A comparison of dynamic panel data estimators: Monte Carlo evidence and an application to the investment function

Andreas Behr
(University of Frankfurt)

Discussion paper 05/03
Economic Research Centre
of the Deutsche Bundesbank

February 2003

Deutsche Bundesbank, Wilhelm-Epstein-Strasse 14, 60431 Frankfurt am Main,
Postfach 10 06 02, 60006 Frankfurt am Main

Tel +49 69 95 66-1
Telex within Germany 4 1 227, telex from abroad 4 14 431, fax +49 69 5 60 10 71

Please address all orders in writing to: Deutsche Bundesbank,
Press and Public Relations Division, at the above address or via fax No. +49 69 95 66-30 77

ISBN 3–935821–46–8

Summary

In our analysis we discuss several dynamic panel data estimators proposed in the literature and assess their performance in Monte Carlo simulations. It is a well known fact that the natural choice, the least squares dummy variable estimator is biased in the context of dynamic estimation. The estimators taking into account the resulting bias can be grouped broadly into the class of instrumental estimators and the class of direct bias corrected estimators.

The simulation results clearly favour the direct bias corrected estimators, especially the estimator proposed by Hansen (2001). The superiority of these estimators decreases with growing numbers of individuals in the simulation. This is the well known fact of large sample properties of the GMM-methods. In the case of endogenous predetermined regressors, the system-estimator proposed by Blundell and Bond is unbiased and most efficient, while direct bias corrected estimators perform similar to the GMM-estimator proposed by Arellano and Bond (1991).

Turning to the empirical comparison, we find that the different estimators lead to the same conclusions concerning the investment behaviour of German manufacturing firms based on the Deutsche Bundesbank's Corporate Balance Sheet Statistics. Investment is strongly positive dependent on lagged investment and *Q*. Nevertheless, in detail the differences of the estimated parameters are not negligible.

JEL-code: C 15, C 23, E 22

Keywords: dynamic panel data estimation, GMM, bias correction, investment

Zusammenfassung

In der vorliegenden Arbeit werden verschiedene in der Literatur vorgeschlagen dynamische Schätzer für Paneldaten diskutiert und im Rahmen einer Monte Carlo-Studie verglichen. Es ist wohlbekannt, dass der Least Squares Dummy Variable-Estimator für den Fall verzögerter endogener erklärender Variablen einen Bias aufweist. Die diskutierten Schätzer lassen sich zwei unterschiedlichen Klassen zuordnen, einer Klasse von Instrumentenschätzern und einer Klasse von biaskorrigierten Schätzern.

Den Ergebnissen der Simulationsstudie zufolge sind die biaskorrigierten Schätzer leicht überlegen, insbesondere die von Hansen (2001) vorgeschlagene Biaskorrektur. Die Überlegenheit nimmt jedoch mit wachsender Zahl der beobachteten Einheiten ab. Hier spiegeln sich die bekannt günstigen Eigenschaften von GMM-Schätzern bei großer Beobachtungszahl wider. Im Falle endogener vorherbestimmter Regressoren weist der von Blundell und Bond (1998) vorgeschlagene System-GMM-Schätzer die höchste Effizienz auf. Biaskorrigierte Schätzer führen hier zu vergleichbaren Ergebnissen wie der GMM-Schätzer von Arellano und Bond (1991).

Bei der empirischen Anwendung zur Schätzung von dynamischen Q-Invstitionsfunktionen für Unternehmen des deutschen Verarbeitenden Gewerbes auf Grundlage der Bilanzstatistik der Deutschen Bundesbank, zeigt sich eine starke positive Abhängigkeit der Investitionen, sowohl von den Vorjahresinvestitionen als auch von Q. Bei gleicher ökonomischer Grundaussage weisen die mittels der verschiedenen diskutierten Methoden geschätzten Parameter jedoch nicht zu vernachlässigende Unterschiede auf.

Contents

List of Tables

List of Figures

A comparison of dynamic panel data estimators: Monte Carlo evidence and an application to the investment function

1. Introduction

The paper discusses methods of dynamic panel data estimation. It is well known that the use of the lagged dependent variable as a right hand side variable introduces specific estimation problems, especially the fixed effects estimator becoming biased.

In our analysis we compare several unbiased or near unbiased estimators suggested in the literature. By a Monte Carlo study we assess the bias and efficiency of various proposed estimators under different data generating processes. Especially two classes of estimators will be compared, the class of instrumental estimators, e.g. the well-known Generalized Methods of Moments (GMM) estimator (Arellano/Bond 1991), and the class of direct bias correcting estimators, e.g. the estimator suggested by Kiviet (1995). While in some simulation studies (Kiviet (1995), Judson and Owen (1999), Hansen (2001)) a corrected LSDV estimator is found superior compared to GMM-estimators these simulations take no account of System-GMM-estimators proposed by Blundell and Bond (1998).

After the Monte Carlo study we apply all discussed estimators to estimate a dynamic *Q*-investment function. This application is of interest in several respects. The *Q*-theory can be seen as the standard approach of empirical investment research and the explanation of firms investment behaviour is one of the central issues in empirical economics. Because there exists now a variety of suggested dynamic panel data estimators, it is of interest to assess the differences of the results due to the chosen procedure.

Following the introduction a brief presentation of the problems caused by lagged dependent variables included as right hand side variables in regression functions and the basic idea of GMM is given. In section 3 we discuss several dynamic panel data estimators suggested in the literature. The Monte Carlo simulation is contained in section 4. In section 5 we apply the discussed dynamic panel data estimators to estimate *Q*-investment functions for German manufacturing firms and section 6 concludes.

2. The problem of bias caused by lagged dependant variables

The following section explains in short the problem of correlation between explanatory variables and the error term leading to biased estimators.

2.1. The idea of instrumentation

The method of instrumentation is one possible way to prevent the bias resulting from correlation between the regressor x and the error term ε. The idea of instrumentation can be stated as follows:

"Find a variable Z, that is highly correlated with X, but does not correlate with ε. Use as the new regressor only that part of the observable variable X which correlates with Z and is orthogonal to ε."

Starting with the problem of correlation between the observable and the error term

$$p\lim\left(\frac{1}{n}X'\varepsilon\right)\neq 0$$

in the linear regression case

$$y = X\beta + \varepsilon \text{ with } \operatorname{var}(\varepsilon) = \sigma^2 I$$

the bias can be circumvented using an instrument Z that correlates with X but is orthogonal to ε:

$$p\lim\left(\frac{1}{n}Z'X\right) = \Sigma_{ZX} \neq 0$$

$$p\lim\left(\frac{1}{n}Z'\varepsilon\right) = 0$$

Premultiplying the regression with Z leads to the residual $Z'\varepsilon$ and the following variance:

$$Z'y = Z'X\beta + Z'\varepsilon$$

$$\operatorname{var}(Z'\varepsilon) = Z' \operatorname{var}(\varepsilon) Z = \sigma^2 Z'Z$$

Making use of the Generalized Least Square-Estimator (GLS) with $V^{-1} = (Z'Z)^{-1}$ and dropping σ^2 leads to the following instrumental variable estimator:

$$b_{IV} = \left[(Z'X)' V^{-1} Z'X\right]^{-1} (Z'X)' V^{-1} Z'y = \left(X'Z(Z'Z)^{-1}Z'X\right)^{-1} X'Z(Z'Z)^{-1}Z'y$$

$$b_{IV} = (X'PX)^{-1} X'Py \qquad \text{with } P = Z(Z'Z)^{-1}Z'$$

Inserting $y = X\beta + \varepsilon$ results in

$$b_{IV} = (X'PX)^{-1} X'P(X\beta + \varepsilon) = (X'PX)^{-1} X'PX\beta + (X'PX)^{-1} X'P\varepsilon$$

$$= \left(X'Z(Z'Z)^{-1}Z'X\right)^{-1} X'Z(Z'Z)^{-1}Z'X\beta + \left(X'Z(Z'Z)^{-1}Z'X\right)^{-1} X'Z(Z'Z)^{-1}Z'\varepsilon$$

$$= \beta + \left(\frac{1}{n}X'Z(Z'Z)^{-1}Z'X\right)^{-1} \frac{1}{n}X'Z(Z'Z)^{-1}Z'\varepsilon$$

Now taking probability limits shows the estimator being unbiased:

$$p\lim b_{IV} = \beta + P\lim\left[\left(\frac{1}{n}X'Z(Z'Z)^{-1}Z'X\right)^{-1}\frac{1}{n}X'Z(Z'Z)^{-1}Z'\varepsilon\right]$$

$$p\lim b_{IV} = \beta + \left(\Sigma_{XZ}\Sigma_{ZZ}^{-1}\Sigma_{ZX}\right)^{-1}\Sigma_{XZ}\Sigma_{ZZ}^{-1}\Sigma_{Z\varepsilon} = \beta$$

where we make use of $\Sigma_{Z\varepsilon} = 0$.

Therefore it is evident, that basic to the idea of instrumenting is the assumed uncorrelatedness of Z and ε.

The instrumentation can be made transparent through the exposition as a two stage procedure. In the first step the explanatory variable X is regressed on the instrument Z. The regression values $\hat{X}$ containing the linear dependent part of X are used as explanatory variables in the second step.

