Economics Letters 76 (2002) 35–45

economics letters

www.elsevier.com/locate/econbase

Jackknife minimum distance estimation

Gábor Kézdi[a], Jinyong Hahn[b,*], Gary Solon[a]

[a]Department of Economics, University of Michigan, Ann Arbor, MI, USA
[b]Department of Economics, Brown University, Box B, Providence, RI 02912, USA

Received 11 November 2001; accepted 27 November 2001

Abstract

We propose a jackknife minimum distance estimator designed to reduce the finite-sample bias of the optimal minimum distance estimator. Monte Carlo results indicate that our jackknife minimum distance estimator is a promising alternative to existing minimum distance procedures.

Keywords: Minimum distance; Jackknife

JEL classification: C1

1. Introduction

Minimum distance estimation has been used in various contexts to estimate structural coefficients based on estimates of reduced form coefficients. Given a set of estimators r_n that converges in probability to some function $f(\theta)$ of the structural parameter θ, minimum distance estimation tries to find a value $\hat{\theta}$ such that the distance between r_n and $f(\hat{\theta})$ is as small as possible (see Chamberlain (1984) and Abowd and Card (1989) for applications to panel models). In some sense, the method of indirect inference developed by Gourieroux et al. (1993) could also be understood as a simulation based approach to minimum distance estimation.

The asymptotically efficient approach to minimum distance estimation is the optimal minimum distance (OMD) estimator developed by Hansen (1982). Empirical researchers, however, have become increasingly hesitant to use OMD because it sometimes is severely biased in finite samples (see Altonji and Segal (1996) and Clark (1996)). Altonji and Segal proposed a split-sample approach for reducing the OMD estimator's bias, but their Monte Carlo experiments indicated that this estimator has poor efficiency properties. Indeed, Altonji and Segal concluded that the split-sample estimator is dominated by the equally weighted minimum distance (EWMD) estimator.

*Corresponding author. Tel.: +1-401-863-3883; fax: +1-401-863-1970.
E-mail address: Jinyong_Hahn@brown.edu (J. Hahn).

0165-1765/02/$ – see front matter
PII: S0165-1765(02)00016-2

To treat the related problem of finite sample bias in instrumental variables (IV) estimation, Angrist and Krueger (1995) similarly proposed a split-sample approach. Angrist et al. (1999) subsequently devised a more efficient jackknife IV estimator. In this paper, we pursue the suggestion in Angrist et al.'s closing sentence that their jackknife idea might be extended to the problem of reducing bias in OMD estimation. Our Monte Carlo experiments indicate that our jackknife minimum distance estimator (JMD) does succeed in reducing the OMD estimator's bias and also is more efficient than Altonji and Segal's split-sample estimator. The results suggest that jackknife minimum distance estimation is a promising alternative to existing minimum distance procedures.[1]

2. Intuition for jackknife minimum distance estimation

Consider the simple problem where we are given a sequence of random vectors $z_1, \dots, z_n$ satisfying $E[z_i] = \varphi(\theta_0)$ for some structural parameter of interest θ_0. The optimal minimum distance estimator $\hat{\theta}_{\text{OMD}}$ is defined as the solution to the problem:

$$\left(n^{-1}\sum_{i=1}^{n} z_i - \varphi(\theta)\right)' \hat{V}^{-1}\left(n^{-1}\sum_{i=1}^{n} z_i - \varphi(\theta)\right) \tag{1}$$

where $\hat{V}$ denotes some preliminary estimator of the covariance matrix V of z_i. It is well known that $\hat{\theta}_{\text{OMD}}$ is $\sqrt{n}$-consistent and asymptotically normal under a standard set of regularity conditions:

$$\sqrt{n}\left(\hat{\theta}_{\text{OMD}} - \theta_0\right) \to N\left(0, (D'V^{-1}D)^{-1}\right)$$

where $D \equiv \partial\varphi(\theta_0)/\partial\theta'$.

Unfortunately, asymptotics provide a poor approximation to the finite sample properties of the OMD estimator. Altonji and Segal (1996) and Clark (1996) found that the finite sample bias of $\hat{\theta}_{OMD}$ can be quite substantial. To illustrate the problem and how it might be relieved by our jackknife procedure, we begin with a problem similar to the ones considered by Altonji and Segal. Consider the simple panel model where:

$$y_{it} \overset{\text{i.i.d.}}{\sim} (0, \sigma^2), \qquad i = 1, \dots, n;\ t = 1, \dots, T$$

The goal is to estimate σ^2 by minimum distance. One version of OMD is given by:

$$\underset{\sigma^2}{\operatorname{argmin}}\left(n^{-1}\sum_{i=1}^{n} z_i - \ell_T\sigma^2\right)' \hat{V}^{-1}\left(n^{-1}\sum_{i=1}^{n} z_i - \ell_T\sigma^2\right)$$

where $\hat{V}$ is an estimated covariance matrix V of $z_i \equiv \left(y_{i1}^2, \dots, y_{iT}^2\right)'$, and ℓ_T is a T-dimensional column vector of ones.[2] In this case, the optimization problem is linear, so the solution has a closed form.

