

Ronning, Gerd; Schneeweiss, Hans

Working Paper

## Panel regression with random noise

CESifo Working Paper, No. 2608

**Provided in Cooperation with:**

Ifo Institute – Leibniz Institute for Economic Research at the University of Munich

*Suggested Citation:* Ronning, Gerd; Schneeweiss, Hans (2009) : Panel regression with random noise, CESifo Working Paper, No. 2608, Center for Economic Studies and ifo Institute (CESifo), Munich

This Version is available at:

<https://hdl.handle.net/10419/26653>

**Standard-Nutzungsbedingungen:**

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

**Terms of use:**

*Documents in EconStor may be saved and copied for your personal and scholarly purposes.*

*You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.*

*If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.*

# Panel Regression with Random Noise

GERD RONNING  
HANS SCHNEEWEISS

CESIFO WORKING PAPER NO. 2608  
CATEGORY 12: EMPIRICAL AND THEORETICAL METHODS  
APRIL 2009

*An electronic version of the paper may be downloaded*

- *from the SSRN website:* [www.SSRN.com](http://www.SSRN.com)
- *from the RePEc website:* [www.RePEc.org](http://www.RePEc.org)
- *from the CESifo website:* [www.CESifo-group.org/wp](http://www.CESifo-group.org/wp)

# Panel Regression with Random Noise

## Abstract

The paper explores the effect of measurement errors on the estimation of a linear panel data model. The conventional fixed effects estimator, which ignores measurement errors, is biased. By correcting for the bias one can construct consistent and asymptotically normal estimators. In addition, we find estimates for the asymptotic variances of these estimators. The paper focuses on multiplicative errors, which are often deliberately added to the data in order to minimize their disclosure risk. They can be analyzed in a similar way as additive errors, but with some important and consequential differences.

JEL Code: C13.

Keywords: panel regression, multiplicative measurement errors, bias correction, asymptotic variance, disclosure control.

*Gerd Ronning*  
*Faculty of Economics*  
*University of Tübingen*  
*Mohlstrasse 36*  
*72074 Tübingen*  
*Germany*  
*gerd.ronning@uni-tuebingen.de*

*Hans Schneeweiss*  
*University of Munich*  
*Statistical Institute*  
*Akademiestrasse 1*  
*80799 Munich*  
*Germany*  
*schneew@stat.uni-muenchen.de*

March 30, 2009

Financial support by Bundesministerium für Bildung und Forschung (project "Wirtschaftsstatistische Paneldaten und faktische Anonymisierung") is gratefully acknowledged.

# 1 Introduction

The paper explores the effect of measurement errors on the estimation of a linear panel data regression model. The conventional fixed effects least squares estimator, which ignores measurement errors, is biased. By correcting for the bias we can construct consistent and asymptotically normal estimators, where asymptotically here means that the number of sample units tends to infinity.

Measurement errors can be additive or multiplicative. Additive measurement errors in panel data models have been extensively studied in the literature, Griliches and Hausman (1986), Hsiao and Taylor (1991), Wansbeek and Kooning (1991), Biørn (1996), Wansbeek (2001), Biørn and Krishnakumar (2008). But multiplicative measurement errors, though not uncommon in other models (see, e.g., Hwang (1986), Lin (1989), Carroll *et al.* (2006)), have not found much attention in the context of panel data models.

The present paper is a first attempt to fill this gap. It was motivated by the various worldwide endeavors to find methods for masking data so that their disclosure risk becomes negligible, see, e.g., Domingo-Ferrer and Saygin (2008). Data and in particular panel data that are released to the public should be not only nominally but also factually anonymous. Making them anonymous in this sense can be done by (slightly) distorting them. The distortion, of course, should be such that the disturbing effects on any subsequent scientific analysis of the data should be minimal or should be amenable to correction. One way of perturbing data is to mix them with random noise, see Kim (1986) as an early reference. This can be done by adding random measurement errors to the data or by multiplying them with measurement errors. The latter procedure is often preferred, as it takes automatically into account that large values of a sensitive variable are more prone to disclosure and hence need to be better protected. In contrast to an additive error, a multiplicative error will distort large values more than small values.

Another aspect of statistical disclosure control techniques is that the procedure used is typically made known to the public. In our case, this means that the measurement error variance is known to the statistician working with the data.

Although linear panel regressions can also be estimated without this knowledge, we here assume that the error variance is known. This assumption not only leads to simpler estimators but also to more efficient ones. Indeed, we use this knowledge as prior information to construct consistent estimators. In addition, we find estimates for the asymptotic variances of these estimators.

We only deal with one type of estimators, the familiar “within” LS estimator and its corrections. It uses the “within” variances and covariances for each sample unit over time instead of the overall (total) variances and covariances. In doing so, the unobserved heterogeneity which is present in the panel data is eliminated. There are other estimators that can do the same, especially instrumental variable estimators that use lagged values of the variables as instruments. But in order for them to function properly some assumptions about the variables must be satisfied. E.g., for instrumental variables to work the variables must be autocorrelated. No such assumptions are needed for the “within” estimators.

Although this paper is mainly concerned with multiplicative measurement errors, we also deal briefly with the additive case, just to show the parallel development in both cases.

While iid measurement errors are the main subject of our investigation, we also explore a special case of autocorrelated errors, a case which has been suggested especially for masking panel data: In addition to an iid part, the measurement error has a component which is random over the sample units but constant in time. It thus has a common factor structure, see Biørn(1996) and Hühne (2008).

The principles involved for constructing consistent estimators can be most easily presented in the context of a simple linear model with only one slope parameter. The procedures developed can then be extended to the case of a multiple regression, which, however, is not done in the present paper.

In Section 2 the linear panel model with measurement errors is presented. Section 3 introduces the within LS (naive) estimator of the slope parameter. In Section 4 the bias of the naive estimator is derived, and in Section 5 a corrected estimator is constructed. A corrected estimator of the regression error variance is presented in Section 6. Section 7 deals with the case of a common factor structure in the error process. In Section 8 asymptotic variances for the various corrected slope estimators are presented. Section 9 has a simulation study, where the asymptotic properties of the estimators are studied under small to medium size samples. Section 10 concludes. In an appendix, we demonstrate the equivalence of the delta method and the sandwich formula.

## 2 The model

### 2.1 Linear panel regression model

A panel consists of a sample of  $i = 1, \dots, N$  sample units (individuals, households, companies) observed over  $t = 1, \dots, T$  points of time (waves). Typically  $T$  is small and  $N$  is large. Asymptotics in this context therefore means asymptotics for  $N \rightarrow \infty$ .

For each individual and each wave, two variables are observed, a regressor (or independent) variable  $x$  and a response (or dependent) variable  $y$ . The panel data consists of the set  $(x_{it}, y_{it})$ ,  $i = 1, \dots, N$ ,  $t = 1, \dots, T$ . We assume a linear relation between  $x$  and  $y$  as follows:

$$y_{it} = \gamma_i + \beta x_{it} + \varepsilon_{it}, \quad i = 1, \dots, N, t = 1, \dots, T, \quad (1)$$

where  $\varepsilon$  is the error in the equation, and  $\gamma_i$  is the individual effect, giving rise to unobserved heterogeneity.  $\beta$  is the slope parameter to be estimated.

All variables including the  $\gamma_i$  are assumed to be random. However, the subsequent results, which are all of an asymptotic nature, remain valid if instead the  $\gamma_i$  or  $x_{it}$  are taken to be deterministic as long as their empirical moments behave asymptotically (i.e., for  $N \rightarrow \infty$ ) in the same way as if they were random. (Exact conditions could be formulated, but are not presented in this paper.)

The  $\varepsilon_{it}$  follow the usual assumptions: they are iid with mean 0 and variance  $\sigma_\varepsilon^2$  and they are independent of the  $\gamma_i$  and  $x_{it}$ . As to the latter, we assume that the vectors  $(\gamma_i, x_{i1}, \dots, x_{iT})$ ,  $i = 1, \dots, N$ , are iid. We make no assumptions about the joint distribution of  $(\gamma_i, x_{i1}, \dots, x_{iT})$  except that the moments as far as necessary exist. In particular, the  $x_{it}$  may be autocorrelated and the  $\gamma_i$  may be correlated with the  $x_{it}$ .

### 2.2 Measurement errors

The variables  $x$  and  $y$  are not known to the statistician. Instead surrogate variables  $x^a$  and  $y^a$ , which are the original variables intermingled with error, are released to him. In the case of *additive errors* these are given by

$$x_{it}^a = x_{it} + u_{it}, \quad y_{it}^a = y_{it} + v_{it} \quad (2)$$

and for *multiplicative errors* by

$$x_{it}^a = x_{it}U_{it}, \quad y_{it}^a = y_{it}V_{it} \quad (3)$$

with

$$U_{it} = 1 + u_{it}, \quad V_{it} = 1 + v_{it}.$$

In both cases, the pairs  $(u_{it}, v_{it})$  are iid with mean  $(0,0)$  and variances  $\sigma_u^2$ ,  $\sigma_v^2$  and covariance  $\sigma_{uv}$ . They are independent of all the  $(x_{it}, y_{it})$ . If the measurement errors  $u$  and  $v$  are used to mask the data, they will typically be uncorrelated. However, to cover the general case, we do not assume  $\sigma_{uv} = 0$ .

### 3 Estimators

The conventional “within” estimator of  $\beta$  constructed from the true data is given by:

$$\hat{\beta} = \frac{S_{xy}}{S_{xx}}, \quad (4)$$

where

$$\begin{aligned} S_{xy} &:= \frac{1}{NT} \sum_i \sum_t (x_{it} - \bar{x}_i)(y_{it} - \bar{y}_i) \\ S_{xx} &:= \frac{1}{NT} \sum_i \sum_t (x_{it} - \bar{x}_i)^2 \end{aligned}$$

with  $\bar{x}_i = \frac{1}{T} \sum_t x_{it}$  and  $\bar{y}_i = \frac{1}{T} \sum_t y_{it}$ . The use of the within variance  $S_{xx}$  and within covariance  $S_{xy}$  results in the elimination of the individual effects  $\gamma_i$ .  $S_{xx}$  and  $S_{xy}$  can also be written as

$$\begin{aligned} S_{xy} &= \overline{s_{xy}} \\ S_{xx} &= \overline{s_{xx}}, \end{aligned}$$

where

$$\begin{aligned} s_{xy} &:= (s_{xy})_i = \frac{1}{T} \sum_t (x_{it} - \bar{x}_i)(y_{it} - \bar{y}_i) \\ s_{xx} &:= (s_{xx})_i = \frac{1}{T} \sum_t (x_{it} - \bar{x}_i)^2 \end{aligned}$$

are, respectively, the covariance of  $x$  and  $y$  and the variance of  $x$  for an individual  $i$  over time  $t$ , and the bar denotes averaging over  $i = 1, \dots, N$ , e.g.:  $\overline{s_{xy}} = \frac{1}{N} \sum_i (s_{xy})_i$ .

This estimator of  $\beta$  is not feasible, as it uses the true data, which are unknown to the statistician. A feasible estimator can be constructed in the same way but with the randomly perturbed data in place of the original data:

$$\hat{\beta}^a = \frac{S_{xy}^a}{S_{xx}^a} \quad (5)$$

with

$$\begin{aligned} S_{xy}^a &:= \frac{1}{NT} \sum_i \sum_t (x_{it}^a - \bar{x}_i^a)(y_{it}^a - \bar{y}_i^a) = \overline{s_{xy}^a} \\ S_{xx}^a &:= \frac{1}{NT} \sum_i \sum_t (x_{it}^a - \bar{x}_i^a)^2 = \overline{s_{xx}^a}, \end{aligned}$$

where the superscript  $a$  indicates the use of the variables  $x^a$  and  $y^a$ . This estimator may be called a naive estimator, as it does not take the perturbations into account. Indeed, the naive estimator turns out to be biased. Derivation of the bias is the subject of the next section.

## 4 Derivation of bias

### 4.1 Probability limits

In order to derive the bias of  $\hat{\beta}^a$ , we need to compute the probability limits of  $S_{xy}^a$  and  $S_{xx}^a$ . We do this separately for the two cases of additive and multiplicative errors.