Based on the first auxiliary regression

$$X = Z\gamma + \nu \,,$$

regression values are obtained

$$\hat{X} = Z\hat{\gamma} = Z(Z'Z)^{-1}Z'X \,,$$

which will be used as new regressors in the second stage

$$y = \hat{X}b_{2S} + \varepsilon$$

$$b_{2S} = \left(\hat{X}'\hat{X}\right)^{-1}\hat{X}'y \,.$$

Inserting $Z(Z'Z)^{-1}Z'X$ for the regression values $\hat{X}$ leads to the instrumental variable estimator b_{IV}:

$$b_{2S} = \left(\left(Z(Z'Z)^{-1}Z'X\right)'Z(Z'Z)^{-1}Z'X\right)^{-1}\left(Z(Z'Z)^{-1}Z'X\right)'y$$

$$= \left(X'Z(ZZ')^{-1}Z'Z(Z'Z)^{-1}Z'X\right)^{-1}X'Z(ZZ')^{-1}Z'y$$

$$\left(X'Z(ZZ')^{-1}Z'X\right)^{-1}X'Z(ZZ')^{-1}Z'y = (X'PX)^{-1}X'Py = b_{IV}$$

2.2. Generalized Methods of Moments (GMM)

During the last decade the concept of Generalized Methods of Moments (GMM) has become increasingly popular. Before discussing some dynamic panel data estimators based on the ideas of GMM, the basic concept is introduced.[1]

The concept of GMM is often a simple alternative, if the explicit Maximum-Likelihood function is difficult to derive. The core of the GMM-estimation is the use of orthogonality conditions. In general GMM can be seen as being especially suited for large data files, while when using only few observations GMM is often less efficient then alternative methods.

The simple OLS-estimation can be represented as an application of the method of moments. The condition of uncorrelatedness of the explanatory variable and the error term is the point to start from:

$$E(X'\varepsilon) = 0$$

Applying this condition to the sample results in the following conditions:

$$\frac{1}{n} X'\left(y - X\hat{\beta}\right) = 0$$

Solving this equation for the parameter vector results in the well know OLS-estimator:

$$\hat{\beta} = (X'X)^{-1} X'y$$

In the same fashion the instrumentation can be expressed as an application of the method of moments where use is made of the assumption that the instrument is orthogonal to the error term:

$$E(Z'\varepsilon) = 0$$

Applying this condition to the sample

$$\frac{1}{n} Z'\left(y - X\hat{\beta}\right) = 0$$

and solving for the parameter vector results in

$$\hat{\beta}_{IV} = b_{2S} = (X'PX)^{-1} X'Py$$

$$\text{with } P = Z(Z'Z)^{-1} Z'$$

1 See e.g. the introduction by Mátyás/Harris (1999).

when applying GLS with $V^{-1} = (Z'Z)^{-1}$. If the number of instruments equals the number of explanatory variables, the estimator simplifies to

$$\hat{\beta}_{IV} = (Z'X)^{-1} Z'y .$$

3. Dynamic Panel Data Estimation

In this section we discuss several suggested estimators for dynamic panel data models. The starting point is the well known bias of the fixed effects model (Nickell 1981) which would be the natural choice when allowing for individual effects.[2]

3.1. The bias of the fixed effects model

The linear model to estimate contains explanatory variables x_t as well as the lagged endogenous variable y_{t-1}.

$$y_{it} = \rho y_{i,t-1} + x'_{it}\beta + \alpha_i + \varepsilon_{it}$$

$$\text{where} \quad \varepsilon_{it} \sim N\left(0, \sigma_\varepsilon^2\right) \qquad \text{and} \qquad |\rho| < 1$$

$i = 1,...,N$ index for individuals

$t = 1,...,T$ index for years

x'_{it} row vector of explanatory variables, dimension k

ρ unknown parameter of the lagged endogenous variable

β unknown parameter vector of the k explanatory variables

α_i individual specific fixed effects

Further we make the following assumptions:

- the error term is orthogonal to the exogenous variables: $E(x'_{it}\varepsilon_{it}) = 0$

- the exogenous variables might be correlated with the individual effect $E(x'_{it}\alpha_i) \neq 0$

- the error term (i.i.d.) is uncorrelated with the lagged endogenous variable: $E\left(y_{i,t-1}\varepsilon_{it}\right) = 0$

Using matrix notation the model can be expressed as

$$y = y_{-1}\rho + X\beta + D\alpha + \varepsilon$$

---

2 For an overview of dynamic panel data estimation see Mátyás/Sevestre (1995) and Baltagi (2001).

where

$$y = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_N \end{pmatrix} \qquad y_i = \begin{pmatrix} y_{i1} \\ y_{i2} \\ \vdots \\ y_{iT} \end{pmatrix} \qquad X = \begin{pmatrix} X_1 \\ X_2 \\ \vdots \\ X_N \end{pmatrix} \qquad X_i = \begin{pmatrix} x'_{i1} \\ x'_{i2} \\ \vdots \\ x'_{iT} \end{pmatrix}$$

$$D = I_N \otimes e \qquad e = \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix} \text{with dimension } T \qquad \alpha = \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_N \end{pmatrix}$$

The simple Least Squares Dummy Variable Estimator (LSDV) is

$$My = My_{-1}\rho + MX\beta + M\varepsilon$$

$$y^* = y^*_{-1}\,\rho + X^*\beta + \varepsilon^*$$

where

$$M = I_{NT} - D(D'D)^{-1}D' = I_N \otimes I_T - I_N \otimes \frac{1}{T}ee' = I_N \otimes \left(I_{T-}\frac{1}{T}ee' \right).$$

Premultiplying with matrix M results in variables measured as deviations from the individual specific means. Because the demeaning procedure makes use of all available time periods,

$$y^*_{i,t-1} = y_{i,t-1} - \frac{1}{T}\sum_{t=1}^{T} y_{i,t-1}$$

$$\varepsilon^*_{it} = \varepsilon_{it} - \frac{1}{T}\sum_{t=1}^{T}\varepsilon_{it}$$

the demeaned lagged endogenous variable correlates with the demeaned error term:

$$E\left(y^*_{i,t-1}\,\varepsilon^*_{it}\right) \neq 0\,.$$

The error term $\varepsilon_{i,t-1}$ is contained with the weight $1 - \frac{1}{T}$ in $y^*_{i,t-1}$ and with the weight $-\frac{1}{T}$ in $\varepsilon^*_{i,t}$. This correlation renders the LSDV-estimators $\hat{\rho}$ and $\hat{\beta}$ biased. It is also obvious that the correlation decreases in T, the number of years available. But since a typical microeconomic panel contains a large number of individuals N but only a few periods of time T, the asymptotic behaviour of the estimator is of special interest when N

tends to infinity ($N \to \infty$). The asymptotic bias of the LSDV-estimator was derived by Nickell (1981):

$$\underset{N\to\infty}{p\lim}\left(\hat{\rho}-\rho\right)=\left(\underset{N\to\infty}{p\lim}\frac{1}{NT}y*'_{-1}\,Py*_{-1}\right)^{-1}\frac{-\sigma_{\varepsilon}^{2}}{T(1-\rho)}\left(1-\frac{1}{T}\frac{\left(1-\rho^{T}\right)}{T\left(1-\rho\right)}\right)$$

where $P = I_{NT} - X*\left(X*'X*\right)^{-1}X*$ is the residual maker.

Inserting P into the bias expressions leads to

$$\underset{N\to\infty}{p\lim}\left(\hat{\rho}-\rho\right)=\left(\underset{N\to\infty}{p\lim}\frac{1}{NT}y*'_{-1}\left(I_{NT}-X*\left(X*'X*\right)^{-1}X*'\right)y*_{-1}\right)^{-1}\frac{-\sigma_{\varepsilon}^{2}}{T(1-\rho)}\left(1-\frac{1}{T}\frac{\left(1-\rho^{T}\right)}{T\left(1-\rho\right)}\right)$$

$$=\left(\underset{N\to\infty}{p\lim}\frac{1}{NT}y*'_{-1}\left(y*_{-1}-X*\hat{\beta}\right)\right)^{-1}\frac{-\sigma_{\varepsilon}^{2}}{T(1-\rho)}\left(1-\frac{1}{T}\frac{\left(1-\rho^{T}\right)}{T\left(1-\rho\right)}\right)$$

$$=\left(\underset{N\to\infty}{p\lim}\frac{1}{NT}y*'_{-1}\,\varepsilon*_{-1}\right)^{-1}\frac{-\sigma_{\varepsilon}^{2}}{T(1-\rho)}\left(1-\frac{1}{T}\frac{\left(1-\rho^{T}\right)}{T\left(1-\rho\right)}\right)$$

The asymptotic bias of the parameter vector $\hat{\beta}$ of the remaining explanatory variables is given by:

$$\underset{N\to\infty}{p\lim}\left(\hat{\beta}-\beta\right)=-\underset{N\to\infty}{p\lim}\left(\left(X*'X*\right)^{-1}X*'y*_{-1}\right)\underset{N\to\infty}{p\lim}\left(\hat{\rho}-\rho\right)$$

$$=-\underset{N\to\infty}{p\lim}\,\hat{\beta}\,\underset{N\to\infty}{p\lim}\left(\hat{\rho}-\rho\right).$$

The asymptotic bias $\hat{\rho}-\rho$ is growing in ρ, N, σ_{ε}^{2} and in the sum of squares $\varepsilon*_{-1}{}'\,\varepsilon*_{-1}$, while it is decreasing in T.