[1] We did not experiment with Hansen et al. (1996) continuous updating estimator because we have experienced computational problems with that estimator in other contexts.

[2] This is the minimum distance estimator based on the functional structure of the expectation of the diagonal elements of $n^{-1}\sum_{i=1}^{n} y_i y_i'$.

Therefore, the optimal minimum distance estimator is given by:

$$\hat{\sigma}^2_{\mathrm{OMD}} = \frac{\ell_T' \hat{V}^{-1} n^{-1} \sum_{i=1}^{n} z_i}{\ell_T' \hat{V}^{-1} \ell_T} = n^{-1} \sum_{i=1}^{n} \frac{\ell_T' \hat{V}^{-1} z_i}{\ell_T' \hat{V}^{-1} \ell_T}$$

Usually, we estimate $\hat{V}$ by:

$$\hat{V} = n^{-1} \sum_{i=1}^{n} (z_i - \ell_T \hat{\sigma}^2)(z_i - \ell_T \hat{\sigma}^2)'$$

where $\hat{\sigma}^2$ is some consistent estimator of σ^2. It is now obvious that $\hat{V}$ and z_i are correlated in general, and such correlation would induce a small sample bias in the resultant minimum distance estimator.[3]

It is relatively easy to fix this problem by modifying the Angrist et al. [3] intuition for their jackknife IV estimator. Let:

$$\hat{V}_{(i)} = (n-1)^{-1} \sum_{j \neq i}^{n} (z_j - \ell_T \hat{\sigma}^2_{(i)})(z_j - \ell_T \hat{\sigma}^2_{(i)})'$$

denote an estimator of V based on all the observations except the ith. Here, $\hat{\sigma}^2_{(i)}$ denotes some consistent preliminary estimator of σ^2 based on the same delete-i sample. Our jackknife minimum distance estimator $\hat{\sigma}^2_{\mathrm{JMD}}$ is given by:

$$\hat{\sigma}^2_{\mathrm{JMD}} = n^{-1} \sum_{i=1}^{n} \frac{\ell_T' \hat{V}_{(i)}^{-1} z_i}{\ell_T' \hat{V}_{(i)}^{-1} \ell_T}$$

Like the split sample based estimator of Altonji and Segal, this estimation strategy basically eliminates the correlation between z_i and the estimated V. Unlike Altonji and Segal's estimator, our $\hat{V}_{(i)}^{-1}$ is based on almost the entire sample, thereby reducing its variability.

Tables 1–4 show our Monte Carlo results for the simple i.i.d. case with $\sigma^2 = 1$ and with various distributions and various values for n and T. We present the bias, standard deviation (S.D.), root mean squared error (RMSE), and mean absolute error (MAE) for the equally weighted minimum distance (EWMD), optimal minimum distance (OMD), split-sample minimum distance (SMD), and jackknife minimum distance (JMD) estimators.

Note that, by construction in this case, the EWMD estimator is optimal because it correctly imposes the restriction that the sample variances in different periods are equally informative. In contrast, the OMD, SMD, and JMD estimators assign differential weights to the sample variances from different periods on the basis of the estimated V matrices, which are nonscalar in finite samples. Nevertheless, this example provides a simple illustration of the operation and properties of the alternative estimators.

[3] The OMD estimator $\hat{\sigma}^2_{\mathrm{OMD}}$ is formally equivalent to the feasible generalized least squares (FGLS) estimator of the regression of $n^{-1} \sum_{i=1}^{n} z_i$ on ℓ_T. Therefore, the bias of the OMD estimator may seem to contradict Rothenberg's (1984) result that FGLS and GLS are higher order equivalent. Rothenberg's derivation, however, required that an odd function and an even function of the 'error term' in the model be uncorrelated, which is clearly violated in the present example.