#### 4.1.1 Additive errors

From the definition of  $S_{xy}^a$  and  $S_{xy}$  and because of (2) it follows that

$$\begin{aligned} S_{xy}^a - S_{xy} &= \overline{s_{xy}^a} - \overline{s_{xy}} \\ &= \overline{s_{xv}} + \overline{s_{yu}} + \overline{s_{uv}}, \end{aligned} \tag{6}$$

where the last three terms are defined in the same way as  $\overline{s_{xy}}$  above. As  $N \rightarrow \infty$  these terms go to the corresponding expected values:

$$\text{plim}(S_{xy}^a - S_{xy}) = \mathbb{E}s_{xv} + \mathbb{E}s_{yu} + \mathbb{E}s_{uv}. \tag{7}$$

Note that, e.g.,  $\mathbb{E}s_{xv}$  is actually  $\mathbb{E}(s_{xv})_i$ , where  $(s_{xv})_i = \frac{1}{T} \sum_t (x_{it} - \bar{x}_i)(v_{it} - \bar{v}_i)$ . But since the expectation of this term is independent of  $i$ , we omit the index  $i$ . Now, because of the independence assumption,  $\mathbb{E}s_{xv} = \mathbb{E}s_{yu} = 0$ . As to  $\mathbb{E}s_{uv}$ , we have

$$\mathbb{E}s_{uv} = (1 - \frac{1}{T})\sigma_{uv}.$$

Consequently,

$$\text{plim}(S_{xy}^a - S_{xy}) = (1 - \frac{1}{T})\sigma_{uv}. \tag{8}$$



In a similar way or just by setting  $y = x$  and  $v = u$ , we obtain

$$\text{plim}(S_{xx}^a - S_{xx}) = (1 - \frac{1}{T})\sigma_u^2 \quad (9)$$

$$\text{plim}(S_{yy}^a - S_{yy}) = (1 - \frac{1}{T})\sigma_v^2 \quad (10)$$

The factor  $1 - \frac{1}{T} = \frac{T-1}{T}$ , which appears in (8) and (9) and in many of the subsequent expressions is of course the correction for degrees of freedom. As typically  $T$  is not large in panel data, this factor can never be suppressed. It is an important distinction of panel data analysis from cross section data analysis.

#### 4.1.2 Multiplicative errors

For multiplicative errors, we replace, in the arguments above,  $u_{it}$  and  $v_{it}$  with  $x_{it}u_{it}$  and  $y_{it}v_{it}$ , respectively. We then obtain in place of (7)

$$\text{plim}(S_{xy}^a - S_{xy}) = \mathbb{E}s_{x(yv)} + \mathbb{E}s_{y(xu)} + \mathbb{E}s_{(xu)(yv)}, \quad (11)$$

where, e.g.,

$$s_{(xu)(yv)} := (s_{(xu)(yv)})_i = \frac{1}{T} \sum_t \left( x_{it}u_{it} - \frac{1}{T} \sum_t x_{it}u_{it} \right) \left( y_{it}v_{it} - \frac{1}{T} \sum_t y_{it}v_{it} \right).$$

Again by the independence assumption,  $\mathbb{E}s_{x(yv)} = \mathbb{E}s_{y(xu)} = 0$ . As to  $\mathbb{E}s_{(xu)(yv)}$ , we obtain

$$\begin{aligned} \mathbb{E}s_{(xu)(yv)} &= \frac{1}{T} \sum_t \mathbb{E}(x_{it}y_{it}u_{it}v_{it}) - \frac{1}{T^2} \sum_t \sum_s \mathbb{E}(x_{it}y_{is}u_{it}v_{is}) \\ &= \sigma_{uv} \mathbb{E}m_{xy} - \frac{1}{T} \sigma_{uv} \mathbb{E}m_{xy} \\ &= (1 - \frac{1}{T}) \sigma_{uv} \mathbb{E}m_{xy}, \end{aligned} \quad (12)$$

where  $m_{xy} := (m_{xy})_i = \frac{1}{T} \sum_t (x_{it}y_{it})$ . Consequently,

$$\text{plim}(S_{xy}^a - S_{xy}) = (1 - \frac{1}{T}) \sigma_{uv} \mathbb{E}m_{xy}. \quad (13)$$

By setting  $y = x$  and  $v = u$ , we finally obtain

$$\text{plim}(S_{xx}^a - S_{xx}) = (1 - \frac{1}{T}) \sigma_u^2 \mathbb{E}m_{xx} \quad (14)$$

$$\text{plim}(S_{yy}^a - S_{yy}) = (1 - \frac{1}{T}) \sigma_v^2 \mathbb{E}m_{yy}. \quad (15)$$

## 4.2 Bias

Substituting the probability limits in (5), we immediately obtain expressions for the probability limit of the naive estimator and thus implicitly for the (asymptotic) bias.

### 4.2.1 Additive errors

By (5), (8), and (9),

$$\begin{aligned}\text{plim}\hat{\beta}^a &= \frac{\text{plim}S_{xy} + (1 - \frac{1}{T})\sigma_{uv}}{\text{plim}S_{xx} + (1 - \frac{1}{T})\sigma_u^2} \\ &= \frac{\beta\mathbb{E}s_{xx} + (1 - \frac{1}{T})\sigma_{uv}}{\mathbb{E}s_{xx} + (1 - \frac{1}{T})\sigma_u^2},\end{aligned}\tag{16}$$

where we used  $\text{plim}S_{xx} = \mathbb{E}s_{xx}$  and  $\text{plim}S_{xy} = \beta\mathbb{E}s_{xx}$ , which follows from (1).

### 4.2.2 Multiplicative errors

By (5), (13), (14), and the model equation (1),

$$\begin{aligned}\text{plim}\hat{\beta}^a &= \frac{\text{plim}S_{xy} + (1 - \frac{1}{T})\sigma_{uv}\mathbb{E}m_{xy}}{\text{plim}S_{xx} + (1 - \frac{1}{T})\sigma_u^2\mathbb{E}m_{xx}} \\ &= \frac{\beta\mathbb{E}s_{xx} + (1 - \frac{1}{T})\sigma_{uv}\mathbb{E}m_{xy}}{\mathbb{E}s_{xx} + (1 - \frac{1}{T})\sigma_u^2\mathbb{E}m_{xx}}.\end{aligned}\tag{17}$$

In the additive case, the bias depends on  $\mathbb{E}s_{xx}$ , and in the multiplicative case, it depends in addition on  $\mathbb{E}m_{xx}$  and, if  $\sigma_{uv} \neq 0$ , on  $\mathbb{E}m_{xy}$ . These expectations in turn depend on the joint distribution of the vector  $(\gamma_i, x_{i1}, \dots, x_{iT})$ , e.g., on whether the  $x_{it}$  are iid or are autocorrelated or whether the individual effects are uncorrelated or correlated with the  $x_{it}$ . To derive more explicit expressions for the bias, one would thus have to know the stochastic law governing the  $\gamma_i$  and the  $x_{it}$ . We do not pursue this line, but see Ronning (2007). It turns out that for constructing consistent estimators it is not necessary to know the distribution of  $(\gamma_i, x_{i1}, \dots, x_{iT})$ . However, the asymptotic variances of the estimators depend on this distribution, see Section 9.

## 5 Bias correction

We can derive correction formulas either by solving the bias formulas (16) and (17), respectively, for  $\beta$  or by using (4) directly. Taking the

latter course, we construct a corrected estimator of  $\beta$  as

$$\hat{\beta}^c = \frac{S_{xy}^c}{S_{xx}^c}, \quad (18)$$

where  $S_{xx}^c$  and  $S_{xy}^c$  are suitable “corrections” of  $S_{xx}$  and  $S_{xy}$ , respectively, which we obtain from Section 4.1 for the additive as well as for the multiplicative case, such that  $\text{plim}(S_{xy}^c - S_{xy}) = 0$  and  $\text{plim}(S_{xx}^c - S_{xx}) = 0$ .

## 5.1 Additive errors

In the additive case, we use (8) and (9) to estimate  $S_{xy}$  and  $S_{xx}$ , respectively, and obtain the following corrected estimator of  $\beta$ :

$$\hat{\beta}^c = \frac{S_{xy}^a - (1 - \frac{1}{T})\sigma_{uv}}{S_{xx}^a - (1 - \frac{1}{T})\sigma_u^2}. \quad (19)$$

## 5.2 Multiplicative errors

In the multiplicative case, we use (13) and (14) to estimate  $S_{xy}$  and  $S_{xx}$ , respectively, and obtain the following corrected estimator of  $\beta$ :

$$\hat{\beta}^c = \frac{S_{xy}^a - (1 - \frac{1}{T})\sigma_{uv}\hat{\mathbb{E}}m_{xy}}{S_{xx}^a - (1 - \frac{1}{T})\sigma_u^2\hat{\mathbb{E}}m_{xx}}, \quad (20)$$

where  $\hat{\mathbb{E}}m_{xy}$  and  $\hat{\mathbb{E}}m_{xx}$  are suitable estimates of  $\mathbb{E}m_{xy}$  and  $\mathbb{E}m_{xx}$ . In order to find these estimates, we consider

$$\overline{m_{xy}^a} := \frac{1}{NT} \sum_i \sum_t x_{it}^a y_{it}^a. \quad (21)$$

We have

$$\overline{m_{xy}^a} = \frac{1}{N} \sum_i \frac{1}{T} \sum_t x_{it} y_{it} U_{it} V_{it},$$

which tends in probability to

$$\frac{1}{T} \sum_t \mathbb{E}(x_{it} y_{it} U_{it} V_{it}) = \mathbb{E} \left( \frac{1}{T} \sum_t x_{it} y_{it} \right) \mathbb{E}(UV) = \mathbb{E}m_{xy} (\sigma_{uv} + 1).$$

It follows that

$$\hat{\mathbb{E}}m_{xy} := \frac{\overline{m_{xy}^a}}{1 + \sigma_{uv}} \quad (22)$$

is a consistent estimate of  $\mathbb{E}m_{xy}$ . Similarly,

$$\hat{\mathbb{E}}m_{xx} := \frac{\overline{m_{xx}^a}}{1 + \sigma_u^2} \quad (23)$$

is a consistent estimate of  $\mathbb{E} m_{xx}$ , where

$$\overline{m_{xx}^a} := \frac{1}{NT} \sum_i \sum_t x_{it}^{a2}. \quad (24)$$

With the abbreviations  $q_{uv} := (1 - \frac{1}{T}) \frac{\sigma_{uv}}{1 + \sigma_{uv}}$  and  $q_u := (1 - \frac{1}{T}) \frac{\sigma_u^2}{1 + \sigma_u^2}$ , we can write the corrected estimator as

$$\hat{\beta}^c = \frac{S_{xy}^a - q_{uv} \overline{m_{xy}^a}}{S_{xx}^a - q_u \overline{m_{xx}^a}}. \quad (25)$$

## 6 Consistent estimation of $\sigma_\epsilon^2$

In a similar way we can also construct estimators of the regression error variance  $\sigma_\epsilon^2$ . For the undisturbed data  $(x_i, y_i)$ , the estimator is

$$\hat{\sigma}_\epsilon^2 = S_{yy} - S_{xy} \hat{\beta}.$$

For the disturbed data  $(x_i^a, y_i^a)$ , the naive estimator is

$$\hat{\sigma}_\epsilon^{a2} = S_{yy}^a - S_{xy}^a \hat{\beta}^a.$$

Using (8) and (10), we find the corrected estimator for the additive case:

$$\hat{\sigma}_\epsilon^{c2} = S_{yy}^a - (1 - \frac{1}{T}) \sigma_v^2 - (S_{xy}^a - (1 - \frac{1}{T}) \sigma_{uv}) \hat{\beta}^c,$$

and using (13) and (15) we have for the multiplicative case:

$$\hat{\sigma}_\epsilon^{c2} = S_{yy}^a - (1 - \frac{1}{T}) \sigma_v^2 \hat{\mathbb{E}} m_{yy} - (S_{xy}^a - (1 - \frac{1}{T}) \sigma_{uv} \hat{\mathbb{E}} m_{xy}) \hat{\beta}^c,$$

where

$$\hat{\mathbb{E}} m_{yy} = \frac{\overline{m_{yy}^a}}{(1 + \sigma_v^2)}, \quad m_{yy}^a = (m_{yy}^a)_i = \frac{1}{T} \sum_{t=1}^T y_{it}^{a2}.$$

## 7 Common factor in the error process

### 7.1 The model

Höhne (2008) suggests a different kind of (additive or multiplicative) random noise. It is characterized by having a common factor structure:

$$\begin{aligned} u_{it} &= d_i + u_{it}^* \\ v_{it} &= e_i + v_{it}^*, \end{aligned} \quad (26)$$

where  $d_i$  and  $e_i$  are random variables with  $\mathbb{E} d_i = \mathbb{E} e_i = 0$  and finite variances.

