_3^2\rho^4$	$\lambda_3^2(4\rho^3\sigma_{v1}^2+\sigma_\varepsilon^2((2\theta)(1+\rho^2)+2\rho K))$	$\lambda_3^2(\sigma_\varepsilon^2(2\rho+2\theta)(1+\rho^2)))$	**0**	**0**
σ_8^2	p_8^2	$p_8^2X_8^2$	$p_8^2 2X_8$	$\lambda_8^2K\left(1+\rho^2+\rho^4 +\ldots+\rho^{12}\right)$	$\lambda_8^2\rho^{14}$	$\lambda_8^2\left(14\rho^{13}\sigma_{v1}^2+\sigma_\varepsilon^2\left((2\theta)\left(1+\rho^2+\rho^4 +..+\rho^{12}\right)+K(2\rho+4\rho^3+\ldots 12\rho^{11})\right)\right)$	$\lambda_8^2\left(\sigma_\varepsilon^2(2\rho+2\theta)\left(1+\rho^2 +\rho^4+..+\rho^{12}\right)\right)$	**0**	**0**
σ_{12}	p_1p_2	$p_1p_2 X_1X_2$	$p_1p_2 (X_1+X_2)$	$\lambda_1\lambda_2\theta$	$\lambda_1\lambda_2\rho$	$\lambda_1\lambda_2(\sigma_{v1}^2)$	$\lambda_1\lambda_2(\sigma_\varepsilon^2)$	$p_2(\sigma_\alpha^2+\sigma_\beta^2X_1X_2 +\sigma_{\alpha\beta}\ (X_1+X_2))$	$\lambda_2\left(\rho\sigma_{v1}^2 +\theta\sigma_\varepsilon^2\right)$
σ_{23}	p_2p_3	$p_2p_3 X_2X_3$	$p_1p_2 (X_2+X_3)$	$\lambda_2\lambda_3(\rho K+\theta)$	$\lambda_2\lambda_3\rho^3$	$\lambda_2\lambda_3(3\rho^2\sigma_{v1}^2+\sigma_\varepsilon^2(\rho 2\theta+K))$	$\lambda_2\lambda_3(\sigma_\varepsilon^2(\rho(2\theta+2\rho)+1)))$	**0**	**0**
σ_{34}	p_3p_4	$p_3p_4 X_3X_4$	$p_3p_4 (X_3+X_4)$	$\lambda_3\lambda_4(\rho K(1+\rho^2) +\theta)$	$\lambda_3\lambda_4\rho^5$	$\lambda_3\lambda_4(5\rho^4\sigma_{v1}^2+\sigma_\varepsilon^2((\rho+\rho^3)(2\theta) +K(1+3\rho^2))$	$\lambda_3\lambda_4(\sigma_\varepsilon^2((\rho+\rho^3)(2\theta+2\rho) +1)))$	**0**	**0**
σ_{24}	p_2p_4	$p_2p_4 X_2X_4$	$p_2p_4 (X_2+X_4)$	$\lambda_2\lambda_4(\rho^2K+\rho\theta)$	$\lambda_2\lambda_4\rho^4$	$\lambda_2\lambda_4(2\rho\sigma_{v1}^2+\sigma_\varepsilon^2(\rho^2 2\theta+K2\rho+\theta))$	$\lambda_2\lambda_4(\sigma_\varepsilon^2(\rho^2(2\theta+2\rho)+\rho))$	**0**	**0**

of identification in this model. This can be most easily seen by assuming that the time trends on the permanent and transitory components are equal and that $\rho = 1$. In this case, the derivatives of the moment conditions associated with σ_α^2 (Table 2, Column 2) and σ_{v1}^2 (Table 2, Column 6)) are identical, resulting in rank deficiency of the derivative matrix. In practice, whenever ρ is close to one, only the sum of these two parameters will be identified and this is insufficient to allow consistent estimation of the transitory and permanent components of the model. Provided ρ is not equal to one, longer panel lengths will aid identification because the term involving ρ to various powers of t in Column 6 will differ increasingly from one for higher t. Although researchers have long been aware that it is higher order covariances that identify these models, particularly when persistence is high, it has never been clear what constitutes strong persistence and/or what panel length is needed to achieve identification. In this paper we use Monte Carlo simulations to illustrate the consequences of different degrees of persistence in data sets with sample sizes and panel lengths typically observed in empirical research.

We further show that identification may depend on the evolution of inequality itself. To see this, suppose that the permanent factor loadings are constant but the transitory factor loadings are falling over time. Considering the derivatives of the moment conditions associated with σ_α^2 and σ_{v1}^2, we now see that even with a relatively high ρ, the declining λs may induce an identification effect that mimics what would be observed with much lower persistence. We use Monte Carlo evidence to demonstrate the effect of alternative time trends on identification and consider the importance of these trends in the light of the available empirical evidence.

3 The data generation process

To generate the data for our Monte Carlo analysis, we calibrate the earnings model described in Section 2 using the parameter estimates from the existing studies outlined in Table 1 as a guide. We draw the initial value of v_1 from a normal distribution with mean zero and variance $\sigma_{v1}^2 = 0.3$. In each period the ε terms are drawn from i.i.d. normal distributions with mean zero and variance $\sigma_\varepsilon^2 = 0.2$. To allow for correlation between the slopes and intercepts of the earnings profiles, α and β are drawn from a bivariate normal distribution with means zero, variances $\sigma_\alpha^2 = 0.5$ and $\sigma_\beta^2 = 0.0004$ and covariance $\sigma_{\alpha\beta} = -0.01$. The MA parameter of the error process, θ, is set equal to -0.5, although our results are robust to variations in θ within the range observed in the empirical literature. The negative value for θ generates a fall in the first order correlation, which is consistent with that observed in real world data. Since the parameter ρ plays a key role in identifying these models, we consider various values of ρ between 0.3 and 0.95; these are in keeping with the range of estimates reported in the literature to date.[1]

As noted in the previous section, the evolution of inequality over time may also play an important role in identification of these models, so we experiment with

[1] Specification of an ARMA(1,1) model introduces a problem if $\rho \approx -\theta$. This 'common factors' problem is well known in time-series econometrics (e.g. [21] page 60). We avoid this complication by not considering values of ρ that are close to $-\theta$.

different patterns of inequality. We first consider a case where both permanent and transitory inequality are rising, by allowing p_t and λ_t to increase by 0.01 and 0.03 per year respectively. These increments are plausible given reported results. Alternative trends are considered later.

To generate heterogeneous growth profiles, we need to specify the ages of the individuals in our sample. In most of our analysis, we consider a balanced sample in which we observe individuals from the start of their working lives. This data structure is typical of cohort studies such as the NCDS in the UK and the NLSY in the US. Other panel data sets such as the BHPS (UK) or the PSID (US) observe people at different stages of their working lives in any given wave. As a result researchers often allow the earnings process to vary among cohorts. We discuss the consequences of introducing cohort heterogeneity and unbalanced data later in the paper.

Having generated the data, the models are estimated using *gmmcovearn*, a user-written Stata program for GMM estimation of earnings covariance models [15]. For each Monte Carlo experiment, we generate 1,000 samples and follow Kocherlakota [26] in starting the minimisation routine at the true parameter values in each sample. The sample sizes and panel lengths are varied across the experiments to examine their impact on identification.

4 Results

We begin by considering identification when the panel is long and the sample is large. In particular the top panel of Table 3 and Fig. 1a and b report the results for a panel length of 25 years and a sample size of 40,000 and $\rho = 0.8$. Data sets of this nature, typically administrative data, have been used by Dickens [12], Gustavsson [17] and Kalwij and Alessie [25]. The means and standard deviations of the estimates from the 1,000 simulations are given in Columns 2 and 3 respectively of the table. For each simulation, the standard GMM formula is used to calculate the asymptotic standard error for each parameter (see for example Cameron and Trivedi [5] page 174), and the average of these asymptotic standard errors is reported in Column 4. Columns 5 to 7 give additional summary details of the empirical distributions of the parameter estimates. Column 8 reports *p*-values for two tests of the hypothesis of normality of the distribution of the estimator, the sktest and the Kolmogorov-Smirnov test. Column 9 reports the empirical size of a 5% two-tailed test of the null hypothesis that the parameter equals the value specified in the DGP. Figure 1a provides a normal quantile plot of the parameter estimates to illustrate the validity of the normal approximation. Since the policy focus of papers in this literature is typically on the evolution of the permanent and transitory components of inequality, it is also important to compare the average predictions from the estimated model with the true profiles, calculated from the parameters of the DGP. This information is provided in Fig. 1b.

The results in Columns 1 and 2 of the top panel show that all the parameters are consistently and precisely estimated with this data structure when ρ is 0.8. As a result, for each of the permanent and transitory components, the estimated and true profiles given in Fig. 1b are very similar.