Table 1
Performance of jackknife minimum distance estimator

$n=50$	$T=10$	EWMD	OMD	SMD	JMD	OMD/JMD	SMD/JMD
Bias	$t(5)$	0.000	−0.148	0.002	0.000	479.49	−7.11
	$t(10)$	0.000	−0.088	0.000	0.000	397.39	−0.50
	Normal	0.000	−0.055	0.000	0.000	161.21	0.57
	Uniform	0.000	−0.009	0.001	0.000	687.29	−103.89
	Lognormal	−0.008	−0.428	0.014	0.025	−17.25	0.55
	Exp	0.000	−0.217	0.006	0.006	−36.18	1.05
S.D.	$t(5)$	0.121	0.108	0.276	0.174	0.62	1.58
	$t(10)$	0.077	0.088	0.148	0.102	0.86	1.45
	Normal	0.063	0.078	0.117	0.078	1.00	1.50
	Uniform	0.040	0.050	0.070	0.044	1.14	1.59
	Lognormal	0.413	0.414	5.338	1.701	0.24	3.14
	Exp	0.125	0.121	0.311	0.218	0.55	1.42
RMSE	$t(5)$	0.121	0.184	0.276	0.174	1.06	1.58
	$t(10)$	0.077	0.124	0.148	0.102	1.22	1.45
	Normal	0.063	0.095	0.117	0.078	1.22	1.50
	Uniform	0.040	0.051	0.070	0.044	1.15	1.59
	Lognormal	0.413	0.595	5.338	1.701	0.35	3.14
	Exp	0.125	0.248	0.311	0.218	1.14	1.42
MAE	$t(5)$	0.088	0.160	0.175	0.123	1.30	1.43
	$t(10)$	0.061	0.104	0.114	0.080	1.30	1.43
	Normal	0.050	0.077	0.093	0.062	1.25	1.50
	Uniform	0.032	0.041	0.056	0.035	1.16	1.59
	Lognormal	0.265	0.495	1.134	0.509	0.97	2.23
	Exp	0.100	0.224	0.224	0.168	1.33	1.33

Note: Results are based on 10 000 Monte Carlo runs.

The results show that both the split-sample and jackknife estimators are successful in removing the bias of the conventional OMD estimator, with especially large improvements in smaller samples. The jackknife estimator regularly dominates the split-sample estimator with respect to standard deviation, RMSE, and MAE. Because the jackknife estimator removes bias while maintaining good precision, it often outperforms conventional OMD with respect to RMSE and MAE.

3. Jackknife minimum distance estimation and an application to earnings dynamics

The jackknife modification in the previous example was relatively straightforward given the linearity of $\varphi(\theta_0)$. We now consider the more general case where $\varphi(\theta)$ is potentially nonlinear in θ. The first order condition for (1) implies that $\hat{\theta}_{\mathrm{OMD}}$ satisfies:

$$\left(\frac{\partial\varphi(\hat{\theta}_{\mathrm{OMD}})}{\partial\theta'}\right)'\hat{V}^{-1}\left(n^{-1}\sum_{i=1}^{n}z_i-\varphi(\hat{\theta}_{\mathrm{OMD}})\right)=0$$

Table 2
Performance of jackknife minimum distance estimator

$n = 100$	$T = 10$	EWMD	OMD	SMD	JMD	OMD/JMD	SMD/JMD
Bias	$t(5)$	0.001	−0.109	0.003	0.000	−374.54	10.98
	$t(10)$	−0.001	−0.057	−0.001	−0.001	105.82	0.95
	Normal	0.000	−0.032	−0.002	−0.001	45.90	2.14
	Uniform	0.000	−0.005	0.001	0.000	−46.88	6.46
	Lognormal	−0.004	−0.414	−0.003	0.005	−81.18	−0.57
	Exp	−0.001	−0.156	−0.001	0.001	−234.41	−2.11
S.D.	$t(5)$	0.084	0.072	0.154	0.111	0.64	1.38
	$t(10)$	0.055	0.060	0.094	0.067	0.90	1.41
	Normal	0.045	0.052	0.072	0.051	1.01	1.42
	Uniform	0.028	0.032	0.044	0.030	1.08	1.49
	Lognormal	0.297	0.158	0.958	0.552	0.29	1.74
	Exp	0.090	0.091	0.177	0.143	0.64	1.24
RMSE	$t(5)$	0.084	0.131	0.154	0.111	1.17	1.39
	$t(10)$	0.055	0.083	0.094	0.067	1.24	1.41
	Normal	0.045	0.061	0.072	0.051	1.19	1.42
	Uniform	0.028	0.033	0.044	0.030	1.09	1.49
	Lognormal	0.297	0.443	0.958	0.552	0.80	1.74
	Exp	0.090	0.180	0.177	0.143	1.26	1.24
MAE	$t(5)$	0.064	0.114	0.110	0.084	1.37	1.32
	$t(10)$	0.044	0.068	0.074	0.053	1.29	1.41
	Normal	0.036	0.049	0.058	0.040	1.21	1.43
	Uniform	0.023	0.026	0.035	0.024	1.09	1.49
	Lognormal	0.202	0.422	0.427	0.339	1.25	1.26
	Exp	0.071	0.160	0.136	0.112	1.43	1.22