In studying disclosure control techniques, H hne assumes a very special distribution for these variables, e.g.,  $d_i$  is of the form  $d_i = \delta D_i$ , where  $D_i$  is a variable which takes the values 1 and  $-1$ , each with probability  $\frac{1}{2}$ , and  $\delta$  is a known constant. Moreover, he sets  $d_i = e_i$ , which is motivated by the idea that then ratios of variables such as  $x/y$  are only slightly affected by the random noise, see Ronning (2009). We do not make these special assumptions, but, of course, this case is covered by our general approach. However, we do require independence and iid assumptions for these new error terms similar to those of Section 2.2.

*Independence assumptions:* As before, the set of pairs  $(u_{it}, v_{it})$  is independent of the set of pairs  $(x_{it}, y_{it})$ . In addition, the set of pairs  $(e_i, d_i)$  is independent of the set of pairs  $(u_{it}^*, v_{it}^*)$ .

*Iid assumptions:* The pairs  $(u_{it}^*, v_{it}^*)$  and the pairs  $(d_i, e_i)$  are iid.

These assumptions differ from the corresponding ones in Section 2.2 in that the  $u_{it}$  and  $v_{it}$  are no more iid, but are serially correlated. Instead, the  $u_{it}^*$  and  $v_{it}^*$  are now iid. This has consequences for the derivation of bias and correction formulas.

As to the relation between  $d$  and  $e$ , there are two main cases considered in the literature. They can be independent or they can be equal.

There is also the possibility that we have iid pairs  $(d_{it}, e_{it})$  instead of  $(d_i, e_i)$ . In this case, the previous theory goes through unchanged.

## 7.2 Additive errors

In the additive case, the within variances and covariances depend on the error terms  $u_{it}$  and  $v_{it}$  only through the differences  $u_{it} - \bar{u}_i$  and  $v_{it} - \bar{v}_i$ , see (6). But since

$$\begin{aligned} u_{it} - \bar{u}_i &= u_{it}^* - \bar{u}_i^* \\ v_{it} - \bar{v}_i &= v_{it}^* - \bar{v}_i^*, \end{aligned}$$

we get the the same results as before except that  $u_{it}$  and  $v_{it}$  have to be replaced with  $u_{it}^*$  and  $v_{it}^*$ , respectively.

Thus the *bias* is given by

$$\text{plim} \hat{\beta}^a = \frac{\beta \mathbb{E} s_{xx} + (1 - \frac{1}{T}) \sigma_{u^* v^*}}{\mathbb{E} s_{xx} + (1 - \frac{1}{T}) \sigma_{u^*}^2},$$

and the *corrected* estimator of  $\beta$  is

$$\hat{\beta}^c = \frac{S_{xy}^a - (1 - \frac{1}{T}) \sigma_{u^* v^*}}{S_{xx}^a - (1 - \frac{1}{T}) \sigma_{u^*}^2}.$$

The corrected estimator of  $\sigma_\epsilon^2$  is given by

$$\hat{\sigma}_\epsilon^2 = S_{yy}^a - (1 - \frac{1}{T})\sigma_{v^*}^2 - (S_{xy}^a - (1 - \frac{1}{T})\sigma_{u^*v^*})\hat{\beta}^c.$$

### 7.3 Multiplicative errors

In the multiplicative case, we need to proceed more carefully. We start with the probability limits of the within variances and covariances as in Section 4.1. The general formula (11) for the probability limit of  $S_{xy}^a$  is still true and, again because of the independence assumption,  $\mathbb{E}s_{x(yv)} = \mathbb{E}s_{y(xu)} = 0$ . However,  $\mathbb{E}s_{(xu)(yv)}$  is different. First note that

$$\begin{aligned} x_{it}u_{it} - (\overline{xu})_i &= d_i(x_{it} - \bar{x}_i) + x_{it}u_{it}^* - (\overline{xu^*})_i \\ y_{it}v_{it} - (\overline{yv})_i &= e_i(y_{it} - \bar{y}_i) + y_{it}v_{it}^* - (\overline{yv^*})_i, \end{aligned}$$

where, e.g.,  $(\overline{xu})_i = \frac{1}{T} \sum_t x_{it}u_{it}$ . It follows that

$$\begin{aligned} \mathbb{E}s_{(xu)(yv)} &= \mathbb{E} \frac{1}{T} \sum_t (x_{it}u_{it} - (\overline{xu})_i)(y_{it}v_{it} - (\overline{yv})_i) \\ &= \mathbb{E} \left[ d_i e_i \frac{1}{T} \sum_t (x_{it} - \bar{x}_i)(y_{it} - \bar{y}_i) \right. \\ &\quad + d_i \frac{1}{T} \sum_t (x_{it} - \bar{x}_i)(y_{it}v_{it}^* - (\overline{yv^*})_i) \\ &\quad + e_i \frac{1}{T} \sum_t (y_{it} - \bar{y}_i)(x_{it}u_{it}^* - (\overline{xu^*})_i) \\ &\quad \left. + \frac{1}{T} \sum_t (x_{it}u_{it}^* - (\overline{xu^*})_i)(y_{it}v_{it}^* - (\overline{yv^*})_i) \right] \\ &= \mathbb{E}(de)\mathbb{E}s_{xy} + \mathbb{E}s_{(xu^*)(yv^*)} \\ &= \sigma_{de}\mathbb{E}s_{xy} + (1 - \frac{1}{T})\sigma_{u^*v^*}\mathbb{E}m_{xy}. \end{aligned}$$

Therefore,

$$\text{plim}(S_{xy}^a - S_{xy}) = \sigma_{de}\mathbb{E}s_{xy} + (1 - \frac{1}{T})\sigma_{u^*v^*}\mathbb{E}m_{xy}. \quad (27)$$

Similarly with  $(x, u, d)$  in place of  $(y, v, e)$ ,

$$\text{plim}(S_{xx}^a - S_{xx}) = \sigma_d^2\mathbb{E}s_{xx} + (1 - \frac{1}{T})\sigma_{u^*}^2\mathbb{E}m_{xx}. \quad (28)$$

(27) and (28) imply the following *bias* formula

$$\begin{aligned} \text{plim}\hat{\beta}^a &= \frac{\text{plim}S_{xy} + \sigma_{de}\mathbb{E}s_{xy} + (1 - \frac{1}{T})\sigma_{u^*v^*}\mathbb{E}m_{xy}}{\text{plim}S_{xx} + \sigma_d^2\mathbb{E}s_{xx} + (1 - \frac{1}{T})\sigma_{u^*}^2\mathbb{E}m_{xx}} \\ &= \frac{\beta(1 + \sigma_{de})\mathbb{E}s_{xx} + (1 - \frac{1}{T})\sigma_{u^*v^*}\mathbb{E}m_{xy}}{(1 + \sigma_d^2)\mathbb{E}s_{xx} + (1 - \frac{1}{T})\sigma_{u^*}^2\mathbb{E}m_{xx}}. \end{aligned} \quad (29)$$

A *bias correction* for estimating  $\beta$  is found by evaluating

$$\hat{\beta}^c = \frac{S_{xy}^c}{S_{xx}^c}. \quad (30)$$

We find an estimator for  $S_{xy}$  from (27) as follows:

$$S_{xy}^c = S_{xy}^a - S_{xy}^c \sigma_{de} - (1 - \frac{1}{T}) \sigma_{u^*v^*} \hat{\mathbb{E}} m_{xy},$$

where  $\hat{\mathbb{E}} m_{xy}$  is a suitable estimator of  $\mathbb{E} m_{xy}$ , and so

$$S_{xy}^c = \frac{S_{xy}^a - (1 - \frac{1}{T}) \sigma_{u^*v^*} \hat{\mathbb{E}} m_{xy}}{1 + \sigma_{de}}. \quad (31)$$

Similarly,

$$S_{xx}^c = \frac{S_{xx}^a - (1 - \frac{1}{T}) \sigma_{u^*}^2 \hat{\mathbb{E}} m_{xx}}{1 + \sigma_d^2} \quad (32)$$

Substituting (31) and (32) in (30) results in

$$\hat{\beta}^c = \frac{S_{xy}^a - (1 - \frac{1}{T}) \sigma_{u^*v^*} \hat{\mathbb{E}} m_{xy}}{S_{xx}^a - (1 - \frac{1}{T}) \sigma_{u^*}^2 \hat{\mathbb{E}} m_{xx}} \cdot \frac{1 + \sigma_d^2}{1 + \sigma_{de}}. \quad (33)$$

It remains to construct estimators of  $\mathbb{E} m_{xy}$  and  $\mathbb{E} m_{xx}$ . We start with

$$\begin{aligned} \overline{m_{xy}^a} &:= \frac{1}{N} \frac{1}{T} \sum_i \sum_t x_{it}^a y_{it}^a \\ &= \frac{1}{N} \sum_i \frac{1}{T} \sum_t x_{it} y_{it} (1 + d_i + u_{it}^*) (1 + e_i + v_{it}^*) \\ &= \frac{1}{N} \sum_i \frac{1}{T} \sum_t x_{it} y_{it} (1 + d_i + e_i + d_i e_i + u_{it}^* + v_{it}^* + d_i v_{it}^* + e_i u_{it}^* + u_{it}^* v_{it}^*), \end{aligned}$$

which converges in probability to  $\mathbb{E} m_{xy} (1 + \sigma_{ed} + \sigma_{u^*v^*})$ . An estimator of  $\mathbb{E} m_{xy}$  is thus given by

$$\hat{\mathbb{E}} m_{xy} = \frac{\overline{m_{xy}^a}}{1 + \sigma_{ed} + \sigma_{u^*v^*}}. \quad (34)$$

Similarly,

$$\hat{\mathbb{E}} m_{xx} = \frac{\overline{m_{xx}^a}}{1 + \sigma_d^2 + \sigma_{u^*}^2}. \quad (35)$$

The corrected estimator of  $\sigma_\epsilon^2$  is given by

$$\hat{\sigma}_\epsilon^2 = \frac{S_{yy}^a - (1 - \frac{1}{T}) \sigma_{v^*}^2 \hat{\mathbb{E}} m_{yy}}{1 + \sigma_\epsilon^2} - \frac{S_{xy}^a - (1 - \frac{1}{T}) \sigma_{u^*v^*} \hat{\mathbb{E}} m_{xy}}{1 + \sigma_{de}} \hat{\beta}^c,$$

where

$$\hat{\mathbb{E}} m_{yy} = \frac{\overline{m_{yy}^a}}{1 + \sigma_\epsilon^2 + \sigma_{v^*}^2}.$$

## 8 Asymptotic variances

### 8.1 Estimating equations

The asymptotic variances of the various estimators can be found from the general theory of unbiased estimating equations. Let  $z_i, i = 1, \dots, N$ , be a set of iid vector-valued random variables and  $\theta$  a  $p$ -dimensional parameter vector pertaining to the distribution of  $z$ . Suppose a  $p$ -dimensional vector-valued function  $\psi$  of  $z$  and  $\theta$  exists such that  $\mathbb{E}\psi(z, \theta) = 0$ . We call  $\psi$  an unbiased estimating function and the equation

$$\sum_{i=1}^N \psi(z_i, \hat{\theta}) = 0$$

an unbiased estimating equation, Heyde (1997). Under general conditions, the solution  $\hat{\theta}$  to this equation exists uniquely, at least for large  $N$ , and is a consistent estimator of  $\theta$ . In addition,  $\hat{\theta}$  is asymptotically normal with an asymptotic covariance matrix given by the sandwich formula

$$\mathbb{V}(\hat{\theta}) = \frac{1}{N} (\mathbb{E}\psi_{\theta})^{-1} \mathbb{E}\psi\psi^{\top} \mathbb{E}(\psi_{\theta})^{-\top},$$

where

$$\psi_{\theta} := \frac{\partial \psi(z, \theta)}{\partial \theta^{\top}},$$

which is a  $(p \times p)$ -matrix. A consistent estimator of the asymptotic covariance matrix is given by

$$\hat{\mathbb{V}}(\hat{\theta}) = \frac{1}{N} (\overline{\psi_{\theta}})^{-1} \overline{\psi\psi^{\top}} (\overline{\psi_{\theta}})^{-\top},$$

where

$$\overline{\psi_{\theta}} = \frac{1}{N} \sum_i \psi_{\theta}(z_i, \hat{\theta}) \tag{36}$$

$$\overline{\psi\psi^{\top}} = \frac{1}{N} \sum_i \psi(z_i, \hat{\theta}) \psi(z_i, \hat{\theta})^{\top} = \frac{1}{N} \Psi^{\top} \Psi, \tag{37}$$

where  $\Psi^{\top}$  is the matrix  $(\psi(z_1, \hat{\theta}), \dots, \psi(z_N, \hat{\theta}))$ .