Table 3 Monte Carlo simulations for DGP with $N = 40,000$ and $T = 25$ (1000 replications)

Parameter	Monte Carlo average	Monte Carlo st.dev	Asymptotic s.error	10th percentile	Median	90th percentile	Normality a.Sktest b.Ksmirnov (p-values)	Emp. size (nominal 5%)
				$\rho = 0.8$				
ρ	.8000	.0035	.0036	.7954	.8001	.8043	a .0135 b .5290	.048
σ_α^2	.4998	.0099	.0095	.4870	.4993	.5129	a .0229 b .4820	.058
σ_ε^2	.1998	.0034	.0035	.1953	.1999	.2043	a .6234 b .9630	.056
σ_{v1}^2	.2999	.0089	.0088	.2885	.3006	.3108	a .0067 b .1920	.052
σ_β^2	.00040	.00001	.00002	.0004	.00040	.00043	a .0000 b .1820	.037
$\sigma_{\alpha\beta}$	−.0100	.0007	.0007	−.0109	−.0100	−.0090	a .0425 b .4030	.048
θ	−.5000	.0034	.0035	−.5042	−.5000	−.4957	a .1451 b .3990	.041
				$\rho = 0.9$				
ρ	.8998	.0032	.0033	.8957	.9000	.9037	a .2856 b .8550	.048
σ_α^2	.5148	.3051	.1144	.4248	.4842	.5789	a .0000 b .0000	.113
σ_ε^2	.1986	.0048	.0048	.1927	.1984	.2045	a .0001 b .6640	.087
σ_{v1}^2	.2860	.2904	.1115	.2226	.3172	.3735	a .0000 b .0000	.112
σ_β^2	.0004	.0003	.0001	.0003	.0004	.0005	a .0000 b .0000	.127
$\sigma_{\alpha\beta}$	−.0107	.0076	.0021	−.0116	−.0102	−.0089	a .0000 b .0000	.023
θ	−.50001	.0026	.0026	−.5033	−.4998	−.4968	a .3216 b .8410	.049
				$\rho = 0.95$				
ρ	.9469	.0045	.0042	.9409	.9472	.9528	a .0148 b .2060	.109
σ_α^2	.7695	1.1027	2.41	.0930	.1864	2.773	a .0000 b .0000	.677
σ_ε^2	.1969	.0078	.0047	.1918	.1967	.2026	a .0000 b .0000	.133
σ_{v1}^2	.0408	1.0585	2.315	−1.884	.5991	.6910	a .0000 b .0000	.674
σ_β^2	.00047	.00059	.0012	.00008	.00018	.0015	a .0000 b .0000	.569
$\sigma_{\alpha\beta}$	−.0157	.0230	.0515	−.0561	−.0035	−.0018	a .0000 b .0000	.676
θ	−.4991	.0024	.0023	−.5020	−.4990	−.4961	a .0606 b .8650	.072

True Parameter Values: $\sigma_\alpha^2 = 0.5$, $\sigma_\varepsilon^2 = 0.2$, $\sigma_{v1}^2 = 0.3$, $\sigma_\beta^2 = 0.0004$, $\sigma_{\alpha\beta}^2 = -0.01$, $\theta = -0.5$. p_t increasing by 0.01 and λ_t increasing by 0.03 in successive periods

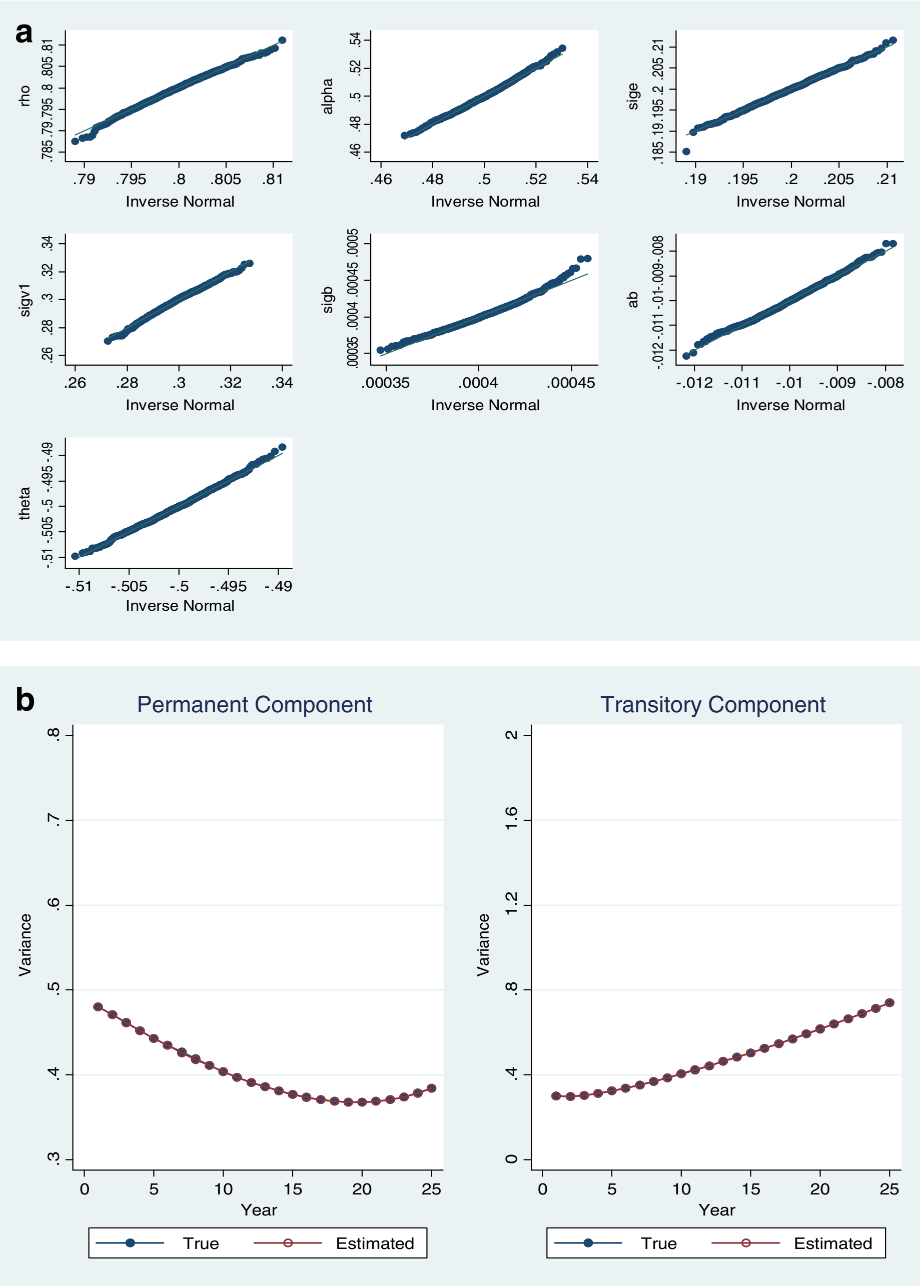

Fig. 1 **a** Normal quantile plots of GMM estimator with long panel, large sample and moderate persistence in transitory earnings. **b** True and estimated components of inequality with long panel, large sample and moderate persistence in transitory earnings

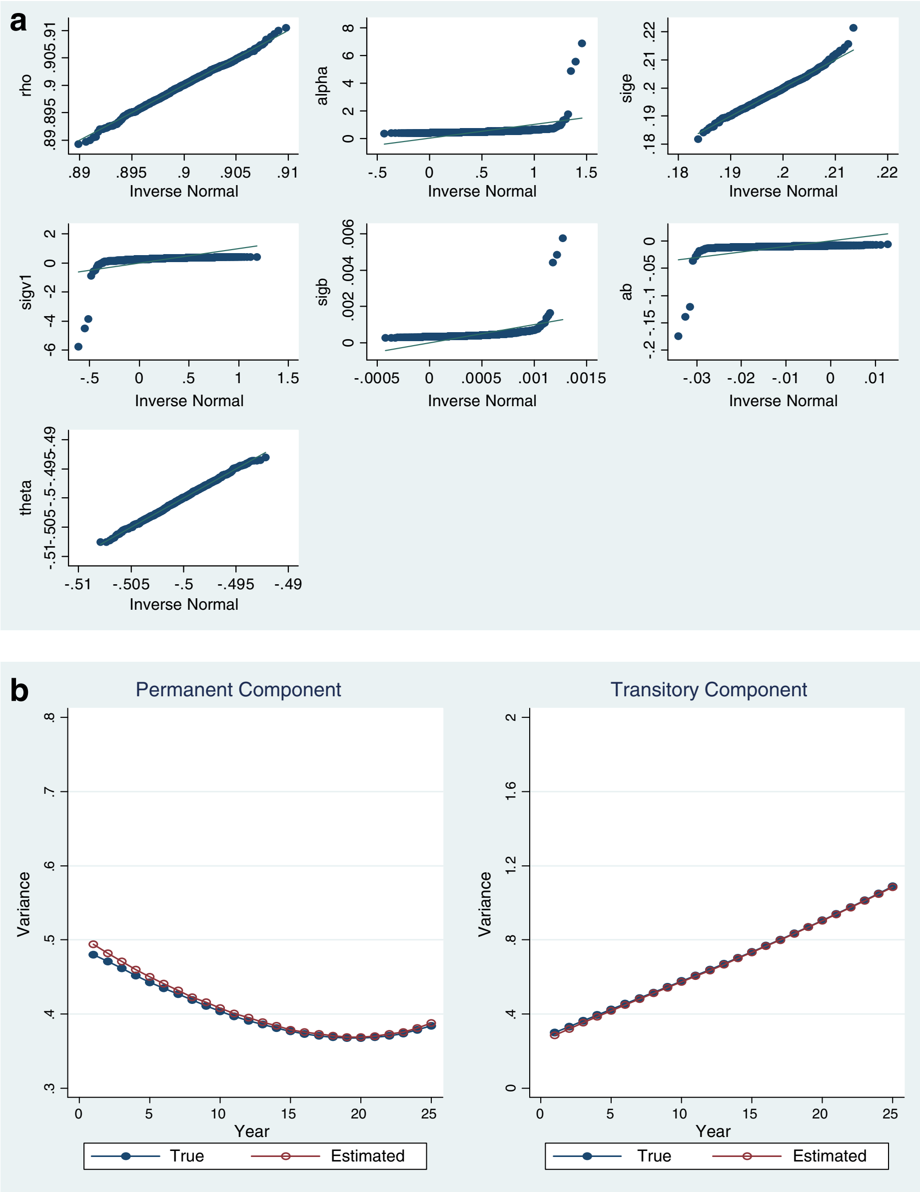

Fig. 2 **a** Normal quantile plots of GMM estimators with long panel, large sample and high persistence in transitory earnings. **b** True and estimated components of inequality with long panel, large sample and high persistence in transitory earnings