3.2. Some proposed dynamic panel data estimators

In the following we discuss some dynamic panel data estimators proposed in the literature which will be examined in a Monte Carlo study. In the following we assume the explanatory variables to be at least predetermined what leads to the assumption of

$$E(x_{it}\varepsilon_{is})=0 \text{ for } s \ge t \text{ but } E(x_{it}\varepsilon_{is}) \ne 0 \text{ for } s < t.$$

3.2.1. The Anderson-Hsiao estimator

The estimator suggested by Anderson and Hsiao (1982) is based on the differenced form of the original equation[3]

$$y_{it} = \rho y_{i,t-1} + x'_{it}\beta + \alpha_i + \varepsilon_{it}$$

$$y_{it} - y_{i,t-1} = \rho\left(y_{i,t-1} - y_{i,t-2}\right) + \left(x'_{it} - x'_{i,t-1}\right)\beta + \varepsilon_{it} - \varepsilon_{i,t-1}$$

which cancels the individual fixed effects assumed to possibly correlate with the exogenous variables $\left(E\left(x'_{it}\alpha_i\right) \neq 0\right)$. Using matrix notation we can write

$$y = y_{-1}\rho + X\beta + D\alpha + \varepsilon$$

$$Fy = Fy_{-1}\rho + FX\beta + F\varepsilon$$

where

$$F = I_N \otimes F_T \text{ and } F_T = \begin{pmatrix} -1 & 1 & 0 & \cdots & 0 & 0 \\ 0 & -1 & 1 & \cdots & 0 & 0 \\ \vdots & & & & & \\ 0 & 0 & 0 & \cdots & -1 & 1 \end{pmatrix} \text{ with dimension } (T-1)\times T .$$

Because $FD = 0$, the individual fixed effects cancel out. But the difference of the lagged endogenous variable

$$y_{i,t-1} - y_{i,t-2} = \rho\left(y_{i,t-2} - y_{i,t-3}\right) + \left(x'_{i,t-1} - x'_{i,t-2}\right)\beta + \varepsilon_{i,t-1} - \varepsilon_{i,t-2}$$

is now obviously correlated with the error term

$$\varepsilon_{it} - \varepsilon_{i,t-1}.$$

Therefore $E\left(dy_{i,t-1}d\varepsilon_{it}\right) \neq 0$ and the estimator will be biased.

Anderson and Hsiao suggest using level instruments y_{t-2} or the lagged difference $y_{i,t-2} - y_{i,t-3}$ as an instrument for the differenced lagged endogenous regressor $y_{i,t-1} - y_{i,t-2}$. These instruments can be expected to be uncorrelated with the differenced error term:

$$E\left(y_{i,t-2}d\varepsilon_{it}\right) = 0 \text{ and } E\left(dy_{i,t-2}d\varepsilon_{it}\right) = 0 \ .$$

3 See also Anderson/Hsiao (1981).

When analysing the properties of the two possible instruments Arellano (1989) found the estimator using level instruments superior because of having smaller variances and no points of singularities. Furthermore the use of the levels as instruments has the advantage of loosing one year less what can be relevant in the practical use, especially when using data files with a large number of individuals and few years.

The differencing imposes a MA(1) structure on the error term, even when the errors ε_{it} originally where not correlated over time.

Estimation will then make use of the following matrices in the case of levels as instruments:

$$\tilde{Z}_i = \begin{bmatrix} y_{i,1} & x'_{i3} - x'_{i2} \\ y_{i,2} & x'_{i4} - x'_{i3} \\ \vdots & \vdots \\ y_{i,T-2} & x'_{iT} - x'_{i,T-1} \end{bmatrix}$$

$$\tilde{X}_i = \begin{bmatrix} y_{i2} - y_{i1} & x'_{i3} - x'_{i2} \\ y_{i3} - y_{i,2} & x'_{i4} - x'_{3} \\ \vdots & \vdots \\ y_{i,T-1} - y_{i,T-2} & x'_{iT} - x'_{i,T-1} \end{bmatrix}$$

$$\tilde{y}_i = \begin{bmatrix} y_{i3} - y_{i,2} \\ y_{i4} - y_{i,3} \\ \vdots \\ y_{i,T} - y_{i,T-1} \end{bmatrix}.$$

And as follows for the use of differenced instruments:

$$\tilde{X}_i = \begin{bmatrix} y_{i3} - y_{i2} & x'_{i4} - x'_{i3} \\ y_{i4} - y_{i,3} & x'_{i5} - x'_{4} \\ \vdots & \vdots \\ y_{i,T-1} - y_{i,T-2} & x'_{iT} - x'_{i,T-1} \end{bmatrix}$$

$$\tilde{Z}_i = \begin{bmatrix} y_{i2} - y_{i1} & x'_{i4} - x'_{i3} \\ y_{i3} - y_{i,2} & x'_{i5} - x'_{i4} \\ \vdots & \vdots \\ y_{i,T-2} - y_{i,T-3} & x'_{iT} - x'_{i,T-1} \end{bmatrix}$$

$$\tilde{y}_i = \begin{bmatrix} y_{i4} - y_{i,3} \\ y_{i5} - y_{i,4} \\ \vdots \\ y_{i,T} - y_{i,T-1} \end{bmatrix}.$$

Stacking the observations for all individuals results in the two estimators:

$$Z = \begin{bmatrix} \tilde{Z}_1 \\ \tilde{Z}_2 \\ \vdots \\ \tilde{Z}_N \end{bmatrix}, \quad X = \begin{bmatrix} \tilde{X}_1 \\ \tilde{X}_2 \\ \vdots \\ \tilde{X}_N \end{bmatrix}, \quad y = \begin{bmatrix} \tilde{y}_1 \\ \tilde{y}_2 \\ \vdots \\ \tilde{y}_N \end{bmatrix}$$

$$\hat{\gamma}^{AH} = (XPX)^{-1} X'Py \quad \text{where} \quad P = Z(Z'Z)^{-1}Z'.$$

We add the symbol L or D to indicate the use of levels or differences as instruments ($\hat{\gamma}^{AH,L}$, $\hat{\gamma}^{AH,D}$).

3.2.2. The Arellano-Bond estimator

In empirical work using firm level or household panel data the Generalized Method of Moments estimator (GMM) suggested by Arrelano and Bond (1991) has become increasingly popular. The estimator is similar to the estimator suggested by Anderson and Hsiao but exploits additional moment restrictions, which enlarges the set of instruments.

The dynamic equation to be estimated in levels is

$$y_{it} = \rho y_{i,t-1} + x'_{it}\beta + \alpha_i + \varepsilon_{it}$$

where differencing eliminates the individual effects α_i:

$$y_{it} - y_{i,t-1} = \rho(y_{i,t-1} - y_{i,t-2}) + (x'_{it} - x'_{i,t-1})\beta + \varepsilon_{it} - \varepsilon_{i,t-1}$$

For each year we now look for the instruments available for instrumenting the difference equation. For $t = 3$ the equation to be estimated is

$$y_{i3} - y_{i2} = \rho(y_{i2} - y_{i1}) + (x'_{i3} - x'_{i2})\beta + \varepsilon_{i3} - \varepsilon_{i2}$$

where the instruments (again assuming x being at least predetermined) $y_{i,1}$, x'_{i2} and x'_{i1} are available.

For $t = 4$ the equation is

$$y_{i4} - y_{i3} = \rho(y_{i3} - y_{i2}) + (x'_{i4} - x'_{i3})\beta + \varepsilon_{i4} - \varepsilon_{i3}$$

and the instruments $y_{i,1}, y_{i,2}, x'_{i1}, x'_{i2}$ and x'_{i3} are available. As can be seen, the time periods valid for instrumentation enlarge and for the equation in the final Period T

$$y_{iT} - y_{i,T-1} = \rho\left(y_{i,T-1} - y_{i,T-2}\right) + \left(x'_{iT} - x'_{i,T-1}\right)\beta + \varepsilon_{iT} - \varepsilon_{i,T-1}$$

the instruments $y_{i,1}, y_{i,2}$, ..., $y_{i,T-2}, x'_{i1}, x'_{i2}$, ..., $x'_{i,T-1}$ are available.

The instrumented equation is

$$W'Fy = W'FX\gamma + W'F\varepsilon$$

where

$$X_i = \begin{bmatrix} y_{i2} - y_{i1} & x'_{i3} - x'_{i2} \\ y_{i3} - y_{i,2} & x'_{i4} - x'_{i3} \\ \vdots & \vdots \\ y_{i,T-1} - y_{i,T-2} & x'_{iT} - x'_{i,T-1} \end{bmatrix}$$

$$X = \left(y_{-1}, X\right), \qquad \gamma' = \left(\rho, \beta'\right), \qquad W = \left(W_1{}', W_2{}', ..., W_N{}'\right)'$$

$$W_i = \begin{bmatrix} [y_{i1}, x'_{i1}, x'_{i2}] & 0 & \cdots & 0 \\ 0 & [y_{i1}, y_{i2}, x'_{i1}, x'_{i2}, x'_{i3}] & \cdots & 0 \\ 0 & 0 & \ddots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & [y_{i1}, y_{i2}, ... y_{i,T-2}, x'_{i1}, x'_{i2}, ..., x'_{iT-1}] \end{bmatrix}$$

As has been shown for the simple case of the instrumental variable estimation, the estimation procedure can be seen as a two-step estimation. First a cross-section auxiliary equation

$$y_{i,t} - y_{i,t-1} = \hat{a}_{1t} y_{i,t-2} + \hat{a}_{2t} y_{i,t-3} + ... + x'_{it-1}\hat{b}_{1t} + x'_{it-2}\hat{b}_{2t} + ... + v_{it}$$

is estimated and in the second step the resulting estimates are used as explanatory variables in the equation of original interest.

In the k-explanatory variable case the maximal number of parameters to be estimated is $T - 2 + k(T-1) = (k+1)(T-1) - 1$ which determines the number of individuals which has to be available to allow estimation.

Because the differencing operation introduces first order autocorrelation into the error term, the first-step estimator makes use of a covariance matrix taking this autocorrelation into account.

$$V = W'GW = \sum_{i=1}^{N} W_i' G_T W_i$$

where $G = \left(I_N \otimes G_T{}'\right)$ and $G_T = F_T F_T{}' = \begin{bmatrix} 2 & -1 & & 0 \\ -1 & 2 & \ddots & \\ & \ddots & \ddots & -1 \\ 0 & & -1 & 2 \end{bmatrix}$.