Note: Results are based on 10 000 Monte Carlo runs.

or

$$\hat{\Lambda}\left(n^{-1}\sum_{i=1}^{n} z_i - \varphi(\hat{\theta}_{\mathrm{OMD}})\right) = 0$$

where

$$\hat{\Lambda} \equiv \left(\frac{\partial\varphi(\hat{\theta}_{\mathrm{OMD}})}{\partial\theta'}\right)' \hat{V}^{-1}$$

Therefore, we may interpret OMD as the method of moments estimator based on:

$$\Lambda E[z_i - \varphi(\theta)] = 0$$

where $\Lambda \equiv D'V^{-1}$ is not known and has to be estimated. Obviously the estimator $\hat{\Lambda}$ will be correlated with every observation, which may lead to a substantial finite sample bias. Our jackknife minimum

Table 3
Performance of jackknife minimum distance estimator

$n=500$	$T=10$	EWMD	OMD	SMD	JMD	OMD/JMD	SMD/JMD
Bias	$t(5)$	0.000	−0.041	−0.001	0.000	380.51	4.78
	$t(10)$	0.000	−0.014	0.000	0.000	−42.94	0.72
	Normal	0.000	−0.007	0.000	0.000	33.67	0.35
	Uniform	0.000	−0.001	0.000	0.000	−15.53	1.55
	Lognormal	0.000	−0.230	−0.001	−0.001	191.83	1.00
	Exp	0.000	−0.044	−0.001	0.000	−95.53	−1.96
S.D.	$t(5)$	0.042	0.034	0.061	0.050	0.68	1.22
	$t(10)$	0.024	0.025	0.037	0.026	0.97	1.41
	Normal	0.020	0.021	0.029	0.021	1.00	1.41
	Uniform	0.013	0.013	0.018	0.013	1.02	1.43
	Lognormal	0.147	0.082	0.300	0.263	0.31	1.14
	Exp	0.040	0.042	0.067	0.048	0.88	1.40
RMSE	$t(5)$	0.042	0.054	0.061	0.050	1.07	1.22
	$t(10)$	0.024	0.029	0.037	0.026	1.11	1.41
	Normal	0.020	0.022	0.029	0.021	1.07	1.41
	Uniform	0.013	0.013	0.018	0.013	1.02	1.43
	Lognormal	0.147	0.245	0.300	0.263	0.93	1.14
	Exp	0.040	0.061	0.067	0.048	1.27	1.40
MAE	$t(5)$	0.030	0.045	0.047	0.035	1.29	1.34
	$t(10)$	0.019	0.023	0.029	0.021	1.12	1.41
	Normal	0.016	0.018	0.023	0.017	1.07	1.41
	Uniform	0.010	0.010	0.015	0.010	1.03	1.44
	Lognormal	0.103	0.231	0.186	0.157	1.47	1.18
	Exp	0.032	0.051	0.053	0.038	1.33	1.39

Note: Results are based on 10 000 Monte Carlo runs.

distance estimator is based on such a method of moments interpretation. Therefore, we can develop a jackknife minimum distance estimator as the method of moments estimator solving:

$$n^{-1}\sum_{i=1}^{n}\left(\frac{\partial\varphi(\theta_{(i)})}{\partial\theta'}\right)'\hat{V}_{(i)}^{-1}\left(z_i-\varphi\left(\hat{\theta}_{\mathrm{JMD}}\right)\right)=0$$

where $\theta_{(i)}$ and $\hat{V}_{(i)}=n^{-1}\Sigma_{j\neq i}^{n}(z_i-\varphi(\theta_{(i)}))(z_i-\varphi(\theta_{(i)}))'$ are based on the delete-i sample.