Consideration of the other results in the top panel of Table 3 shows that the asymptotic standard errors are very close to the Monte Carlo standard deviation. Furthermore, the distribution of the estimator is well approximated by a normal distribution. This is evident from both Column 8 of the table and Fig. 1a. Finally,

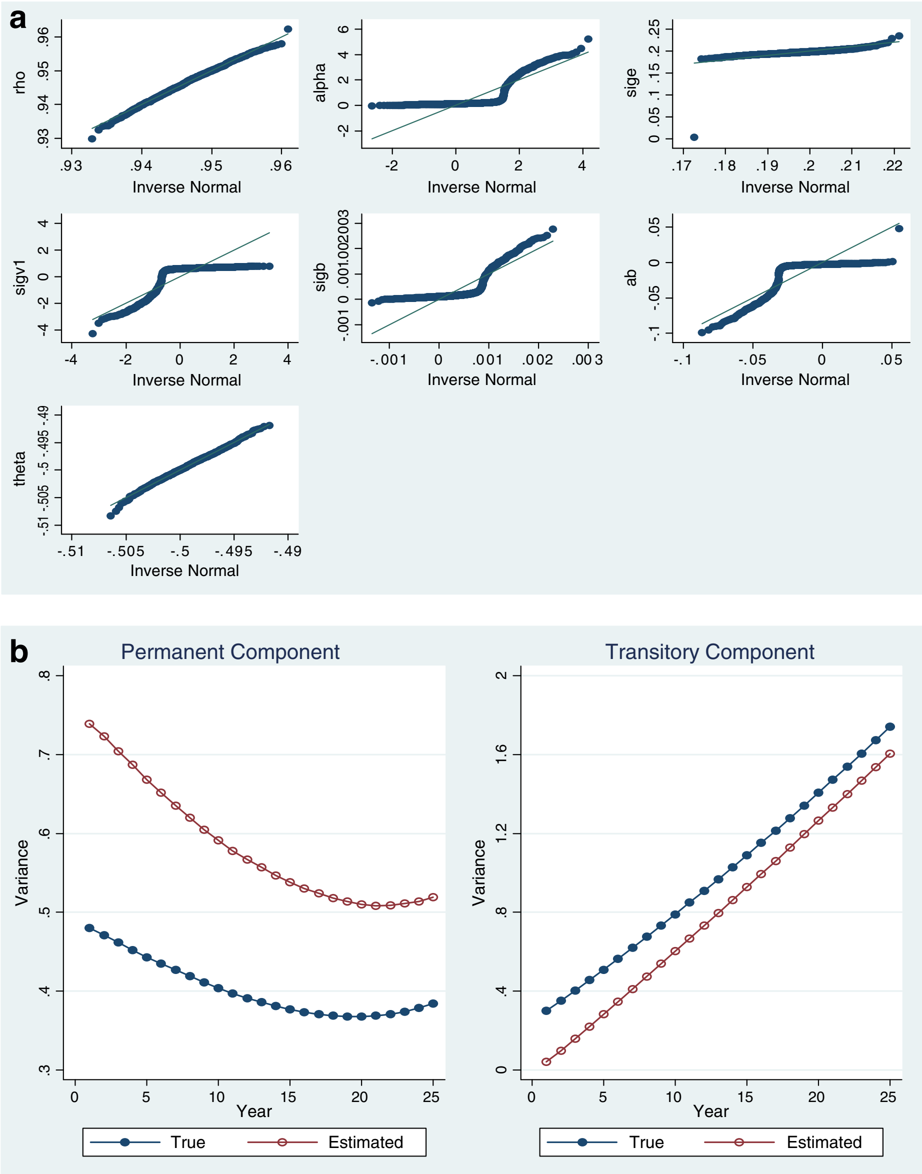

Fig. 3 **a** Normal quantile plots of GMM estimator with long panel, large sample and very high persistence in transitory earnings. **b** True and estimated components of inequality with long panel, large sample and very high persistence in transitory earnings

Column 9 shows that the empirical sizes of our hypothesis tests are close to their nominal values. These all support the validity of inferences based on the GMM estimator for these parameter values and this data structure.

To examine the sensitivity of identification to the degree of persistence in the transitory error term, we next consider models with values of ρ equal to 0.90 and

0.95.[2] When ρ is equal to 0.9 (middle panel, Table 3 and Fig. 2a and b), small biases in the parameters σ_α^2 and σ_{v1}^2 become evident; unreported estimates of the factor loadings also indicate small biases in the estimates of the p_t and λ_t. Consequently, small differences between the estimated and true components of inequality begin to emerge. This is shown in Fig. 2b. In addition the standard errors on σ_α^2 and σ_{v1}^2 are large and the asymptotic standard errors are substantially lower than the Monte Carlo standard deviations. Normality is rejected for most of the parameters, as evidenced by the tests in Table 3 and graphs in Fig. 2a. As a result the empirical sizes of hypothesis tests are distorted.

When ρ is equal to 0.95 (bottom panel, Table 3 and Fig. 3a and b), the results are considerably worse. The estimators for σ_α^2 and σ_{v1}^2 are biased, with average values of 0.77 and 0.04 respectively; the unreported estimates of the factor loadings also indicate biases in the estimates of p_t and λ_t. Consequently the estimated predictions of the permanent and transitory components are misleading; the permanent component is substantially overestimated and the transitory component is underestimated. This is shown in Fig. 3b.

In addition to these large biases, we also see that these parameters are imprecisely estimated. Indeed the lack of precision is such that over 25% of the estimates of σ_{v1}^2 are negative. Reports of negative variances are not unusual in this literature and are often interpreted as a sign that the underlying model is mis-specified or that the researcher has used poor starting values for the optimisation routine. However, our analysis, in which mis-specification cannot be an issue and which uses the true parameter values as starting values, shows that negative variances may be a symptom of weak identification and researchers should be aware of this possibility before considering adjusting their model.

Figure 3a shows that the biases in the asymptotic standard errors and the departures from normality are also more apparent when $\rho = 0.95$ than when $\rho = 0.9$. Indeed the bottom panel of Table 3 reveals empirical sizes as high as 68% for tests with a nominal size of 5%. The tendency to over reject when the model of earnings covariance is weakly identified is consistent with previous work on weak identification of the GMM estimator: in a consumption based asset pricing model, Kocherlakota [26] reports empirical sizes of up to 40% in his Monte Carlo analysis. Our results illustrate the well-known identification problem associated with high persistence in covariance models that was noted in Section 2 and show that even the very long panels and large sample sizes chosen in these experiments may be insufficient to identify the model when ρ is as high as 0.95.

Many panel data sets based on household survey data, such as the PSID, ECHP and BHPS, have considerably fewer than 40,000 observations. To examine identification with smaller samples, we choose a sample size of 3,000 but keep the panel length at 25. This combination is similar to that found in the PSID data used, for example, by Moffitt and Gottschalk [31] and Haider [19]. The results when ρ is equal to 0.8 are given in Table 4 and Fig. 4a and b. The results show that the parameter estimates are reasonable and as a result, the estimated and true profiles shown in Fig. 4b are similar. Therefore with relatively small samples, researchers are likely to

[2] We also considered models with ρ less than 0.8 and all parameters were consistently and precisely estimated.

Table 4 Monte Carlo simulations for DGP with $N = 3000$ and $T = 25$ (1000 Replications)

Parameter	Monte Carlo average	Monte Carlo st.dev	Asymptotic s.error	10th percentile	Median	90th percentile	Normality a.Sktest b.Ksmirnov (p-values)	Emp. size (nominal 5%)
				$\rho = 0.8$				
ρ	.8017	.0129	.0131	.7853	.8015	.8182	a .2855 b .8390	.048
σ^2_α	.5031	.0390	.0386	.4569	.5003	.5492	a .0000 b .0000	.043
σ^2_ε	.1995	.0127	.0131	.1832	.1997	.2159	a .3710 b .7410	.051
σ^2_{v1}	.2967	.0381	.0370	.2522	.2998	.3371	a .0000 b .0000	.060
σ^2_β	.00046	.0002	.00014	.00035	.00043	.00056	a .0000 b .0000	.018
$\sigma_{\alpha\beta}$	−.0098	.0028	.0029	−.0129	−.0102	−.0061	a .0000 b .0000	.081
θ	−.50	.0125	.0126	−.5177	−.5010	−.4850	a .2228 b .6030	.053

True parameter values: $\sigma^2_\alpha = 0.5$, $\sigma^2_\varepsilon = 0.2$, $\sigma^2_{v1} = 0.3$, $\sigma^2_\beta = 0.0004$, $\sigma^2_{\alpha\beta} = -0.01$, $\theta = -0.5$. p_t increasing by 0.01 and λ_t increasing by 0.03 in successive periods

get reliable predictions of the components of inequality, even when persistence is relatively strong, provided they have access to a long panel.[3]

The averages of the asymptotic standard errors are close to the Monte Carlo standard deviations. However the distributions of the estimators deviate from normality, as can be seen from the tests in Table 4 and the graphs in Fig. 4a. This leads to some distortion in the empirical sizes given in Column 9 and therefore care is needed when conducting hypothesis tests based on parameter estimates obtained using this data structure.