Premultiplying the matrix F results in transforming the original observations into differences. Because $Var(Fu) = F\sigma^2 F'$, the covariance matrix $V = FF'$ is used as a first-step approximation to the covariance matrix.

The two-step GMM estimator uses the residuals of the first-step estimation to estimate the covariance matrix as suggested by White (1980):

$$\hat{V} = \sum_{i=1}^{N} W_i' F_T \hat{\varepsilon}_i \hat{\varepsilon}_i{}' F_T{}' W_i$$

The resulting estimator finally is

$$\hat{\gamma}^{GMM} = \left(X W \hat{V}^{-1} W' X\right)^{-1} X' W \hat{V}^{-1} W' y .$$

3.2.3. The Blundell-Bond estimator

The GMM estimator which was suggested by Arellano and Bond (1991) is known to be rather inefficient when instruments are weak because making use of the information contained in differences only. In their 1998 paper Blundell and Bond suggest making use of additional level information beside the differences. The combination of moment restrictions for differences and levels results in an estimator which was called GMM-system-estimator by Arrellano and Bond.

There are T-2 othogonality restrictions in levels which are exploited. The observation t in levels

$$y_{it} = \rho y_{i,t-1} + x_{it}'\beta + \alpha_i + \varepsilon_{it}$$

will be used for the estimation, where differences are used as valid instruments (again assuming x being at least predetermined).

Take for example the last observation T:

$$y_{iT} = \rho y_{i,T-1} + x_{iT}'\beta + \alpha_i + \varepsilon_{iT}$$

where use is made of the instruments $dy_{i,1}, dy_{i,2}, ..., dy_{i,T-1}, dx_{i1}', dx_{i2}', ..., dx_{i,T}'$.

The matrices used for estimation are then defined as:

$$y_i = \begin{bmatrix} y_{i3} - y_{i,2} \\ y_{i4} - y_{i,3} \\ \vdots \\ y_{i,T} - y_{i,T-1} \\ y_{i3} \\ \vdots \\ y_{iT} \end{bmatrix} \qquad X_i = \begin{bmatrix} y_{i2} - y_{i1} & x'_{i3} - x'_{i2} \\ y_{i3} - y_{i,2} & x'_{i4} - x'_{i3} \\ \vdots & \vdots \\ y_{i,T-1} - y_{i,T-2} & x'_{iT} - x'_{i,T-1} \\ y_{i,2} & x'_{i2} \\ \vdots & \vdots \\ y_{i,T-1} & x'_{iT} \end{bmatrix}$$

$$\hat{X} = (y_{-1}, X), \qquad \gamma' = (\rho, \beta'), \qquad W = \left(W_1{}', W_2{}', \ldots, W_N{}'\right)'$$

$$W_i^D = \begin{bmatrix} [y_{i1}, x'_{i1}, x'_{i2}] & 0 & \cdots & 0 \\ 0 & [y_{i1}, y_{i2}, x'_{i1}, x'_{i2}, x'_{i3}] & \cdots & 0 \\ 0 & 0 & \ddots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & [y_{i1}, y_{i2}, \ldots y_{i,T-2}, x'_{i1}, x'_{i2}, \ldots, x'_{iT-1}] \end{bmatrix}$$

$$W_i^L = \begin{bmatrix} [dy_{i2}, dx'_{i2}, dx'_{i3}] & 0 & \cdots & 0 \\ 0 & [dy'_{i2}, dy'_{i3}, dx'_{i2}, dx'_{i3}, dx'_{i4}] & \cdots & 0 \\ 0 & 0 & \ddots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & [dy'_{i2}, \ldots dy_{i,T-2}, dx'_{i2}, \ldots, dx'_{iT-1}] \end{bmatrix}$$

$$W_i = \begin{bmatrix} W_i^D & 0 \\ 0 & W_i^L \end{bmatrix}$$

The first-step estimator makes use of a covariance matrix taking this autocorrelation into account enlarged for the level equations.

$$V = W'GW = \sum_{i=1}^{N} W_i' G_T W_i$$

where $G = \left(I_N \otimes G^{D,L}{}' \right)$ and $G^D = \begin{bmatrix} 2 & -1 & & 0 \\ -1 & 2 & \ddots & \\ & \ddots & \ddots & -1 \\ 0 & & -1 & 2 \end{bmatrix}$.

$$G^{L} = \begin{bmatrix} 1 & 0 & & 0 \\ 0 & 1 & \ddots & \\ & \ddots & \ddots & 0 \\ 0 & & 0 & 1 \end{bmatrix}$$

$$G^{D,L} = \begin{bmatrix} W_i^{D} & 0 \\ 0 & W_i^{L} \end{bmatrix}$$

The two-step GMM estimator uses the residuals of the first-step estimation to estimate the covariance matrix as suggested by White (1980):

$$\hat{V} = \sum_{i=1}^{N} W_i' F_T \hat{\varepsilon}_i \hat{\varepsilon}_i' F_T' W_i$$

The resulting estimator finally is

$$\hat{\gamma}^{GMM-SYS} = \left(XW\hat{V}^{-1}W'X \right)^{-1} X'W\hat{V}^{-1}W'y \, .$$

3.2.4. The direct bias correction

Making use of the asymptotic bias expression derived by Nickell, Kiviet (1995) proposed a direct bias correction method. The basic idea is the approximation of the unknown bias by a two-step procedure. While in the first round empirical estimates are derived, in the second step by a plug-in-procedure an empirical estimation of the bias is derived which leads to a correction of the biased fixed effects estimator.

The motivation for the direct correction lies in the well known fact, that the Least-Squares-Dummy-Variable estimator (LSDV) is biased but has a variance much smaller compared to instrumental variables estimators, like the Anderson-Hsiao estimator.

Kiviet derives the following approximation for the expected bias:

$$\hat{\gamma} - \gamma = -\sigma_{\varepsilon}^{2} (D)^{-1} \begin{pmatrix} \frac{N}{T} (j_t' C j_t) \left(2q - \overline{W}' A \overline{W} (D)^{-1} q \right) \\ + tr\left(\overline{W}' (I_N \otimes A_T C A_T) \overline{W} (D)^{-1} \right) q \\ + \overline{W}' (I_N \otimes A_T C A_T) \overline{W} (D)^{-1} q + \sigma_{\varepsilon}^{2} N q' (D)^{-1} q \\ \times \left(-\frac{N}{T} (j_t' C j_t) tr(C' A_T C) + 2 tr(C' A_T C A_T C) \right) q \end{pmatrix}$$

with

$$D = \overline{W}' A \overline{W} + \sigma_{\varepsilon}^{2} N tr(C' A_T C) q q'$$

$$A\overline{W} = E(AW)$$

$$W = [y_{-1} \ \vdots \ X]$$

$$\gamma' = [\rho, \beta']$$

$$A_T = I_T - \frac{1}{T} j_T j_T{}'$$

$$A = I_N \otimes A_T$$

$$C = \begin{bmatrix} 0 & \cdot & \cdot & \cdot & \cdot & \cdot & 0 \\ 1 & 0 & & & & & \cdot \\ \lambda & 1 & 0 & & & & \cdot \\ \lambda^2 & & 1 & \cdot & & & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & & \cdot \\ \cdot & & & \cdot & \cdot & \cdot & \cdot \\ \lambda^{T-2} & \cdot & \cdot & \cdot & \lambda & 1 & 0 \end{bmatrix}$$

where λ is a first round estimate of ρ, the parameter of the lagged endogenous variable and

$q = (1,0,...,0)'$ with dimension k+1, where k is the number of exogenous variables.

Kiviet suggests the use of a consistent first round estimator, e.g. the Anderson-Hsiao instrumental variable estimator. In our simulations we will make use of the LSDV-estimator and the first step GMM estimator in the first round respectively.

3.2.5. An alternative bias correction method

Based on the estimator proposed by Kiviet (1995) Hansen (2001) suggested an alternative bias correction method. The basic idea is the approximation of the unknown bias by making use of the first step biased fixed effects estimator. As the starting point the biased fixed effects estimators $\hat{\rho}$ and $\hat{\beta}$ are obtained. The asymptotic bias expression is then approximated by making use of first round regression results.

The term

$$\left(\underset{N\to\infty}{p\lim} \frac{1}{NT} y^{*\prime}_{-1} \varepsilon^{*}_{-1} \right)^{-1}$$

is approximated by

$$\frac{NT}{\varepsilon^{*\prime}_{-1} \varepsilon^{*}_{-1}}$$

and

$$\frac{-\sigma_{\varepsilon}^{2}}{T(1-\rho)}\left(1-\frac{1}{T}\frac{(1-\rho')}{T(1-\rho)}\right)$$

by

$$\frac{-\hat{\sigma}_{\varepsilon}^{2}}{T(1-\hat{\rho})}\left(1-\frac{1}{T}\frac{(1-\hat{\rho}')}{T(1-\hat{\rho})}\right)$$

using $\varepsilon *_{i,t-1} = y*_{-1} - X*\hat{\beta}$ and $\hat{\beta} = (X*' X*)^{-1} X*' y*_{-1}$.

Now the parameter $\hat{\rho}_c$ is estimated, which minimises the quadratic difference between the unknown bias and the approximated bias on the basis of the first step fixed effects estimation.

$$\hat{\rho}_c : \underset{\rho}{Min}\left[Bias - B\hat{i}as\right]^2$$

$$\hat{\rho}_c : \underset{\rho}{Min}\left[(\hat{\rho}-\rho) - \frac{NT}{\varepsilon*'_{-1}\,\varepsilon*_{-1}}\frac{-\hat{\sigma}_{\varepsilon}^{2}}{T(1-\rho)}\left(1-\frac{1}{T}\frac{(1-\rho')}{T(1-\rho)}\right)\right]^2$$

The problem has to be solved iteratively. Because the unknown parameter ρ is expected to be in a rather narrow interval, a grid-search is applied.