One obvious application for our jackknife minimum distance estimator is to variance components models of earnings dynamics (see Lillard and Willis (1978), MaCurdy (1982), Abowd and Card (1989), Baker (1997), and Baker and Solon (1999)). To illustrate, consider the model:

$$\varepsilon_{it}=u_i+v_{it}\qquad i=1,\ldots,n;\ t=1,\ldots,T \tag{2}$$

which expresses log earnings ε_{it} for worker i in year t as the sum of two orthogonal components, a time-invariant component u_i and a transitory component v_{it}. For simplicity, we abstract from the experience profile of earnings, and we assume without loss of generality that $E[u_i]=E[v_{it}]=0$.

Table 4
Performance of jackknife minimum distance estimator

$n=1000$	$T=10$	EWMD	OMD	SMD	JMD	OMD/JMD	SMD/JMD
Bias	$t(5)$	0.000	−0.026	0.000	0.000	−224.76	1.33
	$t(10)$	0.000	−0.008	0.000	0.000	63.61	1.06
	Normal	0.000	−0.004	0.000	0.000	21.59	−0.13
	Uniform	0.000	−0.001	0.000	0.000	4.09	1.03
	Lognormal	0.002	−0.161	0.003	0.002	−66.92	1.08
	Exp	0.000	−0.024	0.000	0.000	77.84	0.51
S.D.	$t(5)$	0.027	0.025	0.043	0.031	0.79	1.37
	$t(10)$	0.017	0.018	0.025	0.018	0.98	1.42
	Normal	0.014	0.015	0.021	0.014	1.01	1.42
	Uniform	0.009	0.009	0.013	0.009	1.01	1.42
	Lognormal	0.105	0.068	0.202	0.159	0.42	1.27
	Exp	0.028	0.029	0.045	0.032	0.93	1.42
RMSE	$t(5)$	0.027	0.036	0.043	0.031	1.15	1.37
	$t(10)$	0.017	0.019	0.025	0.018	1.07	1.42
	Normal	0.014	0.015	0.021	0.014	1.04	1.42
	Uniform	0.009	0.009	0.013	0.009	1.01	1.42
	Lognormal	0.105	0.174	0.202	0.159	1.10	1.27
	Exp	0.028	0.038	0.045	0.032	1.21	1.42
MAE	$t(5)$	0.021	0.030	0.033	0.024	1.24	1.36
	$t(10)$	0.014	0.015	0.020	0.014	1.08	1.43
	Normal	0.011	0.012	0.016	0.012	1.03	1.41
	Uniform	0.007	0.007	0.010	0.007	1.01	1.42
	Lognormal	0.076	0.161	0.132	0.111	1.46	1.20
	Exp	0.023	0.031	0.036	0.025	1.24	1.42

Note: Results are based on 10 000 Monte Carlo runs.

Let $\theta_1 \equiv \sigma_u^2 \equiv \text{Var}(u_i)$. Suppose that $v_i \equiv (v_{i1}, \ldots, v_{iT})'$ follows a particular time series structure parameterized by a finite dimensional parameter θ_2. Our parameter of interest is $\theta \equiv (\theta_1, \theta_2)$. This would imply that the $T \times T$ covariance matrix Ω of $\varepsilon_i = (\varepsilon_{i1}, \ldots, \varepsilon_{iT})'$ can be regarded as a function of θ. Therefore, we may write $\Omega = \Omega(\theta)$. We may also define $f(\theta) \equiv v(\Omega(\theta))$, where $v(\Omega)$ denotes the $\frac{1}{2}T(T+1)$ dimensional column vector that is obtained from $\text{vec}(\Omega)$ by eliminating all supradiagonal elements of Ω.[4] For example, if it is assumed that v_i follows an AR(1) model, then:

[4]Suppose that $T = 3$, and

$$\Omega = \begin{bmatrix} \omega_{11} & \omega_{12} & \omega_{13} \\ \omega_{21} & \omega_{22} & \omega_{23} \\ \omega_{31} & \omega_{23} & \omega_{33} \end{bmatrix}$$

Then, $v(\Omega) = (\omega_{11}, \omega_{21}, \omega_{31}, \omega_{22}, \omega_{23}, \omega_{33})'$. For example, see Magnus and Neudecker (1988) for more discussion.

$$\Omega = \sigma_v^2 \begin{pmatrix} 1 & \rho & \cdots & \rho^{T-1} \\ \rho & 1 & \cdots & \rho^{T-2} \\ \vdots & \vdots & \cdots & \vdots \\ \rho^{T-1} & \rho^{T-2} & \cdots & 1 \end{pmatrix} + \sigma_u^2 \begin{pmatrix} 1 & 1 & \cdots & 1 \\ 1 & 1 & \cdots & 1 \\ \vdots & \vdots & \cdots & \vdots \\ 1 & 1 & \cdots & 1 \end{pmatrix} \quad (3)$$

where $\sigma_v^2 \equiv \text{Var}(v_{it})$ and ρ denotes the AR(1) coefficient. Our parameter of interest in this example is $(\rho, \sigma_u^2, \sigma_v^2)$.