In the context of our panel model,  $z_i = (x_{i1}, \dots, x_{iT}, y_{i1}, \dots, y_{iT})$  is the data vector for individual  $i$  of the panel population. The parameter vector  $\theta$  has  $\beta$  in the first component as our parameter of interest and in addition possibly other (nuisance) parameters in the remaining components. The asymptotic variance of  $\hat{\beta}$  is then given by

$$\hat{\mathbb{V}}(\hat{\beta}) = e_1^{\top} \hat{\mathbb{V}}(\hat{\theta}) e_1,$$



where  $e_1 = (1, 0, \dots, 0)^\top$  is the  $p$ -dimensional first unit vector.

We apply this theory to the various error corrected estimators of our panel model by constructing an estimating function for each of the estimators.

## 8.2 Additive errors

From the corrected estimator of  $\beta$  (19), we obtain the estimating function

$$\psi = \psi(x, y, \beta) = (s_{xx}^a - (1 - \frac{1}{T})\sigma_u^2)\beta - (s_{xy}^a - (1 - \frac{1}{T})\sigma_{uv})$$

and

$$\psi_\beta = s_{xx}^a - (1 - \frac{1}{T})\sigma_u^2.$$

In this case,  $\psi$  is a scalar function, and so  $\psi\psi^\top = \psi^2$ .

## 8.3 Multiplicative errors

In the case of multiplicative errors,  $\theta = (\beta, \mathbb{E}m_{xx}, \mathbb{E}m_{xy})^\top$ , and the estimating function for  $\theta$ , as implied by (18), (20), and (21), is given by a vector  $\psi = (\psi_1, \psi_2, \psi_3)^\top$  with

$$\begin{aligned}\psi_1 &= [s_{xx}^a - (1 - \frac{1}{T})\sigma_u^2 \mathbb{E}m_{xx}]\beta - [s_{xy}^a - (1 - \frac{1}{T})\sigma_{uv} \mathbb{E}m_{xy}] \\ \psi_2 &= (1 + \sigma_u^2)\mathbb{E}m_{xx} - m_{xx}^a \\ \psi_3 &= (1 + \sigma_{uv})\mathbb{E}m_{xy} - m_{xy}^a.\end{aligned}$$

For  $\psi_\theta$ , we find

$$\psi_\theta = \begin{pmatrix} s_{xx}^a - (1 - \frac{1}{T})\sigma_u^2 \mathbb{E}m_{xx} & -(1 - \frac{1}{T})\sigma_u^2 \beta & (1 - \frac{1}{T})\sigma_{uv} \\ 0 & 1 + \sigma_u^2 & 0 \\ 0 & 0 & 1 + \sigma_{uv} \end{pmatrix}.$$

In case  $\sigma_{uv} = 0$ ,  $\psi_3$  and the last row and last column of  $\psi_\theta$  can be dropped.

## 8.4 Common factor structure

If the measurement errors have a common factor structure, the estimators of Section 7 are relevant.

For *additive* errors, we have, according to Section 7.2, the same estimating function as in Subsection 8.2 except that  $\sigma_u^2$  and  $\sigma_{uv}$  have to be replaced with  $\sigma_{u^*}^2$  and  $\sigma_{u^*v^*}$ , respectively. Thus

$$\psi = \psi(x, y, \beta) = (s_{xx}^a - (1 - \frac{1}{T})\sigma_{u^*}^2)\beta - (s_{xy}^a - (1 - \frac{1}{T})\sigma_{u^*v^*}).$$

For *multiplicative* errors, we find the estimating function from (33). We have  $\theta = (\beta, \mathbb{E}m_{xx}, \mathbb{E}\overline{xy})^\top$  and  $\psi = (\psi_1, \psi_2, \psi_3)^\top$  with

$$\begin{aligned}\psi_1 &= [s_{xx}^a - (1 - \frac{1}{T})\sigma_{u^*}^2 \mathbb{E}m_{xx}]\beta - [s_{xy}^a - (1 - \frac{1}{T})\sigma_{u^*v^*} \mathbb{E}m_{xy}] \frac{1 + \sigma_d^2}{1 + \sigma_{de}} \\ \psi_2 &= (1 + \sigma_d^2 + \sigma_{u^*}^2) \mathbb{E}m_{xx} - m_{xx}^a \\ \psi_3 &= (1 + \sigma_{de} + \sigma_{u^*v^*}) \mathbb{E}m_{xy} - m_{xy}^a.\end{aligned}$$

For  $\psi_\theta$ , we find

$$\psi_\theta = \begin{pmatrix} s_{xx}^a - (1 - \frac{1}{T})\sigma_{u^*}^2 \mathbb{E}m_{xx} & -(1 - \frac{1}{T})\sigma_{u^*}^2 \beta & (1 - \frac{1}{T})\sigma_{u^*v^*} \frac{1 + \sigma_d^2}{1 + \sigma_{de}} \\ 0 & 1 + \sigma_d^2 + \sigma_{u^*}^2 & 0 \\ 0 & 0 & 1 + \sigma_{de} + \sigma_{u^*v^*} \end{pmatrix}.$$

Again, if  $\sigma_{u^*v^*} = 0$ ,  $\psi_3$  and the last row and last column of  $\psi_\theta$  can be dropped.

It may be noted that in all the above cases the asymptotic variance of  $\hat{\beta}$  can also be computed by the delta method. An example is given in the appendix.

## 9 Simulation

In our simulation study we analyze the performance of both the naive estimator (4) and the corrected estimators when both  $x$  and  $y$  are observed with multiplicative measurement errors. We do not present results for the more familiar case of additive errors.

For the iid case, as given in (3), the corrected estimator is given in (20) and for the (multiplicative) common factor model (26) the corrected estimator is given in (33). In all scenarios we used 2000 replications.

For the regressor variables we assumed the stationary AR(1) process

$$x_{it} = \phi + \varrho x_{i,t-1} + \omega_{it} \quad (38)$$

with  $|\varrho| < 1$  and  $\mathbb{E}\omega_{it} = 0, \mathbb{V}\omega_{it} = \sigma_\omega^2$  for all  $i = 1, \dots, N$  and  $t = 1, \dots, T$ . We use

$$\phi = (1 - \varrho) \mathbb{E}x_{it}, \sigma_\omega^2 = \mathbb{V}x_{it}(1 - \varrho^2),$$

so that for each  $\varrho$  (see below) we have the same expected value and variance of all  $x_{it}$ .

In order to study the effect of correlation of the individual effect  $\gamma$  with the regressor  $x$  we assume that the effect is given by

$$\gamma_i = [\bar{x}_i - \mathbb{E}\bar{x}_i] \lambda + w_i,$$

where  $w_i$  is normal white noise with expectation 0 and variance  $\sigma_w^2$  distributed independently of  $x$ . The variance of  $w_i$  is defined by

$$\sigma_w^2 = \mathbb{V}\gamma_i - \lambda^2 \sigma_{\bar{x}}^2,$$

so that the variance of the individual component is kept constant for any  $\lambda$  (see below). This specification of correlated individual effects has been proposed by Biørn (1996) p. 260. Note that, for positive  $\lambda$ , the correlation of  $\gamma_i$  with  $\bar{x}_i$  is given by

$$\varrho_{\bar{x}\gamma} = \frac{1}{\sqrt{1 + \frac{\sigma_w^2}{\lambda^2 \sigma_{\bar{x}}^2}}},$$

where

$$\sigma_{\bar{x}}^2 = \frac{\sigma_x^2}{T^2} \{T + 2(T-1)\rho^1 + 2(T-2)\rho^2 + \dots + 2 \cdot 2\rho^{T-2} + 2 \cdot 1\rho^{T-1}\}$$

under the autoregressive scheme (38), Hamilton (1994). This correlation will tend to 1 for  $\lambda \rightarrow \infty$  and to 0 for  $T \rightarrow \infty$ .

In our simulations we fixed  $\varrho_{\bar{x}\gamma}$  and  $\sigma_{\gamma}^2$  and derived the two remaining parameters as follows:

$$\lambda = \frac{\varrho_{\bar{x}\gamma} \sigma_{\gamma}}{\sigma_{\bar{x}}},$$

$$\sigma_w^2 = \sigma_{\gamma}^2 (1 - \varrho_{\bar{x}\gamma}^2).$$

When studying the common factor structure we use the special specification of Hühne(2008), which we already mentioned in section 7.1: We set  $d_i = e_i$  in (26) and use the special structure

$$e_i = \delta D \quad \text{with} \quad D = \begin{cases} 1 & \text{with probability 0.5} \\ -1 & \text{with probability 0.5} \end{cases}.$$

This special specification implies  $\sigma_d^2 = \sigma_{de}$  so that the corrected estimator from (33) is given by

$$\hat{\beta}^c = \frac{S_{xy}^a - (1 - \frac{1}{T})\sigma_{u^*v^*}\hat{\mathbb{E}}\overline{xy}}{S_{xx}^a - (1 - \frac{1}{T})\sigma_{u^*}^2\hat{\mathbb{E}}\overline{x^2}}.$$

The following parameters were fixed throughout the whole simulation study: In the linear panel model (1) we used

$$\beta = 1, \mathbb{E}x_{it} = 2, \mathbb{V}x_{it} = 1.5^2, \mathbb{V}\varepsilon_{it} = 0.5^2, \mathbb{V}\gamma_i = 1,$$

and for the measurement errors we used

$$\mathbb{V}u_{it} = 0.2^2, \mathbb{V}v_{it} = 0.2^2$$

for the iid case (3) and

$$\delta = 0.14, \mathbb{V}u_{it}^* = 0.14^2, \mathbb{V}v_{it}^* = 0.14^2$$

for the common factor case (3). Note that the total variance of measurement error in the latter case is given by  $\mathbb{V}u_{it} = \delta^2 + \mathbb{V}u_{it}^* = 0.14^2 + 0.14^2 = 0.0392$ , which is (almost) equal to the variance in case of (3).

For both the iid case and the common factor case, we studied the effects of varying the sample size  $N$ , the number of waves  $T$ , the autoregressive parameter  $\varrho$ , the correlation between  $u$  and  $v$ , which we denote by  $\rho_{uv}$  ( $\rho_{u^*v^*}$  for the common factor model), and the correlation between  $\gamma$  and  $\bar{x}$ . Table 1 has the details. We use alternatively a "moderate" and a large sample size. The number of waves is kept very small. The autoregressive parameter specifies both positive and negative autocorrelation besides the case of no autocorrelation. Additionally, we consider zero and non-zero correlation between  $u$  and  $v$  ( $u^*$  and  $v^*$  for the common factor model). Finally, a non-zero parameter  $\varrho_{\bar{x}\gamma}$  indicates correlation of individual effects with the regressor  $x$ .