In addition to having moderate sample sizes, some of the data sets that have been used for analysis of the earnings covariance structure are relatively short. For example, the European Community Household Panel (ECHP), which is the only available panel data for some European countries, has at most 8 years of data. We run a Monte Carlo simulation for this type of data structure, by setting $T = 8$ and $N = 3{,}000$. The top panel of the Table 5, along with Fig. 5a and b consider the case where $\rho = 0.8$. The estimators of σ^2_α, σ^2_{v1} σ^2_β and $\sigma^2_{\alpha\beta}$ all exhibit large biases. The unreported factor loadings also show substantial biases. Consequently the estimated profiles differ from the true profiles, particularly in the early period; this is seen in Fig. 5b. The identification problems again give rise to a very large number of negative variances, with over 40% of the simulations reporting a negative variance for at least one of σ^2_α, σ^2_{v1} or σ^2_β.

As well as problems of inconsistency, the standard errors on the estimated parameters are very large, and are not well estimated by the asymptotic standard errors. With the exception of ρ, normality is overwhelmingly rejected by the formal

[3] Additional simulations indicate that a panel of at least 20 years is needed to obtain good point estimates and predictions.

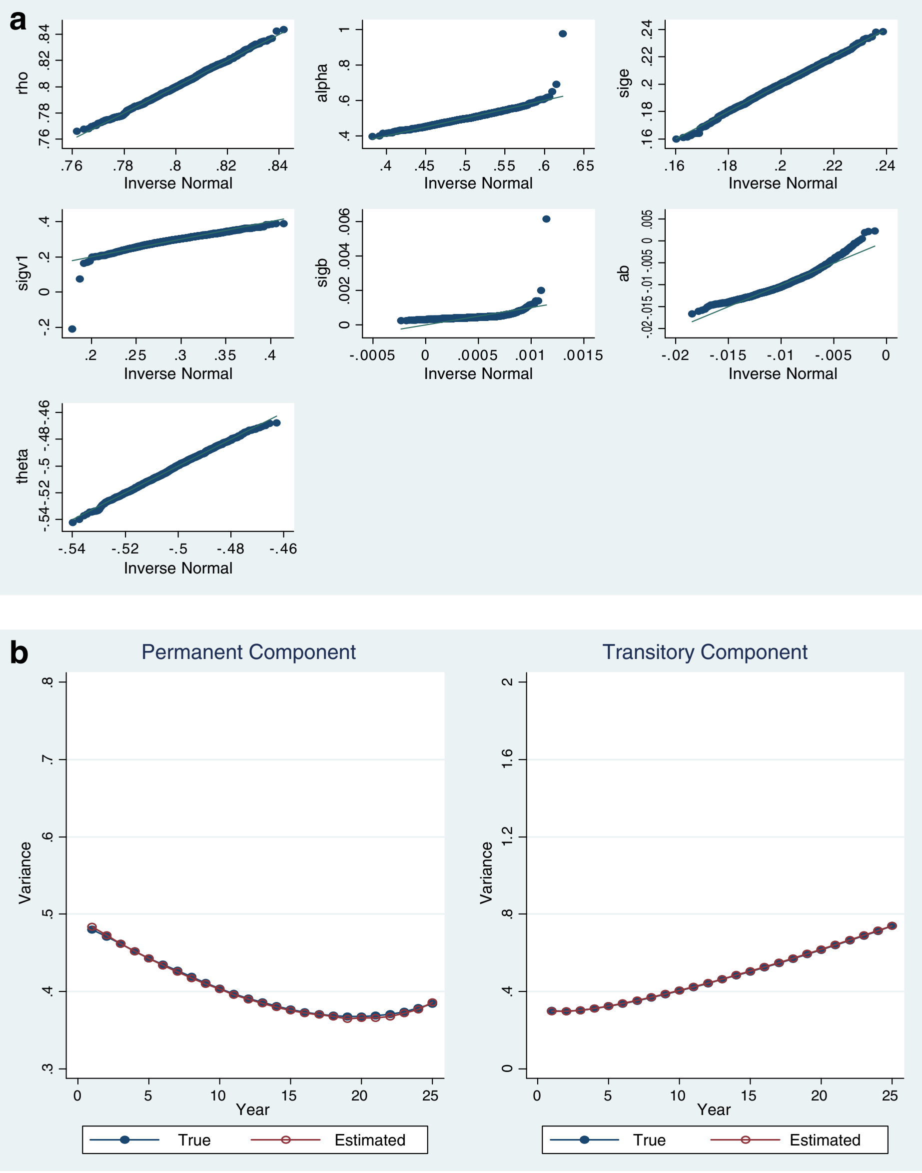

Fig. 4 **a** Normal quantile plots of GMM estimators with long panel, small sample and moderate persistence in transitory earnings. **b** True and estimated components of inequality with long panel, small sample and moderate persistence in transitory earnings

normality tests; inspection of the graphs in Fig. 5a confirm that the asymptotic approximation of the estimator is poor for all these parameters. As a result, hypothesis tests will be unreliable. This is confirmed in Column 9 of Table 5's top panel, where the empirical sizes are as large as 41% for a 5% test. Overall, the results for the estimator when applied to this data structure are unsatisfactory when the magnitude of ρ is of the order of 0.8.

Table 5 Monte Carlo simulations for DGP with $N = 3000$ and $T = 8$ (1000 replications)

Parameter	Monte Carlo average	Monte Carlo st.dev	Asymptotic s.error	10th percentile	Median	90th percentile	Normality a.Sktest b.Ksmirnov (p-values)	Emp. size (nominal 5%)
				$\rho = 0.8$				
ρ	.7835	.1112	.0897	.6352	.7850	.9170	a .6645	.277
							b .0000	
σ_α^2	.6804	1.231	2.9881	−.0324	.5922	.9071	a .0000	.418
							b .0000	
σ_ε^2	.2014	.0255	.0261	.1739	.1998	.2285	a .0000	.154
							b .0040	
σ_{v1}^2	.1432	1.0927	1.7199	−.1196	.2826	.7701	a .0000	.245
							b .0000	
σ_β^2	.0404	.3853	1.9041	−.0018	.0055	.0241	a .0000	.251
							b .0000	
$\sigma_{\alpha\beta}$	−.0419	.2854	.9798	−.0775	−.0190	.0471	a .0000	.307
							b .0000	
θ	−.4832	.0574	.0514	−.5449	−.4946	−.4007	a .0000	.160
							b .0000	
				$\rho = 0.3$				
ρ	.2950	.0926	.0999	.1735	.2931	.4217	a .0044	.040
							b .4710	
σ_α^2	1.5376	4.6871	11.30	.1528	.5420	3.0503	a .0000	.116
							b .0000	
σ_ε^2	.1882	.0435	.0624	.1509	.1939	.2186	a .0000	.026
							b .0000	
σ_{v1}^2	.2795	.1032	.1157	.18230	.2978	.3183	a .0000	.046
							b .0000	
σ_β^2	1.8180	6.1038	15.69	.0011	.0091	5.7842	a .0000	.070
							b .0000	
$\sigma_{\alpha\beta}$	−1.4274	5.3642	12.97	−4.1136	−.0441	.0932	a .0000	.129
							b .0000	
θ	−.5038	.1282	11.90	−0.6825	−.4923	−.3443	a .0021	.052
							b .0001	

True parameter values: $\sigma_\alpha^2 = 0.5, \sigma_\varepsilon^2 = 0.2, \sigma_{v1}^2 = 0.3, \sigma_\beta^2 = 0.0004, \sigma_{\alpha\beta}^2 = -0.01, \theta = -0.5$. p_t increasing by 0.01 and λ_t increasing by 0.03 in successive periods

For $\rho = 0.8$, 10 out of the 1000 simulations failed to converge after reaching the Stata limit of 16000 iterations. For $\rho = 0.3$, 117 out of the 1000 simulations failed to converge after reaching the Stata limit

A number of studies using data sets of this size report estimates of ρ that are much smaller than 0.8. For instance, Cervini-Pla and Ramos [9] report an estimated ρ of 0.27, while Ramos [34] reports a range of estimates from 0.3 to 0.4. In the bottom panel of Table 5 and Fig 6a and b we consider estimation of the model with eight years of data and $N = 3{,}000$ and $\rho = 0.3$. These results are surprisingly poor given the low persistence of the transitory shock. The estimators of σ_α^2, σ_β^2 and $\sigma_{\alpha\beta}$ are biased and imprecisely estimated. Furthermore, the asymptotic approximations are poor.[4]

[4] Although the results improve with larger samples, even with a sample size of 250,000, the average of the sample estimates differ from the true values.