Tables 5–7 report Monte Carlo results on the finite sample behavior of the EWMD, OMD, and jackknife estimators. The experiments set $\sigma_u^2 = \sigma_v^2 = 1$, reflecting the empirically relevant case in

Table 5
Performance of jackknife minimum distance estimator

$n=100$	$T=10$	$\hat{\sigma}_u^2$			$\hat{\sigma}_v^2$			$\hat{\rho}$		
		EWMD	OMD	JMD	EWMD	OMD	JMD	EWMD	OMD	JMD
Bias	$\rho=0.0$	0.003	−0.174	0.102	0.001	−0.206	0.033	−0.002	0.000	0.002
	$\rho=0.1$	−0.001	−0.174	0.103	0.001	−0.205	0.034	0.002	0.001	0.002
	$\rho=0.3$	−0.003	−0.175	0.101	0.003	−0.202	0.038	0.003	0.001	0.002
	$\rho=0.5$	−0.011	−0.179	0.095	0.010	−0.193	0.050	0.003	0.002	0.002
	$\rho=0.7$	−0.043	−0.206	0.050	0.044	−0.160	0.101	0.000	0.002	0.001
	$\rho=0.9$	−0.961	−2.866	−0.020	0.964	2.508	−0.314	−0.009	0.001	−0.038
S.D.	$\rho=0.0$	0.154	0.220	0.323	0.048	0.071	0.107	0.042	0.068	0.070
	$\rho=0.1$	0.156	0.224	0.329	0.050	0.074	0.110	0.045	0.069	0.071
	$\rho=0.3$	0.164	0.236	0.345	0.060	0.086	0.124	0.057	0.071	0.073
	$\rho=0.5$	0.180	0.257	0.369	0.088	0.115	0.161	0.070	0.071	0.074
	$\rho=0.7$	0.244	0.315	0.452	0.198	0.210	0.331	0.075	0.070	0.076
	$\rho=0.9$	2.270	12.051	1.384	2.319	11.746	1.469	0.075	0.064	0.147
RMSE	$\rho=0.0$	0.163	0.279	0.339	0.050	−0.154	0.112	0.042	0.068	0.070
	$\rho=0.1$	0.156	0.286	0.345	0.051	0.227	0.117	0.045	0.070	0.071
	$\rho=0.3$	0.164	0.296	0.365	0.113	0.312	0.142	0.057	0.072	0.073
	$\rho=0.5$	0.180	0.315	0.410	0.335	0.539	0.228	0.070	0.072	0.075
	$\rho=0.7$	0.248	0.377	0.639	0.938	1.142	0.647	0.075	0.070	0.093
	$\rho=0.9$	2.465	13.407	2.955	4.032	13.126	3.182	0.076	0.064	5.093
MAE	$\rho=0.0$	0.122	0.232	0.262	0.038	0.206	0.089	0.033	0.054	0.056
	$\rho=0.1$	0.124	0.235	0.267	0.040	0.205	0.091	0.036	0.055	0.056
	$\rho=0.3$	0.130	0.243	0.279	0.048	0.203	0.104	0.045	0.056	0.057
	$\rho=0.5$	0.143	0.258	0.296	0.069	0.200	0.134	0.056	0.056	0.058
	$\rho=0.7$	0.191	0.303	0.343	0.141	0.222	0.229	0.060	0.054	0.058
	$\rho=0.9$	1.239	3.030	0.726	1.248	2.936	0.820	0.060	0.049	0.071

Note: Results for EWMD and OMD are based on 5000 Monte Carlo runs. Results for JMD are based on 500 Monte Carlo runs. The Monte Carlo was carried out using Matlab on a Sun Enterprise 250 server. The Monte Carlo took about 5 h for each value of ρ.