By making use of the bias corrected parameter of the lagged endogenous variable $\hat{\rho}_c$ the bias corrected estimator for the exogenous variables $\hat{\beta}_c$ is estimated making again use of the first step regression results:

$$\underset{N\to\infty}{p\lim}(\hat{\beta}-\beta) = -\underset{N\to\infty}{p\lim}\left((X*' X*)^{-1} X*' y*_{-1}\right)\underset{N\to\infty}{p\lim}(\hat{\rho}-\rho)$$

$$\hat{\beta}_c = \hat{\beta} + \hat{\beta}(\hat{\rho} - \hat{\rho}_c).$$

4. Monte Carlo study

The following Monte Carlo study compares the behaviour of the different discussed estimators under different circumstances. We vary the size of the data set as well as various key parameters in the simulation setting. Beside analysing the bias of the estimators the study enables to assess the reliability of the estimated standard deviations. The Root Mean Square Error (RMSE) criterion is used to assess the efficiency of the estimators.

All simulations were carried out using estimation routines written in the Interactive Matrix Language (IML) contained in the SAS software package by the author.

4.1. The case of a strictly exogenous explanatory variable

The simulation is based on the following model:

$$y_{it} = \rho y_{i,t-1} + x'_{it-1}\beta + \alpha_i + \varepsilon_{it} \qquad \text{where } \varepsilon_{it} \sim N(0,\sigma_\varepsilon^2) \text{ and } \alpha_i \sim N(0,\sigma_\alpha^2)$$

$$x_{it} = \eta\, x_{i,t-1} + \zeta_{it} \qquad \text{where } \zeta_{it} \sim N(0,\sigma_\zeta^2)$$

$$\gamma = (\rho\ \beta)' \qquad \beta = 1 \quad \rho = \{0.1, 0.5, 0.9\} \qquad \eta = \{0.1, 0.5, 0.9\}$$

Because all the estimators, except the simple pooled estimator in levels, allow for a possible correlation of the individual effects a_i and the explanatory variable $\bar{x}_i$, we do not consider such correlation in the simulation. For each individual the first 20 simulated data were dropped.

The following tables contain simulation results. The means of the estimators $\left(\overline{X}\right)$ as well as the mean of the estimated standard deviations $\overline{s\hat{t}d}$, the empirical standard deviation (*std*) of the estimators in the simulation runs as well as the RMSE. Table 1 contains the results of the simulation for 100 individuals.

We consider the following estimators:

γ^{Pooled}	Pooled estimator
γ^{LSDV}	Least Squares Dummy Variable Model (LSDV)
$\gamma^{AH,L}$	Anderson-Hsiao estimator using lagged levels as instruments (AH,L)
$\gamma^{AH,D}$	Anderson-Hsiao estimator using lagged differences as instruments
$\gamma^{BC,H}$	Bias corrected estimator using proposed by Hansen (BC,H)
$\gamma^{BC,K1}$	Bias corrected estimator using proposed by Kiviet, using LSDV in first step (BC,K1)
$\gamma^{BC,K2}$	Bias corrected estimator using proposed by Kiviet, using GMM1 in first step (BC,K2)
γ^{GMM1}	First step Arellano-Bond estimator (GMM1)
γ^{GMM2}	Second step Arellano-Bond estimator using estimated covariance matrix (GMM2)
γ^{SYS1}	First step system-estimator using level and differences as instruments proposed by Blundell and Bond (SYS1)
γ^{SYS2}	Second step system-estimator using estimated covariance matrix proposed by Blundell and Bond (SYS2)

Table 1: *Simulation results, T*=10, *N*=100 , $\sigma_{\varepsilon}^{2}=1$, $\rho=0.5$, $\eta=0.5$

	ρ				β			
	$\overline{X}$	$\overline{s\hat{t}d}$	std	$RMSE$	$\overline{X}$	$\overline{s\hat{t}d}$	std	$RMSE$
γ^{Pooled}	0.708	0.014	0.021	0.209	0.655	0.029	0.043	0.348
γ^{LSDV}	0.429	0.021	0.021	0.074	1.019	0.034	0.033	0.038
$\gamma^{AH,L}$	0.498	0.178	0.083	0.082	1.001	0.073	0.05	0.05
$\gamma^{AH,D}$	0.96	63.08	5.489	5.494	1.101	12.716	1.103	1.105
$\gamma^{BC,H}$	0.497	0.021	0.022	0.022	1.002	0.034	0.032	0.032
$\gamma^{BC,K1}$	0.481	0.021	0.022	0.029	0.959	0.034	0.031	0.051
$\gamma^{BC,K2}$	0.484	0.021	0.022	0.027	0.957	0.034	0.031	0.053
γ^{GMM1}	0.475	0.033	0.033	0.042	0.986	0.059	0.063	0.064
γ^{GMM2}	0.474	0.013	0.036	0.044	0.984	0.024	0.066	0.068
γ^{SYS1}	0.504	0.029	0.031	0.031	1.006	0.05	0.054	0.054
γ^{SYS2}	0.504	0.004	0.031	0.031	1.007	0.007	0.053	0.054

To ease the comparison of the results for the different estimators, the following figure illustrates the bias and the RMSE for the estimators.

Fig. 1: Bias and Root Mean Square Error of ρ, N=100[4]

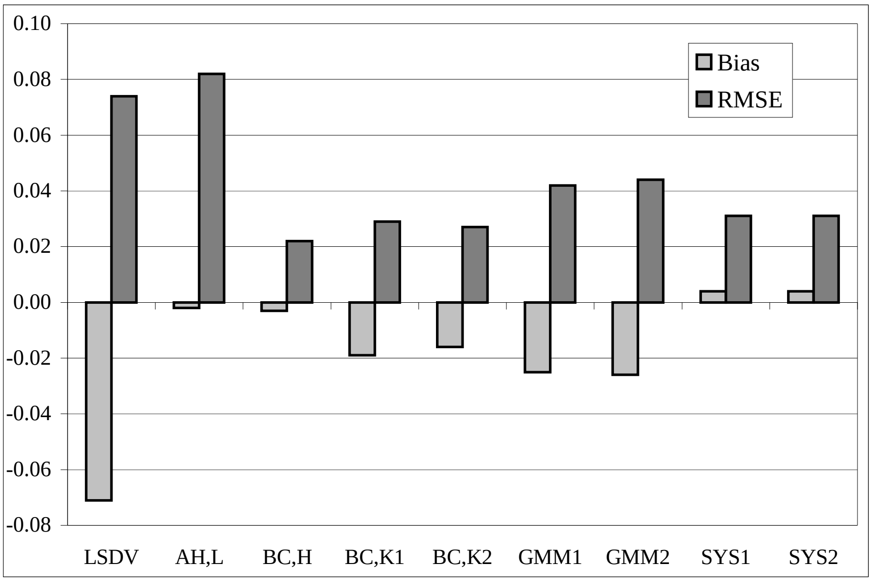
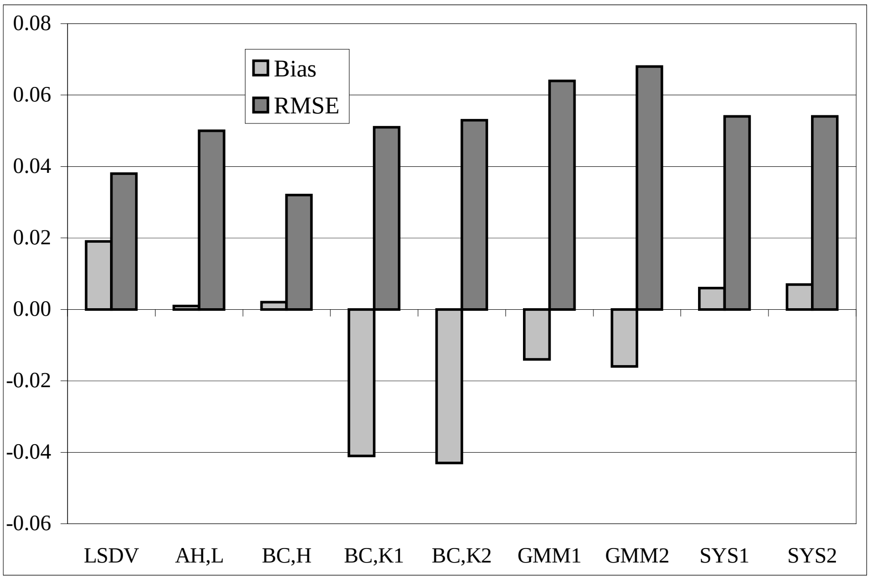


Fig. 2: Bias and Root Mean Square Error of β, N=100



We find that in the simulation with 100 individuals the bias corrected estimator proposed by Hansen performed best according to the RMSE-criterion. The estimator is practically unbiased. The Anderson-Hsiao estimator using lagged levels as instruments is in average practically unbiased but according to the large standard deviation, rather inefficient. The system-estimator clearly outperforms the GMM estimator using only lagged levels as instrument. Quite surprisingly the bias corrected estimator proposed by Kiviet as well as

---

4 The estimator proposed by Anderson and Hsiao making use of differences is clearly outperformed and not shown in the figures to ease comparability of the remaining estimators.

the GMM estimators have a downward bias for the lagged endogenous variable as well as for the exogenous variable. Table 2 contains the simulation results for 1000 individuals.