Table 1: Parameter variation in the simulation study

parameter	values used
$N$	100 ; 1000
$T$	3 ; 10
$\varrho$	-0.5 ; 0 ; +0.5
$\rho_{uv}(\rho_{u^*v^*})$	-0.9 ; 0 ; +0.9
$\varrho_{\bar{x}\gamma}$	0 ; 0.975

The four tables contain the simulation results for the iid case (tables 2 and 3) and for the common factor case (tables 4 and 5). For each case, the second table reports results concerning correlation of individual effects with the regressor. In all four tables, we use the following notation:  $\hat{\beta}$  and  $s_{\hat{\beta}}$  give mean and standard deviation of the 2000 replications concerning the naive estimator,  $\hat{\beta}^c$  and  $s_{\hat{\beta}^c}$  the corresponding results for the corrected estimator.  $\hat{\sigma}_{\hat{\beta}^c}$  denotes the estimate of the theoretical (asymptotic) standard deviation discussed in section 8, and  $s_{\hat{\sigma}_{\hat{\beta}^c}}$  reports the standard error of this estimate. Finally,  $q_{\hat{\beta}^c}^\alpha$  is the  $\alpha$ -quantile of the corrected estimator for three different levels of  $\alpha$ .

Tables 2 , 3, 4 and 5 about here

For large samples ( $N = 1000$ ), our simulations support our theoretical findings: The corrected slope estimator  $\hat{\beta}^c$  shows practically no bias, and the average estimate of the theoretical (asymptotic) standard deviation of the estimator  $\hat{\beta}^c$  corresponds very accurately to the empirical variance of the estimates  $\hat{\beta}^c$  in the simulation runs. The asymptotic results seem to apply almost as well to samples of small to medium size ( $N = 100$ ): the corrected slope estimator shows hardly any bias, and the theoretical standard deviation still corresponds rather closely to the empirical standard deviation. Of course, for smaller  $N$ , these standard deviations are (about three times) larger.

The simulations also highlight the considerable amount of bias in the uncorrected (naive) estimator of the slope parameter. The bias tends to increase for increasing  $\rho$  and for decreasing  $\rho_{uv}$  – the latter in accordance with (17). The bias is considerably smaller for errors with a common factor structure, which is plausible considering the fact that, by using inner variances and covariances for constructing the estimator, the common factor is largely eliminated – it is completely eliminated in the additive case – so that the, much smaller, remaining error components  $u^*$  and  $v^*$  are now relevant for the bias. The presence of a medium correlation between individual effect and regressor ( $\varrho_{\bar{x}\gamma} = 0.5$ ) has almost no effect on the bias (not shown in the tables). But for  $\varrho_{\bar{x}\gamma} = 0.975$ , the bias is clearly smaller than in the case of no correlation if  $\sigma_{uv} \neq 0$ . This is in agreement with the theoretical result (17) on the bias.

The standard deviation of the corrected estimator can also be seen to depend on the various model parameters. It decreases for increasing  $T$ , decreasing  $\rho$ , and increasing  $\rho_{uv}$ . It is a good deal smaller in the common factor case. The dependence on  $\lambda$  is negligible. The standard deviation of the corrected estimator is, of course, larger than for the naive estimator, but not very much. The increase in variance is outweighed by the elimination of bias. Finally, it may be noted that the estimate of the standard deviation is very precise in view of its own standard deviation, in particular for large  $N$ .

## 10 Conclusion

Measurement errors in a linear regression result in biased estimates of the slope parameter when Least Squares is applied without regard to the measurement errors. This is true for panel data models just as for cross sectional models, except that, in panel models, within LS estimators instead of the ordinary LS estimators are used in order to get rid of the unobservable individual effects. As a result of this difference, a degree of freedoms factor enters the bias formula, which

is not present in the well-known bias formula for cross sectional data. When correcting for the bias in panel data models, this factor has again to be taken into account.

We focus our investigation on multiplicative iid. errors. They can be treated in a similar way as the more conventional additive errors, but with some characteristic differences. In the bias formula as well as in the expression for the bias corrected estimator, nuisance parameters appear, which have to be estimated, too. Their presence results in a substantially more complicated computation of the asymptotic variance of the slope estimator than in the additive case. The variance is computed with the help of the sandwich formula, which, however, has to take the nuisance parameters into account.

The results for iid errors can be generalized to the case of a common factor structure in the error process. Again, bias corrected estimators and asymptotic variances can be derived both for the additive and for the multiplicative case.

An extensive simulation study was carried out. It fully corroborates our theoretical findings on the asymptotics of our estimators and shows that the asymptotic results seem to apply almost as well to samples of small to medium size ( $N = 100$ ). The simulations also make evident the dependence of the asymptotic variance on the various model parameters, e.g., on the autocorrelation of the regressor variable or on the correlation of regressor variable and individual effect variable.

Finally, they show how close the asymptotic variance of the corrected estimator may come to that of the uncorrected estimator, at least for large  $N$ . Thus the correction is fully justified both on the ground that it eliminates the bias and that it implies only a small increase in variance.

## 11 Literature

- Biørn, E. (1996): Panel data with measurement errors, in: Mátyás, L. and P. Sevestre (eds.): *The Econometrics of Panel Data: A Handbook of the Theory with Applications*, 2nd. ed., Kluwer: Dordrecht, 236-279.
- Biørn, E. and Krishnakumar, J. (2008): Measurement errors and simultaneity. Chapter 10 in: Mátyás, L. and P. Sevestre (eds.): *The Econometrics of Panel Data*, 3rd. ed., Springer: Heidelberg, 323 - 367.
- Carroll, R.J., Ruppert, D., Stefanski, L.A., and Crainiceanu, C.M. (2006): *Measurement Error in Nonlinear Models*, Chapman and Hall: London.

- Domingo-Ferrer, J. and Y. Saygin (2008) (eds): *Privacy in Statistical Databases*. Unesco Chair in Data Privacy International Conference, PSD 2008. Istanbul, Turkey, September 2008, Proceedings. Springer: Berlin.
- Griliches, Z. and Hausman, J.A. (1986): Errors in variables in panel data, *Journal of Econometrics* **31**, 93-118.
- Hwang, J.T. (1986): Multiplicative errors-in-variables models with applications to recent data released by the U.S. Department of Energy. *Journal of the American Statistical Association* **81**, 680-688.
- Kim, J.J. (1986): A Method For Limiting Disclosure in Microdata Based on Random Noise and Transformation, *Proceedings of the Survey Research Methods Section*, American Statistical Association, 370-374.
- Lin, A. (1989): Estimation of multiplicative measurement error models and some simulation results. *Economics Letters* **31**, 13-20.
- Hamilton, J.D. (1994): *Time Series Analysis*, Princeton University Press: Princeton, NJ.
- Heyde C.C. (1997): *Quasi-Likelihood and Its Application*, Springer: Berlin, Heidelberg, New York.
- Höhne, J. (2008). Anonymisierungsverfahren für Paneldaten. *Wirtschafts- und Sozialstatistisches Archiv* **2**, 259-275.
- Hsiao, C. and Taylor, G. (1991): Some remarks on measurement errors and the identification of panel data models, *Statistica Neerlandica* **45**, 187-194.
- Ronning, G. (2007). *Stochastische Überlagerung mit Hilfe der Mischungsverteilung*. IAW Discussion Paper 30 (April 2007, revised October 2007). <http://www.iaw.edu/pdf/dp2007-30.pdf>
- Ronning, G. (2009). Measuring Research Intensity From Anonymized Data: Does Multiplicative Noise With Factor Structure Save Results Regarding Quotients? *Jahrbücher für Nationalökonomie und Statistik* **228**, (to appear).
- Wansbeek, T.J. (2001): GMM estimation in panel data models with measurement errors, *Journal of Econometrics* **104**, 259-268.
- Wansbeek, T.J. and Koning, R.H. (1991): Measurement error and panel data, *Statistica Neerlandica* **45**, 85-92.

## Appendix: Equivalence of delta method and sandwich formula

The delta method and the sandwich formula are two methods to estimate the asymptotic variance of  $\hat{\beta}^c$ . We show their equivalence by way of example in a special case. Consider the corrected estimator (20) of  $\beta$  in a model with multiplicative iid. measurement errors for the special case  $\sigma_{uv} = 0$ :

$$\hat{\beta}^c = \frac{\overline{s_{xy}^a}}{\overline{s_{xx}^a} - q\overline{m_{xx}^a}}, \quad q := \frac{T-1}{T} \frac{\sigma_u^2}{1 + \sigma_u^2}.$$

By the delta method, an estimate of the asymptotic variance of  $\hat{\beta}^c$  is given by

$$\hat{V}\hat{\beta}^c = \frac{\hat{\beta}^{c2}}{S_{xy}^{a2}} \begin{pmatrix} 1 & -\hat{\beta}^c & q\hat{\beta}^c \end{pmatrix} \begin{pmatrix} v_{11} & v_{12} & v_{13} \\ v_{21} & v_{22} & v_{23} \\ v_{31} & v_{32} & v_{33} \end{pmatrix} \begin{pmatrix} 1 \\ -\hat{\beta}^c \\ q\hat{\beta}^c \end{pmatrix},$$

where

$$\begin{aligned} v_{11} &= \hat{V}(\overline{s_{xy}^a}) = \frac{1}{N}(\overline{s_{xy}^{a2}} - \overline{s_{xy}^a}^2) \\ v_{12} &= \hat{Cov}(\overline{s_{xy}^a}, \overline{s_{xx}^a}) = \frac{1}{N}(\overline{s_{xy}^a s_{xx}^a} - \overline{s_{xy}^a} \overline{s_{xx}^a}) \\ v_{13} &= \hat{Cov}(\overline{s_{xy}^a}, \overline{m_{xx}^a}) = \frac{1}{N}(\overline{s_{xy}^a m_{xx}^a} - \overline{s_{xy}^a} \overline{m_{xx}^a}) \\ v_{22} &= \hat{V}(\overline{s_{xx}^a}) = \frac{1}{N}(\overline{s_{xx}^{a2}} - \overline{s_{xx}^a}^2) \\ v_{23} &= \hat{Cov}(\overline{s_{xx}^a}, \overline{m_{xx}^a}) = \frac{1}{N}(\overline{s_{xx}^a m_{xx}^a} - \overline{s_{xx}^a} \overline{m_{xx}^a}) \\ v_{33} &= \hat{V}(\overline{m_{xx}^a}) = \frac{1}{N}(\overline{m_{xx}^{a2}} - \overline{m_{xx}^a}^2). \end{aligned}$$

This is the same as what we would get by the sandwich formula. To see this, write the estimated vector  $\psi$  for individual  $i$  as

$$\hat{\psi}_i = \begin{pmatrix} [(s_{xx}^a)_i - q\overline{m_{xx}^a}]\hat{\beta}^c & -(s_{xy}^a)_i \\ \overline{m_{xx}^a} - (m_{xx}^a)_i \end{pmatrix},$$

from which the inner part of the sandwich follows as

$$\widehat{\psi\psi'} = \begin{pmatrix} \frac{[(s_{xx}^a - q\overline{m_{xx}^a})\hat{\beta}^c - s_{xy}^a]^2}{[(s_{xx}^a - q\overline{m_{xx}^a})\hat{\beta}^c - s_{xy}^a][\overline{m_{xx}^a} - m_{xx}^a]} & \frac{[(s_{xx}^a - q\overline{m_{xx}^a})\hat{\beta}^c - s_{xy}^a][\overline{m_{xx}^a} - m_{xx}^a]}{[\overline{m_{xx}^a} - m_{xx}^a]^2} \end{pmatrix}.$$

Furthermore,

$$(\hat{\psi}_\theta)_i = \begin{pmatrix} (s_{xx}^a)_i - q\overline{m_{xx}^a} & -(1 - \frac{1}{T})\sigma_u^2\hat{\beta}^c \\ 0 & 1 + \sigma_u^2 \end{pmatrix}$$



and

$$\overline{\hat{\psi}_\theta} = \begin{pmatrix} \overline{s_{xx}^a} - q\overline{m_{xx}^a} & -(1 - \frac{1}{T})\sigma_u^2\hat{\beta}^c \\ 0 & 1 + \sigma_u^2 \end{pmatrix} = \begin{pmatrix} \frac{s_{xy}^a}{\hat{\beta}^c} & -(1 - \frac{1}{T})\sigma_u^2\hat{\beta}^c \\ 0 & 1 + \sigma_u^2 \end{pmatrix}.$$