Table 6
Performance of jackknife minimum distance estimator

$n=500$	$T=10$	$\hat{\sigma}_u^2$			$\hat{\sigma}_v^2$			$\hat{\rho}$		
		EWMD	OMD	JMD	EWMD	OMD	JMD	EWMD	OMD	JMD
Bias	$\rho=0.0$	0.004	−0.030	0.014	−0.001	−0.041	0.003	0.000	0.000	0.000
	$\rho=0.1$	0.001	−0.033	0.014	0.000	−0.041	0.003	0.000	0.000	0.000
	$\rho=0.3$	0.001	−0.033	0.014	0.001	−0.041	0.004	0.001	0.000	0.001
	$\rho=0.5$	−0.001	−0.033	0.013	0.002	−0.040	0.006	0.001	0.000	0.001
	$\rho=0.7$	−0.006	−0.036	0.008	0.007	−0.036	0.012	0.000	0.001	0.002
	$\rho=0.9$	−0.137	−0.071	−0.055	0.138	0.000	0.069	−0.002	0.001	0.004
S.D.	$\rho=0.0$	0.070	0.074	0.079	0.021	0.023	0.024	0.019	0.019	0.019
	$\rho=0.1$	0.071	0.076	0.081	0.022	0.024	0.026	0.020	0.019	0.019
	$\rho=0.3$	0.075	0.079	0.085	0.026	0.029	0.031	0.025	0.020	0.020
	$\rho=0.5$	0.082	0.086	0.093	0.037	0.038	0.041	0.031	0.020	0.020
	$\rho=0.7$	0.103	0.102	0.109	0.072	0.060	0.067	0.034	0.019	0.020
	$\rho=0.9$	0.545	0.216	0.188	0.556	0.197	0.167	0.036	0.018	0.014
RMSE	$\rho=0.0$	0.070	0.080	0.080	0.021	0.047	0.024	0.019	0.019	0.019
	$\rho=0.1$	0.071	0.083	0.082	0.022	0.048	0.026	0.020	0.019	0.019
	$\rho=0.3$	0.075	0.086	0.086	0.026	0.050	0.031	0.025	0.020	0.020
	$\rho=0.5$	0.082	0.092	0.094	0.037	0.055	0.042	0.031	0.020	0.020
	$\rho=0.7$	0.103	0.108	0.109	0.072	0.070	0.068	0.034	0.019	0.020
	$\rho=0.9$	0.562	0.228	0.196	0.572	0.197	0.180	0.036	0.018	0.014
MAE	$\rho=0.0$	0.056	0.065	0.065	0.017	0.042	0.020	0.015	0.015	0.015
	$\rho=0.1$	0.057	0.066	0.066	0.018	0.042	0.021	0.016	0.015	0.016
	$\rho=0.3$	0.060	0.069	0.070	0.021	0.043	0.025	0.020	0.016	0.016
	$\rho=0.5$	0.065	0.074	0.076	0.030	0.046	0.034	0.025	0.016	0.016
	$\rho=0.7$	0.082	0.086	0.089	0.057	0.057	0.053	0.027	0.015	0.016
	$\rho=0.9$	0.310	0.169	0.142	0.312	0.149	0.123	0.029	0.014	0.010

Note: Results for EWMD and OMD are based on 5000 Monte Carlo runs. Results for JMD are based on 500 Monte Carlo runs. The Monte Carlo was carried out using Matlab on a Sun Enterprise 250 server. The Monte Carlo took about 36 h for each value of ρ.

which the permanent and transitory variance components are of similar magnitude. In the present context, unlike in the *i.i.d.* example in the previous section, EWMD is not optimal (unless $\rho=0$). None of the estimators exhibits much bias for ρ. For the variance components σ_u^2 and σ_v^2, however, OMD again is severely biased, and the jackknife estimator has some success in reducing the bias. As n grows larger, the jackknife estimator shows good precision relative to the EWMD estimator, especially when ρ is large. In such settings, the jackknife estimator often outperforms both EWMD and OMD with respect to RMSE and MAE.