Table 2: Simulation results, T=10, N=1000 , $\sigma_\varepsilon^2 = 1$, $\rho = 0.5$, $\eta = 0.5$

	ρ				β			
	$\overline{X}$	$\overline{\hat{std}}$	std	$RMSE$	$\overline{X}$	$\overline{\hat{std}}$	std	$RMSE$
γ^{Pooled}	0.710	0.004	0.007	0.21	0.649	0.009	0.014	0.351
γ^{LSDV}	0.431	0.007	0.007	0.069	1.018	0.011	0.01	0.021
$\gamma^{AH,L}$	0.498	0.055	0.022	0.022	1.000	0.023	0.016	0.016
$\gamma^{AH,D}$	0.501	0.104	0.102	0.101	1.002	0.038	0.035	0.035
$\gamma^{BC,H}$	0.500	0.007	0.007	0.007	1.001	0.011	0.010	0.010
$\gamma^{BC,K1}$	0.480	0.007	0.007	0.021	0.961	0.011	0.010	0.040
$\gamma^{BC,K2}$	0.484	0.007	0.007	0.018	0.958	0.011	0.010	0.043
γ^{GMM1}	0.498	0.012	0.012	0.012	0.998	0.021	0.020	0.020
γ^{GMM2}	0.498	0.011	0.013	0.013	0.998	0.02	0.021	0.021
γ^{SYS1}	0.501	0.01	0.011	0.011	1.001	0.018	0.016	0.016
γ^{SYS2}	0.502	0.008	0.01	0.01	1.001	0.015	0.017	0.017

Fig. 3: Bias and Root Mean Square Error of ρ, N=1000

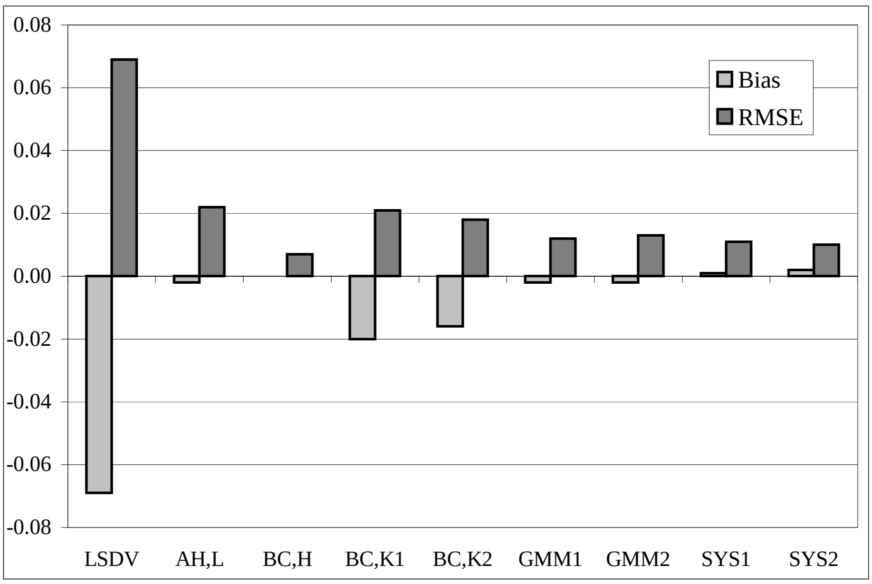
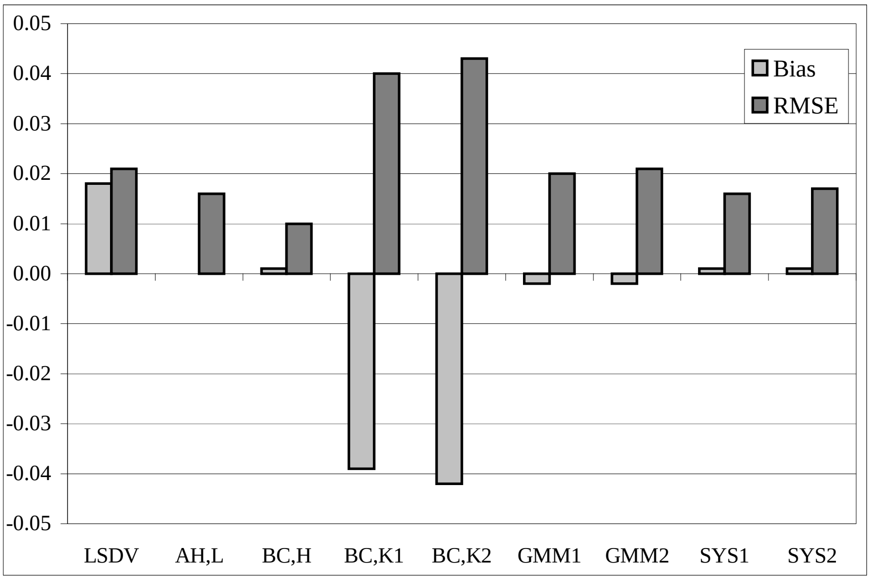


*Fig. 4: Bias and Root Mean Square Error of β, N=*1000



When comparing the simulation results for the case of 1000 individuals the results resemble but the GMM und system estimator improve considerably and perform better than the bias corrected estimators using the correction proposed by Kiviet. The improvement of the instrumented estimators of course had to be expected due to the well known large sample properties of the GMM methods.

Now we turn to the assessment of the standard deviation of the estimators which are important for statistical inference.

The following figure shows the comparison of the average estimated standard deviation and the empirical standard deviation of the simulation parameters.

*Fig. 5: Measures of variation, ρ, N=*1000

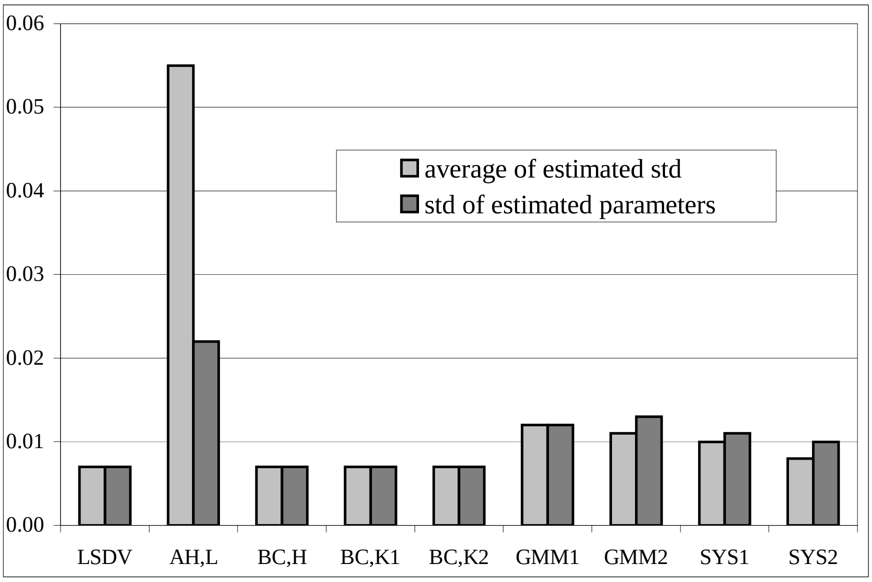
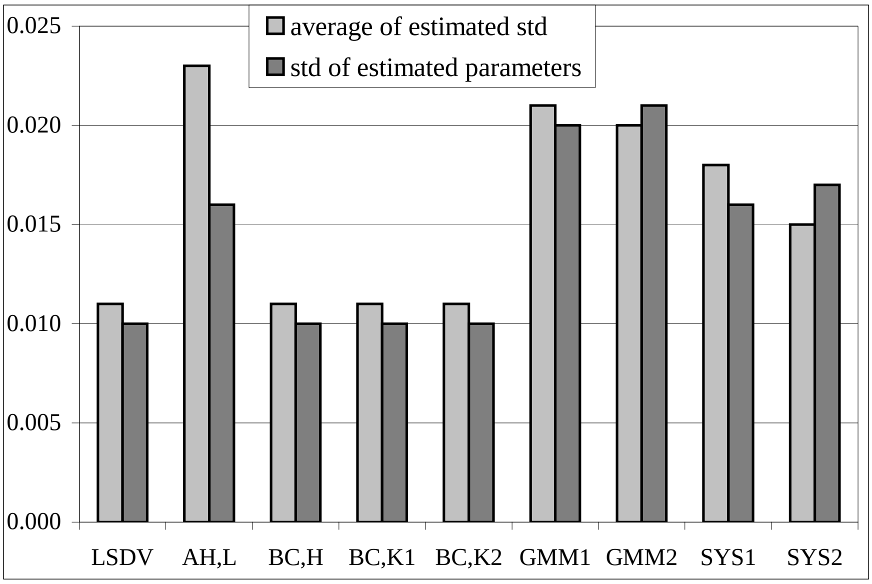


Fig. 6: *Measures of variation, β, N*=1000



The diagrams show that the estimated standard deviation of the bias corrected estimators is very close to the empirical standard deviation of the estimators in the simulation runs. While the bias corrected estimators seem more efficient than the GMM estimators, the system-estimators are almost as efficient. But the results also demonstrate that the estimated standard deviations of the two step GMM and system estimators are strongly downward biased and not reliable.[5]

To sum up the results for the case of an exogenous regressor, we find that bias corrected methods, especially the method proposed by Hansen, seem superior. Especially for smaller samples the GMM and system estimators are less efficient, while the difference narrows when turning to large samples. When comparing the GMM estimator proposed by Arellano and Bond with the system estimator proposed by Blundell and Bond we find that the system estimator is superior, but both second step estimators making use of the estimated covariance matrix face the problem of unreliable standard errors.

4.2. The case of a predetermined endogenous explanatory variable

In this section we want to analyse the performance of the various estimators in the case of a predetermined endogenous explanatory variable *x*.