It follows that

$$\overline{\hat{\psi}_\theta}^{-1} = \frac{\hat{\beta}^c}{s_{xy}^a(1 + \sigma_u^2)} \begin{pmatrix} 1 + \sigma_u^2 & (1 - \frac{1}{T})\sigma_u^2\hat{\beta}^c \\ 0 & \frac{s_{xy}^a}{\hat{\beta}^c} \end{pmatrix}$$

With the first unit vector  $e' := (1, 0)$ , we get

$$e'\overline{\hat{\psi}_\theta}^{-1} = \frac{\hat{\beta}^c}{s_{xy}^a} \begin{pmatrix} 1 & q\hat{\beta}^c \end{pmatrix}.$$

We thus have all the necessary parts to construct the estimate of the asymptotic variance of  $\hat{\beta}^c$ :

$$\hat{\mathbb{V}}(\hat{\beta}^c) = \frac{1}{N} e'\overline{\hat{\psi}_\theta}^{-1} \overline{\hat{\psi}\hat{\psi}'} (e'\overline{\hat{\psi}_\theta}^{-1})'.$$

We now only have to further expand the inner part of the sandwich  $\overline{\hat{\psi}\hat{\psi}'}$ . The upper left corner is given by

$$\begin{aligned} & (\overline{s_{xx}^{a2}} - 2q\overline{s_{xx}^a m_{xx}^a} + q^2\overline{m_{xx}^a}^2)\hat{\beta}^{c2} - 2(\overline{s_{xx}^a s_{xy}^a} - q\overline{m_{xx}^a s_{xy}^a})\hat{\beta}^c + \overline{s_{xy}^{a2}} \\ &= (Nv_{22} + \overline{s_{xx}^a}^2 - 2q\overline{s_{xx}^a m_{xx}^a} + q^2\overline{m_{xx}^a}^2)\hat{\beta}^{c2} - 2(Nv_{12} + \overline{s_{xy}^a} \overline{s_{xx}^a} - q\overline{m_{xx}^a s_{xy}^a})\hat{\beta}^c + Nv_{11} + \overline{s_{xy}^a}^2 \\ &= Nv_{22}\hat{\beta}^{c2} + (\overline{s_{xx}^a} - q\overline{m_{xx}^a})^2\hat{\beta}^{c2} - 2Nv_{12}\hat{\beta}^c - 2(\overline{s_{xx}^a} - q\overline{m_{xx}^a})\overline{s_{xy}^a}\hat{\beta}^c + Nv_{11} + \overline{s_{xy}^a}^2 \\ &= Nv_{22}\hat{\beta}^{c2} + s_{xy}^{a2} - 2Nv_{12}\hat{\beta}^c - 2s_{xy}^{a2} + Nv_{11} + s_{xy}^{a2} \\ &= Nv_{22}\hat{\beta}^{c2} - 2Nv_{12}\hat{\beta}^c + Nv_{11} \\ &= N \begin{pmatrix} 1 & -\hat{\beta}^c \end{pmatrix} \begin{pmatrix} v_{11} & v_{12} \\ v_{21} & v_{22} \end{pmatrix} \begin{pmatrix} 1 \\ -\hat{\beta}^c \end{pmatrix}. \end{aligned}$$

Similarly, the right upper and left lower corners are

$$N \begin{pmatrix} v_{13} & v_{23} \end{pmatrix} \begin{pmatrix} 1 \\ -\hat{\beta}^c \end{pmatrix}.$$

Finally the right lower corner is  $Nv_{33}$ .

Collecting terms we have

$$\overline{\hat{\psi}\hat{\psi}'} = N \begin{pmatrix} (1 \quad -\hat{\beta}^c) \begin{pmatrix} v_{11} & v_{12} \\ v_{21} & v_{22} \end{pmatrix} \begin{pmatrix} 1 \\ -\hat{\beta}^c \end{pmatrix} & (v_{13} \quad v_{23}) \begin{pmatrix} 1 \\ -\hat{\beta}^c \end{pmatrix} \\ (1 \quad -\hat{\beta}^c) \begin{pmatrix} v_{13} \\ v_{23} \end{pmatrix} & v_{33} \end{pmatrix}$$

and

$$\begin{aligned}
\hat{\mathbb{V}}(\hat{\beta}^c) &= \frac{\hat{\beta}^{c2}}{S_{xy}^2} (1 \quad q\hat{\beta}^c) \begin{pmatrix} (1 \quad -\hat{\beta}^c) \begin{pmatrix} v_{11} & v_{12} \\ v_{21} & v_{22} \end{pmatrix} \begin{pmatrix} 1 \\ -\hat{\beta}^c \end{pmatrix} & (v_{13} \quad v_{23}) \begin{pmatrix} 1 \\ -\hat{\beta}^c \end{pmatrix} \\ (1 \quad -\hat{\beta}^c) \begin{pmatrix} v_{13} \\ v_{23} \end{pmatrix} & v_{33} \end{pmatrix} \begin{pmatrix} 1 \\ q\hat{\beta}^c \end{pmatrix} \\
&= \frac{\hat{\beta}^{c2}}{S_{xy}^2} (1 \quad -\hat{\beta}^c \quad q\hat{\beta}^c) \begin{pmatrix} v_{11} & v_{12} & v_{13} \\ v_{21} & v_{22} & v_{23} \\ v_{31} & v_{32} & v_{33} \end{pmatrix} \begin{pmatrix} 1 \\ -\hat{\beta}^c \\ q\hat{\beta}^c \end{pmatrix},
\end{aligned}$$

which is the variance formula according to the delta method.

Table 2: Simulation results for the iid case - uncorrelated individual effects

$\varrho$	$\rho_{uv}$	N	$\hat{\beta}$	$s_{\hat{\beta}}$	$\hat{\beta}^c$	$s_{\hat{\beta}^c}$	$\hat{\sigma}_{\hat{\beta}^c}$	$s_{\hat{\sigma}_{\hat{\beta}^c}}$	$q_{\hat{\beta}^c}^{0.05}$	$q_{\hat{\beta}^c}^{0.50}$	$q_{\hat{\beta}^c}^{0.95}$
T = 3											
-0.50	-0.90	100	0.84283	0.05066	0.99924	0.05891	0.05782	0.00872	0.90154	0.99830	1.09681
		1,000	0.84489	0.01648	0.99997	0.01889	0.01871	0.00100	0.96914	0.99956	1.03066
-0.50	0.00	100	0.91859	0.04105	1.00083	0.04540	0.04439	0.00637	0.92737	1.00093	1.07725
		1,000	0.91875	0.01339	1.00050	0.01475	0.01439	0.00070	0.97608	1.00041	1.02393
-0.50	0.90	100	0.99148	0.02554	0.99966	0.02693	0.02673	0.00354	0.95457	0.99940	1.04433
		1,000	0.99182	0.00800	0.99999	0.00843	0.00858	0.00036	0.98640	0.99997	1.01394
0.00	-0.90	100	0.81154	0.05494	1.00371	0.06589	0.06434	0.00956	0.89923	1.00156	1.11200
		1,000	0.80980	0.01765	1.00030	0.02094	0.02083	0.00112	0.96597	0.99979	1.03525
0.00	0.00	100	0.89991	0.04435	1.00107	0.05037	0.04927	0.00701	0.92314	0.99979	1.08438
		1,000	0.90033	0.01392	1.00044	0.01589	0.01592	0.00076	0.97440	1.00034	1.02637
0.00	0.90	100	0.98919	0.02832	0.99923	0.03047	0.03000	0.00371	0.95045	0.99839	1.04919
		1,000	0.99029	0.00883	1.00030	0.00946	0.00962	0.00039	0.98470	1.00054	1.01586
0.50	-0.90	100	0.69584	0.06767	1.00409	0.09125	0.08734	0.01473	0.86124	1.00040	1.16018
		1,000	0.69625	0.02224	1.00038	0.02893	0.02830	0.00166	0.95370	0.99998	1.04780
0.50	0.00	100	0.84122	0.05552	1.00288	0.06816	0.06568	0.00985	0.89316	1.00238	1.11608
		1,000	0.84001	0.01798	1.00043	0.02185	0.02133	0.00115	0.96501	1.00066	1.03558
0.50	0.90	100	0.98248	0.03604	0.99825	0.04094	0.04056	0.00539	0.93236	0.99821	1.06745
		1,000	0.98331	0.01117	0.99930	0.01280	0.01293	0.00056	0.97738	0.99917	1.02029
T = 10											
-0.50	-0.90	100	0.82120	0.02626	1.00055	0.03112	0.03101	0.00334	0.94925	1.00073	1.05115
		1,000	0.82125	0.00822	1.00010	0.00979	0.00991	0.00037	0.98374	1.00042	1.01594
-0.50	0.00	100	0.90569	0.02138	1.00000	0.02363	0.02341	0.00240	0.96126	0.99982	1.03805
		1,000	0.90610	0.00665	1.00034	0.00745	0.00753	0.00025	0.98817	1.00035	1.01254
-0.50	0.90	100	0.99015	0.01343	0.99960	0.01398	0.01391	0.00131	0.97677	0.99939	1.02260
		1,000	0.99062	0.00421	1.00001	0.00440	0.00443	0.00013	0.99263	1.00004	1.00721
0.00	-0.90	100	0.81098	0.02709	1.00155	0.03199	0.03193	0.00344	0.94922	1.00060	1.05344
		1,000	0.80996	0.00852	0.99984	0.01021	0.01022	0.00035	0.98369	0.99964	1.01726
0.00	0.00	100	0.90057	0.02235	1.00076	0.02511	0.02423	0.00235	0.95949	1.00064	1.04204
		1,000	0.89992	0.00694	0.99994	0.00781	0.00776	0.00025	0.98693	0.99970	1.01265
0.00	0.90	100	0.98946	0.01424	0.99945	0.01490	0.01441	0.00127	0.97498	0.99964	1.02414
		1,000	0.99005	0.00425	1.00010	0.00443	0.00459	0.00013	0.99278	1.00028	1.00704
0.50	-0.90	100	0.77289	0.03006	1.00004	0.03464	0.03513	0.00406	0.94285	0.99969	1.05504
		1,000	0.77404	0.00960	1.00014	0.01135	0.01127	0.00042	0.98176	1.00005	1.01912
0.50	0.00	100	0.88009	0.02384	0.99906	0.02697	0.02662	0.00286	0.95383	0.99827	1.04515
		1,000	0.88097	0.00761	1.00004	0.00866	0.00852	0.00029	0.98526	1.00003	1.01384
0.50	0.90	100	0.98794	0.01498	0.99993	0.01579	0.01593	0.00152	0.97297	0.99996	1.02554
		1,000	0.98815	0.00473	1.00006	0.00499	0.00508	0.00015	0.99181	1.00017	1.00841

Table 3: Simulation results for the iid case - correlated individual effects  
( $\varrho_{\bar{x}\gamma} = 0.975$ )