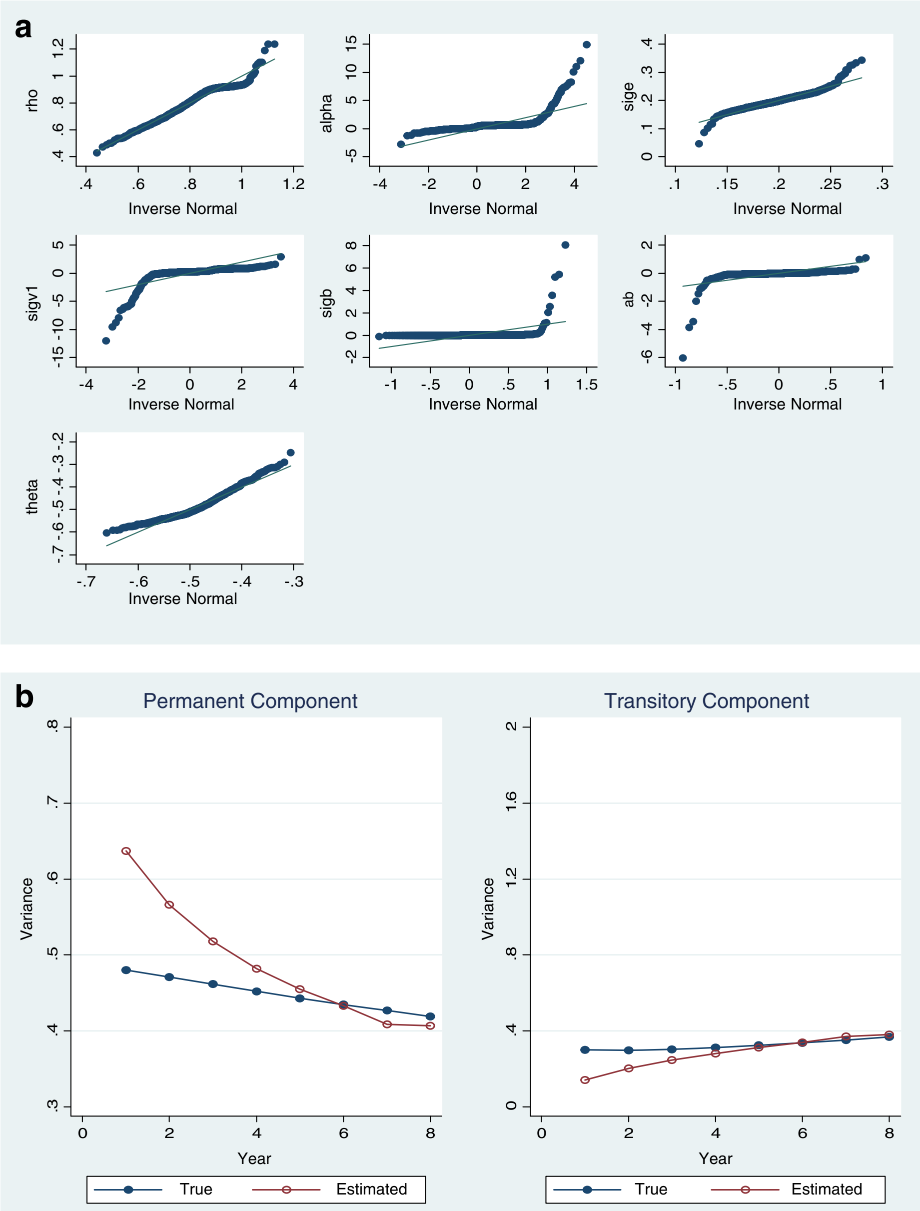

Fig. 5 **a** Normal quantile plots of GMM estimators with short panel, small sample and moderate persistence in transitory earnings. **b** True and estimated components of inequality with short panel, small sample and moderate persistence in transitory earnings

The estimator also had difficulty estimating the factor loadings, with both the λ_ts and the p_ts exhibiting large biases. However, when we consider the predicted profiles associated with these estimates, shown in Fig. 6b, we see that these are surprisingly good. On closer inspection, the identification problem in this case appears to differ

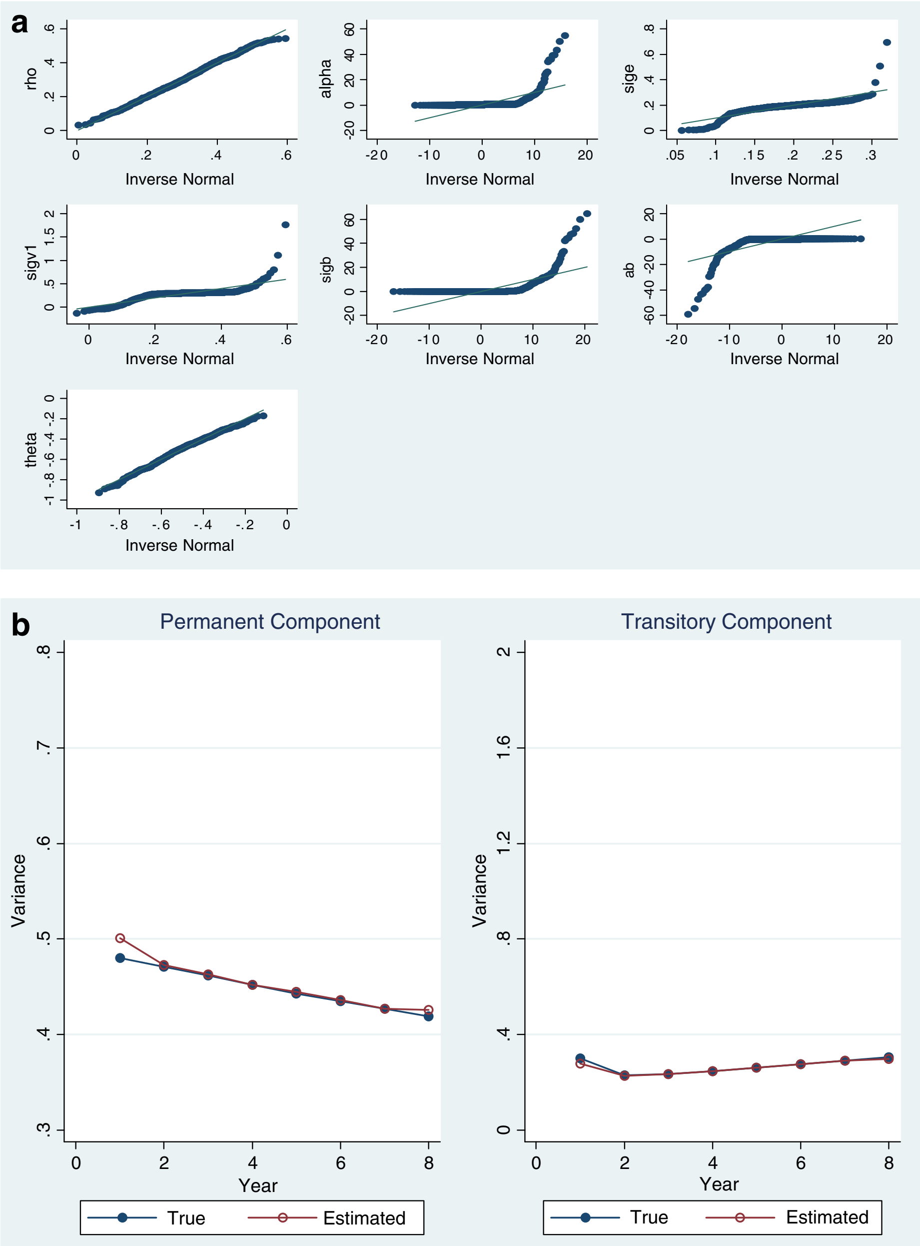

Fig. 6 **a** Normal quantile plots of GMM estimators with short panel, small sample and low persistence in transitory earnings. **b** True and estimated components of inequality with short panel, small sample and low persistence in transitory earnings

from that associated with higher persistence. Unlike the previous case, the estimates of σ_{v1}^2 and σ_ε^2 are now reasonable. The problem now is not in distinguishing the permanent from the transitory component, but rather in distinguishing among the components of permanent inequality, namely σ_α^2, σ_β^2 and $\sigma_{\alpha\beta}$.

The source of this identification problem is not immediately obvious from inspection of the derivatives of the moment conditions. The problem in part reflects the very low value of σ_β^2 typically observed in the literature. When we run additional simulations with the same data structure and level of persistence but with a substantially higher value for σ_β^2, the individual parameters of the permanent component are consistently estimated. However the value needed is of the order of 0.04 rather than the value of 0.0004 typically observed.[5] The problem also reflects the limited extent to which age variation is used in the estimation of the above model; X_t changes very little over such a short panel. When we extend the panel length to 25 years but maintain $\rho = 0.3$, the problem is no longer evident. This is because the moment conditions for the longer panel now span a sufficiently wide age range to allow identification of the heterogeneous growth pattern. We return to this point in Section 5.2.

5 Extensions

As noted in Section 2, the evolution of inequality can play a role in identifying earnings covariance models. We consider this issue in Section 5.1. Sections 5.2 and 5.3 consider the impact of cohort heterogeneity and unbalanced data on identification. Finally Section 5.4 examines the sensitivity of our findings to changes in the specification of the earnings process.

5.1 Time trends

As illustrated in Sections 2 and 4, high persistence in the transitory error process can make identification of the permanent and transitory components difficult. This problem is well known in the literature and we showed in Table 3 that with a reasonable parameterization, the model is poorly identified when ρ is equal to 0.95, even with long panels and large sample sizes. However what seems less well known is the role of time trends in determining identification. In Section 2, we noted that certain patterns of inequality may accommodate identification. To illustrate this, we now consider a DGP which is identical to that used in the bottom panel of Table 3 ($\rho = 0.95$), except that λ_t *falls* by 0.035 in each successive period, This is similar to the trend reported by Sologon and O'Donoghue [37].