5 See Windmeijer (2000) for a discussion of these fact and a proposed small sample correction for the standard deviation. For the case of 1000 individuals the small sample correction is neglectable.

The simulation is based on the following model:

$$y_{it} = \rho \, y_{i,t-1} + \beta \, x_{it-1} + \alpha_i + \varepsilon_{it} \quad \text{where} \quad \varepsilon_{it} \sim N(0, \sigma_\varepsilon^2) \text{ and } \alpha_i \sim N(0,1)$$

$$x_{it} = \eta \, x_{i,t-1} + \delta \, y_{it-1} + \zeta_i + \upsilon_{it} \quad \text{where} \quad \upsilon_{it} \sim N(0, \sigma_\upsilon^2) \text{ and } \zeta_i \sim N(0,1)$$

While the instrumented estimators in this setting still are consistent due to the choice of instruments taking the predetermination into account the bias corrected estimators will become inconsistent. Because the bias corrected estimators proofed to be superior in the case of exogenous variables, we want to assess the trade off between the efficiency and the resulting bias. The adequate measure to take both into account is the RMSE-criterion.

Table 3: Simulation results, T=10, *N*=1000 , $\sigma_\varepsilon^2 = 1$, $\sigma_\upsilon^2 = 1$, $\rho = 0.5$, $\eta = 0.5$, $\delta = 0.1$

	ρ				β			
	$\overline{X}$	$\overline{\hat{std}}$	*std*	*RMSE*	$\overline{X}$	$\overline{\hat{std}}$	*std*	*RMSE*
γ^{Pooled}	0.711	0.004	0.007	0.211	0.634	0.009	0.014	0.367
γ^{LSDV}	0.414	0.007	0.007	0.086	0.996	0.011	0.012	0.013
$\gamma^{AH,L}$	0.497	0.103	0.038	0.038	0.999	0.043	0.021	0.021
$\gamma^{AH,D}$	0.507	0.081	0.085	0.085	1.002	0.035	0.036	0.036
$\gamma^{BC,H}$	0.486	0.007	0.007	0.016	0.973	0.011	0.012	0.029
$\gamma^{BC,K1}$	0.463	0.007	0.007	0.038	0.941	0.011	0.011	0.060
$\gamma^{BC,K2}$	0.467	0.007	0.007	0.034	0.937	0.011	0.012	0.064
γ^{GMM1}	0.491	0.017	0.018	0.020	0.985	0.033	0.034	0.037
γ^{GMM2}	0.491	0.016	0.019	0.021	0.985	0.031	0.035	0.038
γ^{SYS1}	0.503	0.011	0.012	0.012	1.004	0.019	0.02	0.02
γ^{SYS2}	0.502	0.009	0.01	0.011	1.003	0.016	0.019	0.019

As could have been expected, in the case of predetermined endogenous explanatory variable, the simulation results are in favour of the instrumental variable estimators. The RMSE of the direct bias corrected estimator (BC,H) is about 50% larger than the system-estimator. Still the bias corrected estimator outperforms the GMM estimators and the estimators proposed by Anderson and Hsiao.

5. An empirical application of dynamic panel data estimation

In this section we want to apply the discussed estimators to a large firm level panel data set and return to the simulation results, especially to assess the case of endogeneity of x in face of the empirical results. The model to estimate is a dynamic Q investment function.[6] The variable to be explained is the investment ratio, while the explanatory variable is from the beginning of period Q, which is hence the case of a predetermined variable.

5.1. Deutsche Bundesbank's Corporate Balance Sheet Statistics

The empirical analysis is based on the Deutsche Bundesbank' corporate balance sheet statistics.[7] This data base covers about 50,000 to 70,000 enterprises each year which represent about 4% of the total number of enterprises in Germany. In the context of its rediscount-lending operations the Bundesbank collects the financial statements of firms using trade bills to assess the creditworthiness of the bill-presenting firm.[8]

Because the sample is biased towards larger enterprises about 75% of the total turn over of the corporate sector in western Germany is covered. The time period covered by our sample is from 1987 to 1998.

Starting with a very large data set the number of observations decreases considerably through incomplete balance sheets, outlier control and balancing. Especially the need to use the detailed schedule of fixed asset movements (Anlagespiegel) to apply our algorithm for calculating the capital stock at replacement costs shrinks the available data further. Because we expect sectoral differences between the manufacturing, construction and traders to lead to unreliable results when pooling all the data from all sectors, we focus in the following on manufacturing firms only. This leads to 1,371 firms contained in the final estimations.

5.2. Empirical results

In this section we apply the discussed estimators to the manufacturing data file described above. We estimate a dynamic investment function including Q as the regressor beside the lagged investment ratio.[9]

$$\left(\frac{I}{K}\right)_{it} = a_i + \rho\left(\frac{I}{K}\right)_{i,t-1} + \beta\, Q_{it-1} + \varepsilon_{it}$$

---

6 For an exact description of the variables see Behr/Bellgardt (2002).

7 For an overview of empirical work based on this data base see Stöss (2001).

8 See Deutsche Bundesbank (1998) and Stöss (2001).

9 For a sorrow variable description see Behr/Bellgardt (2002).

The following table contains the empirical results.

Table 4: Empirical results of the dynamic Q-investment function

	ρ	$std(\rho)$	$t(\rho)$	β	$std(\beta)$	$t(\beta)$
γ^{Pooled}	0.243	0.008	29.73	0.015	0.001	17.73
γ^{LSDV}	0.072	0.009	8.47	0.106	0.003	39.84
$\gamma^{AH,L}$	0.107	0.016	6.73	0.141	0.004	36.4
$\gamma^{AH,D}$	0.075	0.022	3.48	0.174	0.005	37.75
$\gamma^{BC,H}$	0.176	0.009	20.58	0.104	0.003	38.86
$\gamma^{BC,K1}$	0.145	0.016	8.87	0.096	0.008	11.74
$\gamma^{BC,K2}$	0.167	0.010	17.43	0.101	0.006	16.35
γ^{GMM1}	0.128	0.016	7.89	0.043	0.003	16.71
γ^{GMM2}	0.165	0.008	20.26	0.042	0.001	29.93
γ^{SYS1}	0.106	0.009	12.44	0.102	0.003	38.06
γ^{SYS2}	0.109	0.009	12.75	0.101	0.003	37.96

Leaving aside the pooled estimator which is clearly not appropriate and the least square dummy variable estimator known to be biased, we still find a large amount of variation in the estimates. The estimates of ρ range from 0.075 for the Anderson-Hsaio estimator using lagged differences as instruments to 0.176 for the bias corrected estimator proposed by Hansen. The parameters for Q show some variation, too. The lowest estimate is obtained using the GMM estimators proposed by Arellano and Bond while the highest parameter value results for the Anderson-Hsiao estimator using lagged differences as instruments.

We now turn to the problem of potential endogeneity of the Q-variable. As was found in the simulation results this endogeneity of Q, even in the case of predetermination, could lead to the conclusion that the system estimator proposed by Blundell and Bond should be favoured. To judge the seriousness of the problem in the empirical data, we estimate the following regression equation for Q:

$$Q_{it} = a_i + \rho\left(\frac{I}{K}\right)_{i,t-1} + \beta\, Q_{it-1} + \varepsilon_{it}$$

Table 5: Empirical results for the regression of Q on lagged values

	ρ	$std(\rho)$	$t(\rho)$	β	$std(\beta)$	$t(\beta)$
γ^{Pooled}	0.23	0.009	25.19	0.008	0.001	9.12
γ^{LSDV}	0.055	0.01	5.5	0.011	0.001	7.06
$\gamma^{AH,L}$	0.159	0.022	7.08	0.000	0.002	-0.09
$\gamma^{AH,D}$	0.135	0.028	4.77	0.001	0.002	0.33
$\gamma^{BC,H}$	0.202	0.01	19.84	0.011	0.002	7.41
$\gamma^{BC,K1}$	0.156	0.02	7.92	-0.011	0.004	-2.63
$\gamma^{BC,K2}$	0.171	0.016	11.00	-0.008	0.003	-2.38
γ^{GMM1}	0.166	0.02	8.49	-0.018	0.005	-3.95
γ^{GMM2}	0.209	0.013	16.54	-0.001	0.003	-0.49
γ^{SYS1}	0.093	0.01	9.27	0.010	0.001	6.62
γ^{SYS2}	0.130	0.009	14.5	0.018	0.003	5.89

We find that Q is significantly related to its lagged value, the parameter is in average about 0.15. The lagged investment ratio does not seem to influence Q significantly. The different estimators are in average about 0 with varying signs.

This empirical finding leads us to the conclusion that the problem of endogeneity in the data is not very serious. Hence the use of a GMM or system estimator instead of the somewhat superior direct bias corrected estimators is not indicated by the empirical findings.[10]

The final estimate making use of the bias corrected estimator based on the quadratic minimisation is:

$$\left(\frac{I}{K}\right)_{it} = a_i + \underset{(20.58)}{0{,}176}\left(\frac{I}{K}\right)_{i,t-1} + \underset{(38.86)}{0{,}104}\; Q_{i,t-1} + \hat{\varepsilon}_{it}$$

10 When using the investment ratio of the same period as the variables used to construct Q, the results resemble. The parameters of the investment ratio have varying signs and have in average a very small negative value with low t-values.

We find that Q influences the investment decision of German manufacturing firms most significantly. The actual investment ratio is also depends significantly positive on last year's investment ratio. This finding shows that the dynamic estimation in this empirical case is appropriate.

6. Conclusion

In our analysis we discussed several linear dynamic panel data estimators proposed in the literature. It is a well known fact that the natural choice, the least squares dummy variable estimator is biased in the context of dynamic estimation. The estimators taking into account the resulting bias can be grouped broadly into the class of instrumental estimators and the class of direct bias corrected estimators. Because there are now various estimators available, the applied researcher faces the problem of choosing among them.