$\varrho$	$\rho_{uv}$	N	$\hat{\beta}$	$s_{\hat{\beta}}$	$\hat{\beta}^c$	$s_{\hat{\beta}^c}$	$\hat{\sigma}_{\hat{\beta}^c}$	$s_{\hat{\sigma}_{\hat{\beta}^c}}$	$q_{\hat{\beta}^c}^{0.05}$	$q_{\hat{\beta}^c}^{0.50}$	$q_{\hat{\beta}^c}^{0.95}$
T = 3											
-0.50	-0.90	100	0.91782	0.04274	1.00002	0.04607	0.04415	0.00662	0.92459	0.99971	1.07690
		1000	0.91896	0.01327	1.00041	0.01421	0.01426	0.00074	0.97694	1.00022	1.02420
-0.50	0.00	100	0.95886	0.03408	1.00023	0.03581	0.03543	0.00493	0.93990	0.99959	1.05895
		1000	0.95916	0.01047	1.00016	0.01100	0.01145	0.00055	0.98205	0.99998	1.01846
-0.50	0.90	100	0.99855	0.02441	0.99909	0.02508	0.02494	0.00324	0.95802	0.99823	1.04062
		1000	0.99947	0.00763	1.00003	0.00785	0.00802	0.00034	0.98723	0.99994	1.01283
0.00	-0.90	100	0.89798	0.04530	1.00163	0.04921	0.04965	0.00730	0.92211	1.00204	1.08206
		1000	0.89771	0.01391	1.00048	0.01514	0.01605	0.00079	0.97592	1.00017	1.02563
0.00	0.00	100	0.95034	0.03779	1.00142	0.04013	0.03985	0.00559	0.93552	1.00022	1.06809
		1000	0.94972	0.01195	1.00049	0.01266	0.01280	0.00061	0.97952	1.00040	1.02096
0.00	0.90	100	1.00093	0.02660	0.99975	0.02761	0.02786	0.00352	0.95343	0.99948	1.04466
		1000	1.00121	0.00842	1.00010	0.00874	0.00896	0.00037	0.98566	1.00017	1.01404
0.50	-0.90	100	0.82617	0.06098	1.00224	0.06854	0.06787	0.01118	0.88785	1.00266	1.11473
		1000	0.82648	0.01934	0.99981	0.02147	0.02184	0.00131	0.96506	0.99961	1.03423
0.50	0.00	100	0.91766	0.04810	1.00294	0.05279	0.05272	0.00754	0.91648	1.00248	1.09066
		1000	0.91644	0.01493	1.00057	0.01640	0.01700	0.00087	0.97272	1.00060	1.02776
0.50	0.90	100	1.00738	0.03376	1.00192	0.03639	0.03718	0.00478	0.94246	1.00035	1.06331
		1000	1.00589	0.01109	1.00046	0.01193	0.01193	0.00050	0.98080	1.00049	1.02032
T = 10											
-0.50	-0.90	100	0.90746	0.02165	1.00006	0.02321	0.02333	0.00253	0.96166	1.00008	1.03801
		1000	0.90774	0.00694	1.00009	0.00749	0.00747	0.00027	0.98814	1.00009	1.01199
-0.50	0.00	100	0.95225	0.01717	0.99991	0.01808	0.01857	0.00197	0.96915	1.00003	1.02980
		1000	0.95226	0.00564	0.99983	0.00595	0.00595	0.00020	0.99002	0.99981	1.00944
-0.50	0.90	100	0.99679	0.01262	0.99955	0.01297	0.01294	0.00125	0.97842	0.99915	1.02165
		1000	0.99728	0.00393	1.00009	0.00404	0.00414	0.00013	0.99342	1.00001	1.00678
0.00	-0.90	100	0.89985	0.02270	0.99969	0.02437	0.02439	0.00260	0.96076	0.99965	1.03900
		1000	0.90014	0.00741	0.99985	0.00792	0.00781	0.00027	0.98686	0.99987	1.01319
0.00	0.00	100	0.94907	0.01826	0.99986	0.01933	0.01923	0.00189	0.96943	0.99925	1.03314
		1000	0.94930	0.00580	1.00001	0.00617	0.00616	0.00020	0.98972	1.00004	1.01010
0.00	0.90	100	0.99849	0.01258	1.00025	0.01298	0.01343	0.00127	0.97862	1.00021	1.02179
		1000	0.99829	0.00398	0.99997	0.00411	0.00429	0.00013	0.99295	0.99999	1.00680
0.50	-0.90	100	0.87789	0.02487	1.00071	0.02645	0.02772	0.00312	0.95739	1.00000	1.04397
		1000	0.87742	0.00795	0.99991	0.00849	0.00886	0.00034	0.98612	0.99977	1.01456
0.50	0.00	100	0.93853	0.01983	0.99962	0.02099	0.02138	0.00228	0.96545	0.99949	1.03381
		1000	0.93903	0.00637	1.00004	0.00676	0.00684	0.00024	0.98908	0.99985	1.01119
0.50	0.90	100	1.00034	0.01417	0.99980	0.01476	0.01491	0.00144	0.97494	0.99980	1.02364
		1000	1.00054	0.00445	1.00008	0.00461	0.00475	0.00014	0.99230	1.00008	1.00766

Table 4: Simulation results for the common factor case - uncorrelated individual effects

$\varrho$	$\rho_{u^* v^*}$	<b>N</b>	$\hat{\beta}$	$s_{\hat{\beta}}$	$\hat{\beta}^c$	$s_{\hat{\beta}^c}$	$\hat{\sigma}_{\hat{\beta}^c}$	$s_{\hat{\sigma}_{\hat{\beta}^c}}$	$q_{\hat{\beta}^c}^{0.05}$	$q_{\hat{\beta}^c}^{0.50}$	$q_{\hat{\beta}^c}^{0.95}$
<b>T = 3</b>											
-0.50	-0.90	100	0.92168	0.04152	1.00010	0.04460	0.04313	0.00642	0.92761	0.99981	1.07544
		1,000	0.92227	0.01282	1.00024	0.01377	0.01400	0.00069	0.97732	1.00019	1.02295
-0.50	0.00	100	0.95940	0.03429	1.00066	0.03576	0.03480	0.00478	0.94435	1.00066	1.06085
		1,000	0.95913	0.01051	1.00011	0.01099	0.01128	0.00053	0.98239	1.00004	1.01839
-0.50	0.90	100	0.99696	0.02475	1.00112	0.02543	0.02474	0.00317	0.96071	1.00055	1.04261
		1,000	0.99606	0.00766	1.00013	0.00787	0.00796	0.00033	0.98694	1.00001	1.01269
0.00	-0.90	100	0.90301	0.04485	1.00063	0.04874	0.04809	0.00727	0.92040	1.00134	1.08118
		1,000	0.90374	0.01370	1.00017	0.01498	0.01543	0.00075	0.97615	1.00032	1.02483
0.00	0.00	100	0.95125	0.03752	1.00247	0.03966	0.03868	0.00525	0.93777	1.00201	1.06933
		1,000	0.94936	0.01151	1.00017	0.01221	0.01247	0.00054	0.97991	1.00024	1.02007
0.00	0.90	100	0.99619	0.02630	1.00137	0.02722	0.02786	0.00357	0.95706	1.00117	1.04480
		1,000	0.99530	0.00845	1.00038	0.00876	0.00891	0.00037	0.98620	1.00042	1.01513
0.50	-0.90	100	0.84030	0.05733	1.00153	0.06412	0.06369	0.00944	0.89518	1.00084	1.10874
		1,000	0.84089	0.01796	1.00054	0.02013	0.02046	0.00108	0.96704	1.00111	1.03269
0.50	0.00	100	0.91643	0.04770	1.00120	0.05248	0.05084	0.00704	0.91432	1.00064	1.08891
		1,000	0.91629	0.01411	1.00023	0.01556	0.01635	0.00079	0.97557	1.00018	1.02679
0.50	0.90	100	0.99077	0.03442	0.99919	0.03700	0.03694	0.00493	0.93823	0.99973	1.06088
		1,000	0.99157	0.01081	0.99994	0.01158	0.01187	0.00049	0.98069	1.00002	1.01906
<b>T = 10</b>											
-0.50	-0.90	100	0.91033	0.02136	1.00091	0.02312	0.02292	0.00243	0.96371	1.00044	1.03878
		1,000	0.90978	0.00689	1.00021	0.00748	0.00736	0.00025	0.98809	1.00011	1.01232
-0.50	0.00	100	0.95266	0.01726	1.00041	0.01821	0.01842	0.00187	0.97034	1.00104	1.03013
		1,000	0.95234	0.00559	0.99992	0.00586	0.00589	0.00019	0.99030	0.99973	1.00974
-0.50	0.90	100	0.99553	0.01250	1.00030	0.01279	0.01291	0.00126	0.97957	1.00032	1.02092
		1,000	0.99529	0.00390	1.00003	0.00398	0.00413	0.00012	0.99345	1.00000	1.00653
0.00	-0.90	100	0.90314	0.02193	0.99979	0.02380	0.02368	0.00246	0.96016	0.99989	1.03824
		1,000	0.90339	0.00709	0.99977	0.00764	0.00760	0.00025	0.98729	0.99955	1.01254
0.00	0.00	100	0.94956	0.01804	1.00038	0.01908	0.01909	0.00188	0.96886	1.00051	1.03246
		1,000	0.94935	0.00577	1.00005	0.00611	0.00607	0.00019	0.98984	0.99996	1.01049
0.00	0.90	100	0.99476	0.01279	0.99987	0.01320	0.01341	0.00126	0.97766	0.99969	1.02234
		1,000	0.99496	0.00404	1.00002	0.00414	0.00427	0.00012	0.99310	1.00014	1.00687
0.50	-0.90	100	0.88376	0.02337	1.00022	0.02520	0.02647	0.00287	0.95938	1.00069	1.04097
		1,000	0.88413	0.00761	1.00005	0.00819	0.00846	0.00030	0.98662	0.99975	1.01396
0.50	0.00	100	0.93864	0.01981	0.99972	0.02080	0.02096	0.00218	0.96598	0.99918	1.03464
		1,000	0.93892	0.00625	0.99990	0.00666	0.00669	0.00022	0.98905	0.99997	1.01089
0.50	0.90	100	0.99406	0.01409	1.00017	0.01468	0.01486	0.00147	0.97606	0.99993	1.02431
		1,000	0.99393	0.00438	0.99999	0.00452	0.00474	0.00014	0.99271	1.00014	1.00745

Table 5: Simulation results for the common factor case - correlated individual effects ( $\varrho_{\bar{x}\gamma} = 0.975$ )