The results are given in Table 6 and Fig. 7a and b. For ease of comparison we also reproduce the results for the original time trends in the bottom panel of the table. In contrast to the earlier results for $T = 25$ and $N = 40{,}000$, we find that with the alternative time trends, the model is well identified by the GMM estimator, even with ρ as high as 0.95.[6] All of the parameters are now precisely and consistently estimated. The analytical standard errors are much closer to the empirical standard deviations, while the departure from the normal approximation is less pronounced. Consequently, the empirical sizes reported in the last column are much closer to the nominal sizes. To the extent that identification has been considered in this literature,

[5] See Baker [2] for a related discussion.

[6] We were unable to find any set of factor loadings that eliminated the identification problems for a data structure with $T = 8$ and $N = 3{,}000$.

Table 6 Monte Carlo simulations for DGP with $N = 40{,}000$ and $T = 25$ (1000 replications)

Parameter	Monte Carlo average	Monte Carlo st.dev	Asymptotic s.error	10th percentile	Median	90th percentile	Normality a.Sktest b.Ksmirnov (p-values)	Emp. size (nominal 5%)
				$\rho = 0.95$, λt falling by 0.035				
ρ	.9500	.0030	.0031	.9464	.9500	.9537	a .0358 b .7100	.038
σ^2_α	.5036	.0252	.0215	.4798	.5004	.5279	a .0000 b .0000	.037
σ^2_ε	.1999	.0025	.0024	.1966	.1998	.2033	a .1348 b .3820	.061
σ^2_{v1}	.2970	.0245	.0206	.2737	.2994	.3203	a .0000 b .0000	.039
σ^2_β	.0004	.00004	.00004	.00037	.0004	.0005	a .0000 b .0000	.053
$\sigma_{\alpha\beta}$	−.0100	.0004	.0003	−.0104	−.0100	−.0097	a .0000 b .0000	.037
θ	−.4999	.0023	.0023	−.5027	−.4999	−.4971	a .4309 b .8410	.053
				$\rho = 0.95$, λt rising by 0.03				
ρ	.9469	.0045	.0042	.9409	.9472	.9528	a .0148 b .2060	.109
σ^2_α	.7695	1.1027	2.41	.0930	.1864	2.773	a .0000 b .0000	.677
σ^2_ε	.1969	.0078	.0047	.1918	.1967	.2026	a .0000 b .0000	.133
σ^2_{v1}	.0408	1.0585	2.315	−1.884	−2.020	−1.716	a .0000 b .0000	.674
σ^2_β	.00047	.00059	.0012	.00007	.00002	.0015	a .0000 b .0000	.569
$\sigma_{\alpha\beta}$	.0156	.0230	.0515	−.0562	−.0035	−.0018	a .0000 b .0000	.676
θ	−.4991	.0024	.0023	−.5020	−.4990	−.4961	a .0606 b .8650	.072

True parameter values: $\sigma^2_\alpha = 0.5, \sigma^2_\varepsilon = 0.2, \sigma^2_{v1} = 0.3, \sigma^2_\beta = 0.0004, \sigma^2_{\alpha\beta} = -0.01, \theta = -0.5$. p_t increasing by 0.01 in successive periods

the focus has been on the degree of transitory earning persistence. The results in this section show that this reasoning is incomplete. Identification in these models depends on the interaction of the degree of persistence with the time pattern of inequality in the economy being studied.

5.2 Cohort effects

While panel data sets such as the NCDS and the NLSY follow a single cohort over time, other data sets such as the BHPS and the PSID include individuals of different ages. Recognising the possibility that earnings dynamics may differ across cohorts, many studies have extended the model described in Section 2 to allow for cohort heterogeneity. This is modelled by including cohort shifters that allow the permanent and transitory components to vary across cohorts; additionally, some researchers

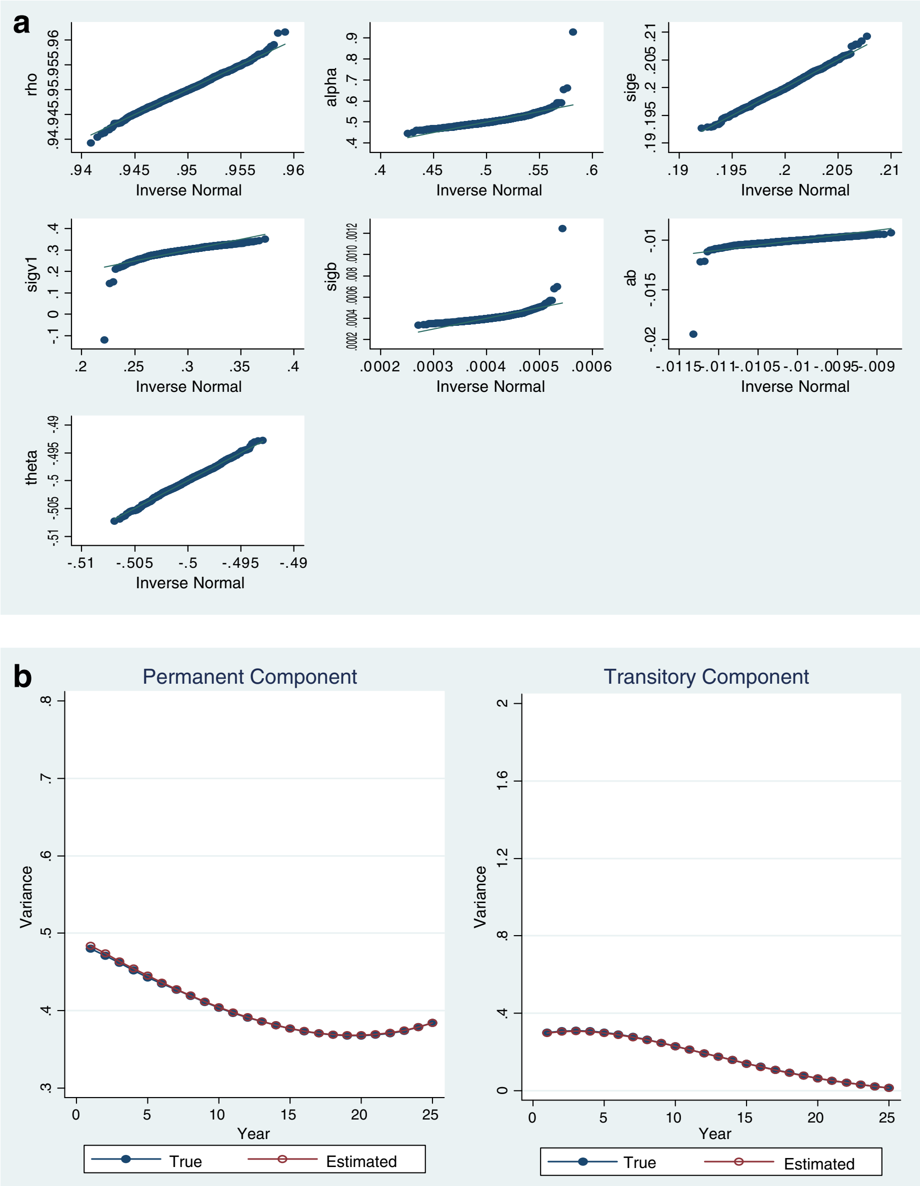

Fig. 7 **a** Normal quantile plots of GMM estimators with long panel, large sample, very high persistence in transitory earnings and diverging time trends. **b** True and estimated components of inequality with long panel, large sample, very high persistence in transitory earnings and diverging time trends

have allowed the initial transitory variances to vary by cohort. Specifically, earnings for individual i, belonging to cohort c, with x years of experience at time t, y_{icxt} is given by

$$y_{icxt} = q_c p_t \alpha_{ix} + s_c \lambda_t v_{ict} \tag{5}$$

where q_c and s_c are cohort shifters on the permanent and transitory components respectively and σ^2_{v1c} varies by cohort. The remainder of the model is as described in Section 2. The resulting model is identical to that used by Sologon and O'Donoghue [36]. This introduction of cohort effects increases the number of parameters by $3(c-1)$, while increasing the number of moments by a factor of c.

To examine identification of the earnings covariance model in the presence of cohort-specific effects, we consider a DGP that is identical to that used in Table 3, except that we now introduce two cohorts.[7] The total sample size, N, is 40,000, with 20,000 in each cohort, each observed for 25 years. Individuals in the young cohort are observed from the beginning of their working lives, while observations on the second, older cohort begin 10 years after this cohort enter the labour market. We normalise $q_1 = s_1 = 1$ and calibrate $q_2 = 1.3$ and $s_2 = 1.1$; these are within the range of the estimates reported in the literature. We set $\sigma^2_{v11} = 0.3$ and $\sigma^2_{v12} = 0.6$. With this parameterisation, we repeat the analysis conducted in Table 3 for $\rho = 0.8$. The results given in the top panel of Table 7 show that all the parameters of the model, including the cohort shifters and initial variances, are well identified. Unreported simulations for $\rho = 0.95$ suggest that the identification problems shown in Table 3 are still evident in the model with cohort effects.

In Section 4 we noted that the limited age variation in short panels contributes to weak identification of the individual components of permanent inequality even when ρ is low. We now show how a model that distinguishes between different cohorts can overcome this problem. The second panel of Table 7 reports simulation results for the model given in Eq. 5, with $\rho = 0.3$, $T = 8$ and $N = 3000$. In contrast to the results in Table 5, the individual components of permanent inequality are now consistently estimated, despite the short panel. This is because the moment conditions for this model are now specified separately for each cohort. Consequently the age variation in the sample is exploited more effectively than in models that combine all cohorts. We emphasise that it is the appropriate exploitation of the age variation across cohorts and not the presence or absence of cohort effects per se that is important in this case. Additional simulations not reported, show that the components of permanent inequality are still identified even if all cohort shifters and cohort variances are in fact equal. As a result researchers working with short panels should consider models that distinguish between cohorts even when structural differences across cohorts are limited.