Table 7
Performance of jackknife minimum distance estimator

$n = 1000$	$T = 10$	$\hat{\sigma}_u^2$			$\hat{\sigma}_v^2$			$\hat{\rho}$		
		EWMD	OMD	JMD	EWMD	OMD	JMD	EWMD	OMD	JMD
Bias	$\rho = 0.0$	0.004	−0.021	0.005	−0.001	−0.021	0.002	0.000	0.000	0.001
	$\rho = 0.1$	0.001	−0.017	0.005	0.000	−0.021	0.002	0.000	0.000	0.001
	$\rho = 0.3$	0.001	−0.017	0.005	0.001	−0.020	0.002	0.001	0.000	0.001
	$\rho = 0.5$	−0.001	−0.017	0.005	0.002	−0.020	0.003	0.001	0.000	0.000
	$\rho = 0.7$	−0.006	−0.018	0.003	0.007	−0.018	0.004	0.000	0.000	0.000
	$\rho = 0.9$	−0.137	−0.032	−0.043	0.138	−0.003	0.046	−0.002	0.000	0.003
S.D.	$\rho = 0.0$	0.070	0.016	0.055	0.021	0.016	0.016	0.019	0.013	0.012
	$\rho = 0.1$	0.071	0.053	0.056	0.022	0.016	0.017	0.020	0.013	0.012
	$\rho = 0.3$	0.075	0.055	0.059	0.026	0.019	0.019	0.025	0.013	0.012
	$\rho = 0.5$	0.082	0.060	0.063	0.037	0.025	0.025	0.031	0.013	0.013
	$\rho = 0.7$	0.103	0.070	0.074	0.072	0.040	0.041	0.034	0.012	0.013
	$\rho = 0.9$	0.545	0.135	0.125	0.556	0.122	0.108	0.036	0.012	0.010
RMSE	$\rho = 0.0$	0.070	0.026	0.055	0.021	0.026	0.016	0.019	0.013	0.012
	$\rho = 0.1$	0.071	0.055	0.056	0.022	0.026	0.017	0.020	0.013	0.012
	$\rho = 0.3$	0.075	0.058	0.059	0.026	0.028	0.019	0.025	0.013	0.012
	$\rho = 0.5$	0.082	0.062	0.063	0.037	0.037	0.025	0.031	0.013	0.013
	$\rho = 0.7$	0.103	0.073	0.074	0.072	0.072	0.041	0.034	0.012	0.013
	$\rho = 0.9$	0.562	0.139	0.132	0.572	0.572	0.117	0.036	0.012	0.010
MAE	$\rho = 0.0$	0.056	0.022	0.043	0.017	0.022	0.013	0.015	0.010	0.009
	$\rho = 0.1$	0.057	0.044	0.044	0.018	0.022	0.013	0.016	0.010	0.010
	$\rho = 0.3$	0.060	0.046	0.046	0.021	0.023	0.015	0.020	0.010	0.010
	$\rho = 0.5$	0.065	0.050	0.049	0.030	0.026	0.020	0.025	0.010	0.010
	$\rho = 0.7$	0.082	0.058	0.058	0.057	0.036	0.033	0.027	0.010	0.010
	$\rho = 0.9$	0.310	0.108	0.102	0.312	0.095	0.085	0.029	0.009	0.008

Note: Results for EWMD and OMD are based on 5000 Monte Carlo runs. Results for JMD are based on 500 Monte Carlo runs. The Monte Carlo was carried out using Matlab on a Pentium III 866 Mhz computer. The Monte Carlo took about 95 h for each value of ρ.

References

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

Altonji, J.G., Segal, L.M., 1996. Small-sample bias in GMM estimation of covariance structures. Journal of Business and Economic Statistics 14, 353–366.

Angrist, J.D., Imbens, G.W., Krueger, A.B., 1999. Jackknife instrumental variables estimation. Journal of Applied Econometrics 14, 57–67.

Angrist, J.D., Krueger, A.B., 1995. Split-sample instrumental variables estimates of the return to schooling. Journal of Business and Economic Statistics 13, 225–235.

Baker, M., 1997. Growth-rate heterogeneity and the covariance structure of life-cycle earnings. Journal of Labor Economics 15, 338–375.

Baker, M., Solon, G., 1999. Earnings Dynamics and Inequality among Canadian Men, 1976–1992: Evidence from Longitudinal Income Tax Records. NBER Working Paper No. 7370.

Chamberlain, G., 1984. Panel data. In: Griliches, Z., Intriligator, M.D. (Eds.). Handbook of Econometrics, Vol. 2. North-Holland, New York.

Clark, T.E., 1996. Small-sample properties of estimators of nonlinear models of covariance structures. Journal of Business and Economic Statistics 14, 367–373.

Gourieroux, C., Monfort, A., Renault, E., 1993. Indirect inference. Journal of Applied Econometrics 8, S85–S118.

Hansen, L.P., 1982. Large sample properties of generalized method of moments estimators. Econometrica 50, 1029–1054.

Hansen, L.P., Heaton, J., Yaron, A., 1996. Finite-sample properties of some alternative GMM estimators. Journal of Business and Economic Statistics 14, 262–280.

Lillard, L.A., Willis, R.J., 1978. Dynamic aspects of earning mobility. Econometrica 46, 985–1012.

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, 83–114.

Magnus, J.R., Neudecker, H., 1988. Matrix Differential Calculus with Applications in Statistics and Econometrics. Wiley, New York.

Rothenberg, T.J., 1984. Approximate normality of generalized least squares estimates. Econometrica 52, 811–825.