$\varrho$	$\rho_{u^*v^*}$	<b>N</b>	$\hat{\beta}$	$s_{\hat{\beta}}$	$\hat{\beta}^c$	$s_{\hat{\beta}^c}$	$\hat{\sigma}_{\hat{\beta}^c}$	$s_{\hat{\sigma}_{\hat{\beta}^c}}$	$q_{\hat{\beta}^c}^{0.05}$	$q_{\hat{\beta}^c}^{0.50}$	$q_{\hat{\beta}^c}^{0.95}$
			<b>T = 3</b>								
-0.50	-0.90	100	0.91782	0.04274	1.00002	0.04607	0.04415	0.00662	0.92459	0.99971	1.07690
		1000	0.91896	0.01327	1.00041	0.01421	0.01426	0.00074	0.97694	1.00022	1.02420
-0.50	0.00	100	0.95886	0.03408	1.00023	0.03581	0.03543	0.00493	0.93990	0.99959	1.05895
		1000	0.95916	0.01047	1.00016	0.01100	0.01145	0.00055	0.98205	0.99998	1.01846
-0.50	0.90	100	0.99855	0.02441	0.99909	0.02508	0.02494	0.00324	0.95802	0.99823	1.04062
		1000	0.99947	0.00763	1.00003	0.00785	0.00802	0.00034	0.98723	0.99994	1.01283
0.00	-0.90	100	0.89798	0.04530	1.00163	0.04921	0.04965	0.00730	0.92211	1.00204	1.08206
		1000	0.89771	0.01391	1.00048	0.01514	0.01605	0.00079	0.97592	1.00017	1.02563
0.00	0.00	100	0.95034	0.03779	1.00142	0.04013	0.03985	0.00559	0.93552	1.00022	1.06809
		1000	0.94972	0.01195	1.00049	0.01266	0.01280	0.00061	0.97952	1.00040	1.02096
0.00	0.90	100	1.00093	0.02660	0.99975	0.02761	0.02786	0.00352	0.95343	0.99948	1.04466
		1000	1.00121	0.00842	1.00010	0.00874	0.00896	0.00037	0.98566	1.00017	1.01404
0.50	-0.90	100	0.82617	0.06098	1.00224	0.06854	0.06787	0.01118	0.88785	1.00266	1.11473
		1000	0.82648	0.01934	0.99981	0.02147	0.02184	0.00131	0.96506	0.99961	1.03423
0.50	0.00	100	0.91766	0.04810	1.00294	0.05279	0.05272	0.00754	0.91648	1.00248	1.09066
		1000	0.91644	0.01493	1.00057	0.01640	0.01700	0.00087	0.97272	1.00060	1.02776
0.50	0.90	100	1.00738	0.03376	1.00192	0.03639	0.03718	0.00478	0.94246	1.00035	1.06331
		1000	1.00589	0.01109	1.00046	0.01193	0.01193	0.00050	0.98080	1.00049	1.02032
			<b>T = 10</b>								
-0.50	-0.90	100	0.90746	0.02165	1.00006	0.02321	0.02333	0.00253	0.96166	1.00008	1.03801
		1000	0.90774	0.00694	1.00009	0.00749	0.00747	0.00027	0.98814	1.00009	1.01199
-0.50	0.00	100	0.95225	0.01717	0.99991	0.01808	0.01857	0.00197	0.96915	1.00003	1.02980
		1000	0.95226	0.00564	0.99983	0.00595	0.00595	0.00020	0.99002	0.99981	1.00944
-0.50	0.90	100	0.99679	0.01262	0.99955	0.01297	0.01294	0.00125	0.97842	0.99915	1.02165
		1000	0.99728	0.00393	1.00009	0.00404	0.00414	0.00013	0.99342	1.00001	1.00678
0.00	-0.90	100	0.89985	0.02270	0.99969	0.02437	0.02439	0.00260	0.96076	0.99965	1.03900
		1000	0.90014	0.00741	0.99985	0.00792	0.00781	0.00027	0.98686	0.99987	1.01319
0.00	0.00	100	0.94907	0.01826	0.99986	0.01933	0.01923	0.00189	0.96943	0.99925	1.03314
		1000	0.94930	0.00580	1.00001	0.00617	0.00616	0.00020	0.98972	1.00004	1.01010
0.00	0.90	100	0.99849	0.01258	1.00025	0.01298	0.01343	0.00127	0.97862	1.00021	1.02179
		1000	0.99829	0.00398	0.99997	0.00411	0.00429	0.00013	0.99295	0.99999	1.00680
0.50	-0.90	100	0.87789	0.02487	1.00071	0.02645	0.02772	0.00312	0.95739	1.00000	1.04397
		1000	0.87742	0.00795	0.99991	0.00849	0.00886	0.00034	0.98612	0.99977	1.01456
0.50	0.00	100	0.93853	0.01983	0.99962	0.02099	0.02138	0.00228	0.96545	0.99949	1.03381
		1000	0.93903	0.00637	1.00004	0.00676	0.00684	0.00024	0.98908	0.99985	1.01119
0.50	0.90	100	1.00034	0.01417	0.99980	0.01476	0.01491	0.00144	0.97494	0.99980	1.02364
		1000	1.00054	0.00445	1.00008	0.00461	0.00475	0.00014	0.99230	1.00008	1.00766

# CESifo Working Paper Series

for full list see [www.cesifo-group.org/wp](http://www.cesifo-group.org/wp)

(address: Poschingerstr. 5, 81679 Munich, Germany, [office@cesifo.de](mailto:office@cesifo.de))

---

- 2545 John Beirne, Guglielmo Maria Caporale, Marianne Schulze-Ghattas and Nicola Spagnolo, Volatility Spillovers and Contagion from Mature to Emerging Stock Markets, February 2009
- 2546 Ali Bayar and Bram Smeets, Economic and Political Determinants of Budget Deficits in the European Union: A Dynamic Random Coefficient Approach, February 2009
- 2547 Jan K. Brueckner and Anming Zhang, Airline Emission Charges: Effects on Airfares, Service Quality, and Aircraft Design, February 2009
- 2548 Dolores Messer and Stefan C. Wolter, Money Matters – Evidence from a Large-Scale Randomized Field Experiment with Vouchers for Adult Training, February 2009
- 2549 Johannes Rincke and Christian Traxler, Deterrence through Word of Mouth, February 2009
- 2550 Gabriella Legrenzi, Asymmetric and Non-Linear Adjustments in Local Fiscal Policy, February 2009
- 2551 Bruno S. Frey, David A. Savage and Benno Torgler, Surviving the Titanic Disaster: Economic, Natural and Social Determinants, February 2009
- 2552 Per Engström, Patrik Hesselius and Bertil Holmlund, Vacancy Referrals, Job Search, and the Duration of Unemployment: A Randomized Experiment, February 2009
- 2553 Giorgio Bellettini, Carlotta Berti Ceroni and Giovanni Prarolo, Political Persistence, Connections and Economic Growth, February 2009
- 2554 Steinar Holden and Fredrik Wulfsberg, Wage Rigidity, Institutions, and Inflation, February 2009
- 2555 Alexander Haupt and Tim Krieger, The Role of Mobility in Tax and Subsidy Competition, February 2009
- 2556 Harald Badinger and Peter Egger, Estimation of Higher-Order Spatial Autoregressive Panel Data Error Component Models, February 2009
- 2557 Christian Keuschnigg, Corporate Taxation and the Welfare State, February 2009
- 2558 Marcel Gérard, Hubert Jayet and Sonia Paty, Tax Interactions among Belgian Municipalities: Does Language Matter?, February 2009
- 2559 António Afonso and Christophe Rault, Budgetary and External Imbalances Relationship: A Panel Data Diagnostic, February 2009

- 2560 Stefan Krasa and Mattias Polborn, Political Competition between Differentiated Candidates, February 2009
- 2561 Carsten Hefeker, Taxation, Corruption and the Exchange Rate Regime, February 2009
- 2562 Jiahua Che and Gerald Willmann, The Economics of a Multilateral Investment Agreement, February 2009
- 2563 Scott Alan Carson, Demographic, Residential, and Socioeconomic Effects on the Distribution of 19<sup>th</sup> Century US White Statues, February 2009
- 2564 Philipp Harms, Oliver Lorz and Dieter Urban, Offshoring along the Production Chain, February 2009
- 2565 Patricia Apps, Ngo Van Long and Ray Rees, Optimal Piecewise Linear Income Taxation, February 2009
- 2566 John Whalley and Shunming Zhang, On the Arbitrariness of Consumption, February 2009
- 2567 Marie-Louise Leroux, Endogenous Differential Mortality, Non-Contractible Effort and Non Linear Taxation, March 2009
- 2568 Joanna Bęza-Bojanowska and Ronald MacDonald, The Behavioural Zloty/Euro Equilibrium Exchange Rate, March 2009
- 2569 Bart Cockx and Matteo Picchio, Are Short-Lived Jobs Stepping Stones to Long-Lasting Jobs?, March 2009
- 2570 David Card, Jochen Kluve and Andrea Weber, Active Labor Market Policy Evaluations: A Meta-analysis, March 2009
- 2571 Frederick van der Ploeg and Anthony J. Venables, Harnessing Windfall Revenues: Optimal Policies for Resource-Rich Developing Economies, March 2009
- 2572 Ondřej Schneider, Reforming Pensions in Europe: Economic Fundamentals and Political Factors, March 2009
- 2573 Jo Thori Lind, Karl Ove Moene and Fredrik Willumsen, Opium for the Masses? Conflict-Induced Narcotics Production in Afghanistan, March 2009
- 2574 Silvia Marchesi, Laura Sabani and Axel Dreher, Agency and Communication in IMF Conditional Lending: Theory and Empirical Evidence, March 2009
- 2575 Carlo Altavilla and Matteo Ciccarelli, The Effects of Monetary Policy on Unemployment Dynamics under Model Uncertainty - Evidence from the US and the Euro Area, March 2009
- 2576 Falko Fecht, Kjell G. Nyborg and Jörg Rocholl, The Price of Liquidity: Bank Characteristics and Market Conditions, March 2009



- 2577 Giorgio Bellettini and Filippo Taddei, Real Estate Prices and the Importance of Bequest Taxation, March 2009
- 2578 Annette Bergemann and Regina T. Riphahn, Female Labor Supply and Parental Leave Benefits – The Causal Effect of Paying Higher Transfers for a Shorter Period of Time, March 2009
- 2579 Thomas Eichner and Rüdiger Pethig, EU-Type Carbon Emissions Trade and the Distributional Impact of Overlapping Emissions Taxes, March 2009
- 2580 Antonios Antypas, Guglielmo Maria Caporale, Nikolaos Kourogenis and Nikitas Pittis, Selectivity, Market Timing and the Morningstar Star-Rating System, March 2009
- 2581 António Afonso and Christophe Rault, Bootstrap Panel Granger-Causality between Government Budget and External Deficits for the EU, March 2009
- 2582 Bernd Süßmuth, Malte Heyne and Wolfgang Maennig, Induced Civic Pride and Integration, March 2009
- 2583 Martin Peitz and Markus Reisinger, Indirect Taxation in Vertical Oligopoly, March 2009
- 2584 Petra M. Geraats, Trends in Monetary Policy Transparency, March 2009
- 2585 Johannes Abeler, Armin Falk, Lorenz Götte and David Huffman, Reference Points and Effort Provision, March 2009
- 2586 Wolfram F. Richter, Taxing Education in Ramsey's Tradition, March 2009
- 2587 Yin-Wong Cheung, Menzie D. Chinn and Eiji Fujii, China's Current Account and Exchange Rate, March 2009
- 2588 Alexander Haupt and Silke Uebelmesser, Voting on Labour-Market Integration and Education Policy when Citizens Differ in Mobility and Ability, March 2009
- 2589 Hans Jarle Kind, Marko Koethenbuerger and Guttorm Schjelderup, Should Utility-Reducing Media Advertising be Taxed?, March 2009
- 2590 Alessandro Cigno, How to Avoid a Pension Crisis: A Question of Intelligent System Design, March 2009
- 2591 Helmut Lutkepohl and Fang Xu, The Role of the Log Transformation in Forecasting Economic Variables, March 2009
- 2592 Rainald Borck, Hyun-Ju Koh and Michael Pflüger, Inefficient Lock-in and Subsidy Competition, March 2009
- 2593 Paolo M. Panteghini, On the Equivalence between Labor and Consumption Taxation, March 2009

- 2594 Bruno S. Frey, Economists in the PITS?, March 2009
- 2595 Natalie Chen and Dennis Novy, International Trade Integration: A Disaggregated Approach, March 2009
- 2596 Frédérique Bec and Christian Gollier, Term Structure and Cyclicity of Value-at-Risk: Consequences for the Solvency Capital Requirement, March 2009
- 2597 Carsten Eckel, International Trade and Retailing, March 2009
- 2598 Gianni De Nicolò and Iryna Ivaschenko, Global Liquidity, Risk Premiums and Growth Opportunities, March 2009
- 2599 Jay Pil Choi and Heiko Gerlach, International Antitrust Enforcement and Multi-Market Contact, March 2009
- 2600 Massimo Bordignon and Guido Tabellini, Moderating Political Extremism: Single Round vs Runoff Elections under Plurality Rule, April 2009
- 2601 Ana B. Ania and Andreas Wagener, The Open Method of Coordination (OMC) as an Evolutionary Learning Process, April 2009
- 2602 Simon Gächter, Daniele Nosenzo, Elke Renner and Martin Sefton, Sequential versus Simultaneous Contributions to Public Goods: Experimental Evidence, April 2009
- 2603 Philippe Jehiel and Andrew Lilico, Smoking Today and Stopping Tomorrow: A Limited Foresight Perspective, April 2009
- 2604 Andreas Knabe, Steffen Rätzel, Ronnie Schöb and Joachim Weimann, Dissatisfied with Life, but Having a Good Day: Time-Use and Well-Being of the Unemployed, April 2009
- 2605 David Bartolini and Raffaella Santolini, Fiscal Rules and the Opportunistic Behaviour of the Incumbent Politician: Evidence from Italian Municipalities, April 2009
- 2606 Erkki Koskela and Jan König, Can Profit Sharing Lower Flexible Outsourcing? A Note, April 2009
- 2607 Michel Beine, Frédéric Docquier and Çağlar Özden, Diasporas, April 2009
- 2608 Gerd Ronning and Hans Schneeweiss, Panel Regression with Random Noise, April 2009