5.3 Unbalanced samples

The DGPs we have examined so far have focused on balanced data, in which each individual is present in every wave of the data and therefore contributes to every moment condition. In the panel data sets used in practice, individuals are typically missing in some waves. In these situations, researchers often work with unbalanced data. Clearly, if movements in and out of the sample are random, the unbalanced nature of the data may affect the precision of the estimator, through smaller sample sizes, but not its consistency. However, care should be taken to ensure that the sample sizes for estimation of long covariances are adequate.

[7] Experiments with larger number of cohorts show that our key results are not affected by the number of cohorts.

Table 7 Monte Carlo simulations for DGP with two cohorts (1000 replications)

Parameter	Monte Carlo average	Monte Carlo st.dev	Asymptotic s.error	10th percentile	Median	90th percentile	Normality a.Sktest b.Ksmirnov (p-values)	Emp. size (nominal 5%)
				$\rho = 0.8$, $T = 25$ and $N = 40{,}000$				
ρ	.8003	.0043	.0043	.7946	.8003	.8056	a .694 b .933	.058
σ^2_α	.5004	.0141	.0141	.4816	.5003	.5191	a .846 b .882	.050
σ^2_ε	.2001	.0042	.0042	.1947	.2001	.2055	a .892 b .977	.054
σ^2_{v11}	.2996	.0142	.0144	.2811	.2992	.3178	a .642 b .863	.049
σ^2_{v12}	.6001	.0108	.0111	.5866	.6000	.6139	a .129 b .360	.040
σ^2_β	.0004	.00002	.00002	.00037	.0004	.00041	a .392 b .338	.047
$\sigma_{\alpha\beta}$	−.0100	.0004	.00038	−.01049	−.0100	−.0095	a .433 b .822	.042
θ	−.5001	.0042	.0043	−.5054	−.5002	−.4932	a .662 b .866	.048
q_2	1.2997	.0185	.0184	1.2776	1.2998	1.3239	a .240 b .524	.050
s_2	1.1000	.0033	.0033	1.0959	1.0999	1.1044	a .054 b .599	.048
				$\rho = 0.3$, $T = 8$ and $N = 3000$				
ρ	.2983	.0513	.0506	.2358	.2966	.3655	a .0389 b .545	.059
σ^2_α	.4976	.0206	.0236	.4705	.4977	.5236	a .5265 b .834	.033
σ^2_ε	.1970	.0188	.0186	.1732	.1984	.2193	a .0000 b .212	.048
σ^2_{v11}	.2985	.0161	.0158	.2785	.2997	.3173	a .0000 b .017	.032
σ^2_{v12}	.5986	.0237	.0238	.5680	.5983	.6280	a .7016 b .790	.056
σ^2_β	.0005	.0003	.0003	.0003	.0005	.0007	a .0000 b .000	.063
$\sigma_{\alpha\beta}$	−.0083	.0056	.0068	−.0142	−.0098	.0001	a .0000 b .0000	.168
θ	−.5024	.0696	.0670	−.5905	−.4972	−.4197	a .0000 b .003	.038
q_2	1.2593	.1852	.2099	.9895	1.2900	1.4818	a .0000 b .000	.144
s_2	1.1000	.0094	.0091	1.0877	1.1002	1.1119	a .2897 b .857	.062

True parameter values: $\sigma^2_\alpha = 0.5$, $\sigma^2_\varepsilon = 0.2$, $\sigma^2_{v11} = 0.3$,$\sigma^2_{v12} = 0.6$, $\sigma^2_\beta = 0.0004$, $\sigma^2_{\alpha\beta} = -0.01$, $\theta = -0.5$, $q_2 = 1.3$, $s_2 = 1.1$. p_t increasing by 0.01 and λ_t increasing by 0.03 in successive periods

However, a common feature of a number of data sets used in this literature is that the probability of being present in a particular wave of a survey is related to an individual's age. This may arise if younger people are absent from the labour force in the early years of the panel while in education, and if older people leave in later years due to retirement. GMM still provides a consistent estimator of the models considered in this paper in this setting because plim $\left(\frac{1}{N}\sum_{i=1}^{N} m_i(y_i;\varphi_0)\right) = 0$ will continue to hold with unbalanced data whenever the model is identified using the balanced data.

To illustrate this, we adjust the DGP discussed in Section 3. We allow for age heterogeneity by considering a situation in which in any year we observe variation in the ages of individuals in our sample. To keep things simple, we generate a 25 year sample of 40,000 individuals, of which half have t years of experience at time t, and the other half have $(t+5)$ years of experience. To generate unbalanced data, we consider a missing mechanism where in the first year, only 50% of the younger sample is observed, with a further 10% entering in each successive year, so that 100% are observed by year 6. For older people, we assume that 10% exit in year 21, with an additional 10% leaving in each subsequent year so that by year 25, 50% of the older sample are missing. This results in a data structure where observations are missing randomly conditional on age. Results for Monte Carlo simulations for this DGP with $\rho = 0.8$ are shown in Table 8. As we can see, all the parameters are consistently and precisely estimated, despite the unbalanced nature of the data.

Clearly if the probability of being missing from the sample is related to unobservable characteristics that are correlated with earnings, the GMM consistency condition is likely to be violated. Consistent estimation of the earnings covariance structure in this case will require specification of a missing mechanism for the data and joint estimation of this mechanism with the earnings process. We know of no study that has carried out such an exercise in this context.

Table 8 Monte Carlo simulations for DGP with $N = 40000$ and $T = 25$ and unbalanced data (1000 replications)

Parameter	Monte Carlo average	Monte Carlo st.dev	Asymptotic s.error	10th percentile	Median	90th percentile	Normality a.Sktest b.Ksmirnov (p-values)	Emp. size (nominal 5%) two tail.
				$\rho = 0.8$				
ρ	.8003	.00466	.0043	.7945	.8001	.8064	a .3899 b .3000	.070
σ^2_α	.5002	.01299	.0127902	.4836	.5000	.5170	a .9521 b .7050	.059
σ^2_ε	.1999	.00416	.0044842	.1944	.1999	.2051	a. 2093 b .3690	.030
σ^2_{v1}	.2997	.00804	.0085531	.2891	.2996	.3099	a .2589 b .6310	.035
σ^2_β	.0004	.00003	.0000275	.00037	.0004	.00044	a .0000 b .3330	.027
$\sigma_{\alpha\beta}$	−.0100	.00096	.0009495	−.0111	−.0100	−.0087	a .0001 b .2520	.054
θ	−.5002	.0043	.0039467	−.5059	−.5002	−.4948	a .6252 b .9280	.076

True parameter values: $\sigma^2_\alpha = 0.5, \sigma^2_\varepsilon = 0.2, \sigma^2_{v1} = 0.3, \sigma^2_\beta = 0.0004, \sigma^2_{\alpha\beta} = -0.01, \theta = -0.5$. p_t increasing by 0.01 and λ_t increasing by 0.03 in successive periods

In this case, asymptotic standard errors are calculated using the adjustment suggested by Haider [19]

5.4 Combined random growth and random walk model

A small number of papers have estimated earnings covariance models that combine the Random Growth and Random Walk models (e.g. [31, 34]). Earnings for individual i, with x years of experience at time t, y_{ixt} is then given by

$$y_{ixt} = p_t\alpha_{ix} + \lambda_t v_{it} \tag{6a}$$

$$\alpha_{ix} = \alpha_{i(x-1)} + \beta_i + w_{ix} \tag{6b}$$

where $E(\alpha_{ix}) = E(\beta_i) = E(w_{ix}) = E(v_{it}) = 0$. α_{i0} and β_i have variances σ_α^2 and σ_β^2 respectively and covariance $\sigma_{\alpha\beta}$. As before, the first two terms of (6b) capture the random growth component of earnings. w_{ix} is a shock that arrives randomly from a distribution with variance σ_w^2, but permanently alters an individual's position in the earnings distribution. The transitory component is modelled using an ARMA process as before. The estimation of the combined Random Growth and Random Walk (RG+ RW) model introduces one extra parameter, σ_w^2, but no new moment conditions.