While in empirical applications instrumental estimators are widely used, simulation results seem to favour direct bias corrected methods. But to our knowledge there is no comparison of up to date instrumental estimators, like the system estimator proposed by Blundell and Bond (1998) and direct bias corrected methods (Kiviet 1995, Hansen 2001).

One special feature of the direct bias corrected methods is that they rely on the assumption of exogenous regressors. In the case of estimating investment functions based on balance sheet data, this assumption can be expected to be violated. The case of a predetermined but endogenous regressor is therefore also assessed by the means of Monte Carlo simulations.

The simulation results clearly favour the direct bias corrected estimators, especially the estimator proposed by Hansen (2001). The superiority of these estimators decreases with growing numbers of individuals in the simulation. This is the well known fact of large sample properties of the GMM-methods. Turning to the case of endogenous predetermined regressors, the system-estimator proposed by Blundell and Bond is unbiased and most efficient, while direct bias corrected estimators perform similar to the GMM-estimator proposed by Arellano and Bond.

Turning to the empirical comparison, we find that the different estimators lead to the same conclusions concerning the investment behaviour of German manufacturing firms based on the Deutsche Bundesbank's Corporate Balance Sheet Statistics. Investment is strongly positive dependent on lagged investment and Q. Nevertheless, in detail the differences of the estimated parameters are not negligible.

To analyse the potential problems caused by endogeneity in the empirical data, the influence of investment on Q was assessed by estimating a dynamic equation. The results do not indicate that the endogeneity in this empirical example is serious. Hence the use of direct corrected estimators as well as the system estimator seems appropriate. This conclusion is also supported by the resemblence of the results obtained by the direct bias corrected and the system estimator.

References

Anderson, T.W./Hsiao, C. (1981): Estimation of Dynamic Models with Error Components, Journal of the American Statistical Association, Vol. 76, p. 598-606.

Anderson, T.W./Hsiao, C. (1982): Formulation and estimation of dynamic models using panel data, Journal of Econometrics, Vol. 18, p. 47-82.

Arellano, M. (1989): A note on the Anderson-Hsiao estimator for panel data, Economics Letters, Vol. 31, p. 337-341.

Arellano, M./Bond, S.R. (1991): Some Tests of Specification for Panel Data: Monte Carlo Evidence and an Application to Employment Equations, Review of Economic Studies, Vol. 58, p. 277-297.

Baltagi, B.H. (2001): Econometric Analysis of Panel Data, , 2nd edition, Chichester.

Behr, A./Bellgardt, E. (2002): Dynamic Q-investment functions for Germany using panel balance sheet data and a new algorithm for the capital stock at replacement values, Discussion Paper 23/02, Deutsche Bundesbank.

Blundell, R./Bond, S. (1998): Initial conditions and moment restrictions in dynamic panel data models, Journal of Econometrics 87, p. 115-143.

Deutsche Bundesbank (1998): The Methodological Basis of the Deutsche Bundesbank's Corporate Balance Sheet Statistics, Monthly Report, October, p. 49-64.

Hansen, G. (2001): A bias-corrected least squares estimator of dynamic panel models, Allgemeines Statistisches Archiv, Vol. 85, p. 127-140.

Judson, R. A. / Owen, A. L. (1999): Estimating dynamic panel data models: a guide for macroeconomics, Economics Letters, Vol. 65, p. 9-15.

Kiviet, J.F. (1995): On bias, inconsistency and efficiency of various estimators in dynamic panel data models, Journal of Econometrics, Vol. 68, p. 53-78.

Mátyás, L. (ed.) (1999): Generalized Method of Moments Estimation, Cambridge University Press.

Mátyás, L./Harris D. (1999): Introduction to the Generalized Method of Moments Estimation, in: Mátyás, L. (ed.) (1999): Generalized Method of Moments Estimation, Cambridge University Press.

Mátyás, L./Sevestre, P. (1995): The Econometrics of Panel Data, Dordrecht.

Nickel, S.J. (1981): Biases in Dynamic Models with Fixed Effects, Econometrica, Vol. 49, p. 1417-1426.

Stöss, E. (2001): Deutsche Bundesbank’s Corporate Balance Sheet Statistics and Areas of Application, Schmollers Jahrbuch, Vol. 121, p. 131-137.

White, H. (1980): A Heteroskedasticity-Consistent Covariance Matrix Estimator and a Direct Test for Heteroskedasticity, in: Econometrica, Vol. 48, May 1980, p. 817-838.

Windmeijer, F. (2000): A Finite Sample Correction For The Variance of Linear Two-Step GMM Estimators, IFS-working paper, 00/19.

The following papers have been published since 2002:

January	2002	Rent indices for housing in West Germany 1985 to 1998	Johannes Hoffmann Claudia Kurz
January	2002	Short-Term Capital, Economic Transformation, and EU Accession	Claudia M. Buch Lusine Lusinyan
January	2002	Fiscal Foundation of Convergence to European Union in Pre-Accession Transition Countries	László Halpern Judit Neményi
January	2002	Testing for Competition Among German Banks	Hannah S. Hempell
January	2002	The stable long-run CAPM and the cross-section of expected returns	Jeong-Ryeol Kim
February	2002	Pitfalls in the European Enlargement Process – Financial Instability and Real Divergence	Helmut Wagner
February	2002	The Empirical Performance of Option Based Densities of Foreign Exchange	Ben R. Craig Joachim G. Keller
February	2002	Evaluating Density Forecasts with an Application to Stock Market Returns	Gabriela de Raaij Burkhard Raunig
February	2002	Estimating Bilateral Exposures in the German Interbank Market: Is there a Danger of Contagion?	Christian Upper Andreas Worms
February	2002	The long-term sustainability of public finance in Germany – an analysis based on generational accounting	Bernhard Manzke

March	2002	The pass-through from market interest rates to bank lending rates in Germany	Mark A. Weth
April	2002	Dependencies between European stock markets when price changes are unusually large	Sebastian T. Schich
May	2002	Analysing Divisia Aggregates for the Euro Area	Hans-Eggert Reimers
May	2002	Price rigidity, the mark-up and the dynamics of the current account	Giovanni Lombardo
June	2002	An Examination of the Relationship Between Firm Size, Growth, and Liquidity in the Neuer Markt	Julie Ann Elston
June	2002	Monetary Transmission in the New Economy: Accelerated Depreciation, Transmission Channels and the Speed of Adjustment	Ulf von Kalckreuth Jürgen Schröder
June	2002	Central Bank Intervention and Exchange Rate Expectations – Evidence from the Daily DM/US-Dollar Exchange Rate	Stefan Reitz
June	2002	Monetary indicators and policy rules in the P-star model	Karl-Heinz Tödter
July	2002	Real currency appreciation in accession countries: Balassa-Samuelson and investment demand	Christoph Fischer
August	2002	The Eurosystem's Standing Facilities in a General Equilibrium Model of the European Interbank Market	Jens Tapking

August 2002	Imperfect Competition, Monetary Policy and Welfare in a Currency Area	Giovanni Lombardo
August 2002	Monetary and fiscal policy rules in a model with capital accumulation and potentially non-superneutral money	Leopold von Thadden
September 2002	Dynamic Q-investment functions for Germany using panel balance sheet data and a new algorithm for the capital stock at replacement values	Andreas Behr Egon Bellgardt
October 2002	Tail Wags Dog? Time-Varying Information Shares in the Bund Market	Christian Upper Thomas Werner
October 2002	Time Variation in the Tail Behaviour of Bund Futures Returns	Thomas Werner Christian Upper
November 2002	Bootstrapping Autoregressions with Conditional Heteroskedasticity of Unknown Form	Sílvia Gonçalves Lutz Kilian
November 2002	Cost-Push Shocks and Monetary Policy in Open Economies	Alan Sutherland
November 2002	Further Evidence On The Relationship Between Firm Investment And Financial Status	Robert S. Chirinko Ulf von Kalckreuth
November 2002	Genetic Learning as an Explanation of Stylized Facts of Foreign Exchange Markets	Thomas Lux Sascha Schornstein
December 2002	Wechselkurszielzonen, wirtschaftlicher Aufholprozess und endogene Realignmentrisiken *	Karin Radeck

* Available in German only.

December 2002	Optimal factor taxation under wage bargaining – a dynamic perspective	Erkki Koskela Leopold von Thadden
January 2003	Testing mean-variance efficiency in CAPM with possibly non-gaussian errors: an exact simulation-based approach	Marie-Claude Beaulieu Jean-Marie Dufour Lynda Khalaf
January 2003	Finite-sample distributions of self-normalized sums	Jeong-Ryeol Kim
January 2003	The stock return-inflation puzzle and the asymmetric causality in stock returns, inflation and real activity	Jeong-Ryeol Kim
February 2003	Multiple equilibrium overnight rates in a dynamic interbank market game	Jens Tapking
February 2003	A comparison of dynamic panel data estimators: Monte Carlo evidence and an application to the investment function	Andreas Behr

Visiting researcher at the Deutsche Bundesbank

The Deutsche Bundesbank in Frankfurt is looking for a visiting researcher. Visitors should prepare a research project during their stay at the Bundesbank. Candidates must hold a Ph D and be engaged in the field of either macroeconomics and monetary economics, financial markets or international economics. Proposed research projects should be from these fields. The visiting term will be from 3 to 6 months. Salary is commensurate with experience.

Applicants are requested to send a CV, copies of recent papers, letters of reference and a proposal for a research project to:

Deutsche Bundesbank
Personalabteilung
Wilhelm-Epstein-Str. 14

D - 60431 Frankfurt
GERMANY