To examine the implications of estimating the RG+RW model, we consider a DGP that is identical to that used in the model analysed in Table 3, but extend the specification of the permanent component to incorporate a random walk as well as a random growth process. We keep $T = 25$ and $N = 40{,}000$ and calibrate $\sigma_w^2 = 0.005$, which is consistent with estimates reported in the literature. We carry out our simulations for $\rho = 0.8$ and present the results in Table 9. We see that the estimator performs reasonably well, but there is evidence of a small bias in the point estimates for σ_α^2 and σ_{v1}^2. The estimator for σ_w^2 is also biased, and the standard error

Table 9 Monte Carlo simulations for DGP with $N = 40000$ and $T = 25$ and combined random growth and random walk model (1000 replications)

Parameter	Monte Carlo average	Monte Carlo st.dev	Asymptotic s.error	10th percentile	Median	90th percentile	Normality a.Sktest b.Ksmirnov (p-values)	Emp. size (nominal 5%) two tail.
				$\rho = 0.8$				
ρ	.8042	.0124	.0111	.7885	.8042	.8197	a .0009 b .4920	.110
σ_α^2	.4893	.0335	.0281	.4481	.4929	.5274	a .0000 b .0040	.077
σ_ε^2	.2019	.0060	.0060	.1943	.2020	.2095	a . 2365 b .7720	.075
σ_{v1}^2	.3112	.0351	.0293	.2705	.3085	.3561	a .0000 b .0010	.085
σ_β^2	.0004	.00009	.00008	.0003	.0004	.0005	a .0002 b .1330	.126
$\sigma_{\alpha\beta}$	−.0096	.0012	.0011	−.0109	−.0096	−.0081	a .0004 b .6110	.097
θ	−.4990	.0047	.0044	−.5046	−.4994	−.4930	a .0006 b .0940	.063
σ_ω^2	.0036	.0043	.0037	−.0020	.0035	.0090	a .1271 b .6250	.105

True parameter values: $\sigma_\alpha^2 = 0.5$, $\sigma_\varepsilon^2 = 0.2$, $\sigma_{v1}^2 = 0.3$, $\sigma_\beta^2 = 0.0004$, $\sigma_{\alpha\beta}^2 = -0.01$, $\sigma_\omega^2 = 0.005$, $\theta = -0.5$. p_t increasing by 0.01 and λ_t increasing by 0.03 in successive periods on this parameter is very large. The inability to estimate this parameter precisely may lead to difficulties in determining the correct model when ρ is high. However, further simulations suggest that the estimator performs well for this RG+RW model provided ρ is less than 0.7. Therefore, although the presence of the random walk in the DGP makes identification more difficult, results are still satisfactory for a wide range of values of ρ.

6 Conclusion

In this paper we examine the performance of the GMM estimator in the context of the covariance structure of earnings. We consider a range of models that are representative of those previously estimated and which capture many of the key features of earnings dynamics discussed in the literature. In particular the models considered allow for heterogenous growth in permanent earnings profiles, persistence in the transitory shocks, time trends in the earnings process and earnings dynamics that differ across age cohorts. We use Monte Carlo simulations to examine the sensitivity of parameter identification to key features such as data structure, the degree of persistence of earnings shocks and model specification. We show that long panels allow the identification of the model, even when persistence in transitory shocks is high. Short panels, on the other hand, are insufficient to identify individual parameters of the model even with moderate levels of persistence.

In terms of practical implications, since the estimate of ρ is reliable in each of the models we consider, this parameter can form the basis of a check on identification. However, it is not possible to establish a single cut-off value below which identification is guaranteed. Researchers with access to long panels should be wary if ρ is above 0.9 and the estimated standard errors on σ_α^2 and σ_{v1}^2 are large. When using shorter panels, problems are more severe. Identification is problematic in these models even for moderate values of ρ and researchers need to be very careful when using the GMM estimator with short panels irrespective of the degree of earnings persistence.

Acknowledgements We would like to thank Peter Gottschalk for helpful comments on an earlier version of this paper, as well as seminar participants at NUI Maynooth and the Geary Institute UCD and participants at the *Income and Earnings Dynamics* conference at NUI Maynooth, the BHPS-2009 conference in Essex, The European Society of Population Economics Annual conference 2010, Essen, and the Irish Economics Association Annual Conference 2010, Belfast. We would also like to thank Steven Haider, Marco Lilla, Xavier Ramos, Denisa Sologon and Robert Valetta, who provided us with additional details of their work that were helpful when conducting our research. We gratefully acknowledge the financial support provided by the Irish Research Council for the Humanities and Social Sciences.

References

1. Altonji, J., Segal, L.: Small-sample bias in GMM estimation of covariance structures. J. Bus. Econ. Stat. **14**(3), 353–366 (1996)
2. Baker, M.: Growth rate heterogeneity and covariance structure of life-cycle earnings. J. Labor Econ. **15**(2), 338–375 (1997)
3. Baker, M., Solon, G.: Earnings dynamics and inequality among Canadian Men, 1976–1992: evidence from longitudinal income tax records. J. Labor Econ. **21**(2), 289–321 (2003)

4. Blundell, R., Bond, S.: GMM estimation with persistent panel data: an application to production functions. Econom. Rev. **19**(4), 321–340 (2000)
5. Cameron, A.C., Trivedi, P.K.: Microeconometrics: Methods and Applications. Cambridge University Press, New York (2005)
6. Cappellari, L.: The covariance structure of Italian male wages. Manch. Sch. **68**(6), 659–684 (2000)
7. Cappellari, L.: Earnings dynamics and uncertainty in Italy: how do they differ between the private and public sectors? Labour Econ. **9**, 477–496 (2002)
8. Cappellari, L.: The dynamics and inequality of Italian Men's earnings: long-term changes or transitory fluctuations? J. Hum. Resour. **39**(2), 475–499 (2004)
9. Cervini-Pla, M., Ramos, X.: Long-term earnings inequality, earnings instability and temporary employment in Spain 1993–2000. IZA Working Paper 3538 (2008)
10. Clark, T.: Small-sample properties of estimators of nonlinear models of covariance structure. J. Bus. Econ. Stat. **14**(4), 367–373 (1996)
11. Daly, M., Valletta, R.: Cross-national trends in earnings inequality and instability. Econ. Lett **99**(2), 215–219 (2008)
12. Dickens, R.: The evolution of individual male earnings in Great Britain: 1975–95. Econ. J. **110**(460), 27–49 (2000)
13. Doris, A., O'Neill, D., Sweetman, O.: Does growth affect the nature of inequality? Ireland 1994–2001. IZA Discussion paper 3701 (2008)
14. Doris, A., O'Neill, D., Sweetman, O.: Aggregate earnings inequality in Europe: permanent differences or transitory fluctuations? NUI Maynooth Working Paper N211-10 (2010)
15. Doris, A., O'Neill, D., Sweetman, O.: GMM estimation of the covariance structure of longitudinal data on earnings. Stata J. **11**(4), 439–459 (2011)
16. Gustavsson, M.: The 1990s rise in Swedish earnings inequality – persistent or transitory? Appl. Econ. **39**, 25–30 (2007)
17. Gustavsson, M.: A new picture of Swedish earnings inequality: persistent and transitory components, 1960–1990. Rev. Income Wealth **54**(4), 324–329 (2008)
18. Guvenen, F.: An empirical investigation of labor income processes. Rev. Econ. Dyn. **12**, 58–79 (2009)
19. Haider, S.: Earnings instability and earnings inequality of males in the United States: 1967–1991. J. Labor Econ. **19**(5), 799–836 (2001)
20. Hall, A.: Generalised Method of Moments. Oxford University Press, Oxford (2005)
21. Hamilton, J.: Time Series Analysis. Princeton University Press, New Jersey (1994)
22. Hansen, L.P.: Large sample properties of generalised method of moments estimators. Econometrica **50**, 1029–1054 (1982)
23. Hansen, L., Heaton, P., Yaron, A.: Finite sample properties of some alternative GMM estimators. J. Bus. Econ. Stat. **14**, 262–280 (1996)
24. Hoffmann, F.: The Different Life-Cycle Income Trajectories of Less and More Educated Workers, mimeo. University of Toronto (2009)
25. Kalwij, A., Alessie, R.: Permanent and transitory wages of British men, 1975–2001: year, age and cohort effects. J. Appl. Econ. **22**, 1063–1093 (2007)
26. Kocherlakota, N.: On tests of representative consumer asset pricing models. J. Monet. Econ. **26**, 285–304 (1990)
27. Lilla, M., Staffolani, S.: The evolution of wage inequality in Italy. Appl. Econ. **41**(15), 1873–1892 (2009)
28. MaCurdy, T.: The use of time-series processes to model the error structure of earnings in a longitudinal data analysis. J. Econom. **18**(1), 83–114 (1982)
29. Moffitt, R., Gottschalk, P.: Trends in the autocovariance structure of earnings in the U.S., 1969–1987, mimeo. Johns Hopkins University (1995)
30. Moffitt, R., Gottschalk, P.: Trends in the transitory variance of earnings in the United States. Econ. J. **112**, 68–73 (2002)
31. Moffitt, R., Gottschalk, P.: Trends in the transitory variance of male earnings in the U.S., 1970–2004. J. Hum. Resour. (2011, forthcoming)
32. Myck, M., Ochmann, R., Qari, S.: Dynamics of earnings and hourly wages in Germany. IZA Discussion Paper 3751 (2008)
33. Pozzi, L.: The coefficient of relative risk aversion: a Monte Carlo study investigating small sample estimator problems. Econ. Model. **20**(6), 923–940 (2003)
34. Ramos, X.: The covariance structure of earnings in Great Britain, 1991–1999. Economica **70**, 353–374 (2003)

35. Santos, A., Souza, A.: Earnings inequality in Brazil: is it permanent or transitory? Brazilian Rev. Econometrics **27**(2), 259–283 (2007)
36. Sologon, D., O'Donoghue, C.: Earnings dynamics and inequality in the EU; 1994–2001. CEPS/INSEAD Working Paper 2010-34 (2010a)
37. Sologon, D., O'Donoghue, C.: Earnings dynamics and inequality among men in Luxembourg, 1988–2004: evidence from administrative data. IZA Discussion Paper 3751 (2010b)
38. Stock, J., Wright, J., Yogo, M.: A survey of weak instruments and weak identification in generalised method of moments. J. Bus. Econ. Stat. **20**(5), 518–529 (2002)
39. Stock, J., Wright, J.: GMM with weak identification. Econometrica **68**, 1055–1096 (2000)
40. Tauchen, G.: Statistical properties of GMM estimators of structural parameters obtained from financial market data. J. Bus. Econ. Stat. **4**, 397–425 (1986)