pdynmc – An R-Package for Estimating Linear Dynamic Panel Data Models Based on Nonlinear Moment Conditions

Markus Fritsch
University of Passau

Andrew Adrian Yu Pua
WISE, Xiamen University

Joachim Schnurbus
University of Passau

Abstract

pdynmc is an R-package for GMM estimation of linear dynamic panel data models that are based on nonlinear moment conditions as proposed by Ahn and Schmidt (1995). In addition, we allow for iterated GMM in light of recent developments in its theory. This paper provides a description of the variety of options regarding instrument type, covariate type, estimation methodology, and general configuration. All functionality is demonstrated through the most popular firm-level dataset in panel data econometrics (Arellano and Bond 1991) and we relate to other software and packages.

Keywords: dynamic panels, generalized method of moments, iterated GMM, quadratic moment conditions, R.

1. Introduction

The formulation of the linear dynamic panel data model accounts for dynamics and unobserved individual-specific heterogeneity simultaneously. Due to the presence of unobserved individual-specific heterogeneity and lagged dependent variables, applying ordinary least squares including individual-specific dummy variables is inconsistent (see, e.g., Hsiao 2014).

A suitable alternative for obtaining parameter estimates in linear dynamic panel data models is deriving moment conditions (or population orthogonality conditions) from the model assumptions. The moment conditions may be linear (Anderson and Hsiao 1982; Holtz-Eakin, Newey, and Rosen 1988; Arellano and Bover 1995) or quadratic (Ahn and Schmidt 1995) in parameters and determine the natural instruments available for estimation. Usually, the number of moment conditions exceeds the number of parameters and the moment conditions need to be aggregated appropriately. This can be achieved by the generalized method of moments (GMM), where (weighted) linear combinations of the moment conditions are employed to obtain parameter estimates.

Theoretical results and evidence from Monte Carlo simulations in the literature suggest that incorporating the quadratic moment conditions proposed by Ahn and Schmidt (1995) may prove valuable for particular data generating processes (DGPs). A relevant example in practice is when the DGP exhibits high persistence and the linear moment conditions fail to identify the model parameters. In this situation, the quadratic moment conditions may still

provide identification (Bun and Kleibergen 2019; Bun and Sarafidis 2015; Gørgens, Han, and Xue 2019), and they are immediate by-products of imposing only the so-called standard assumptions (henceforth SA). Furthermore, we have also been working on other theoretical and practical aspects related to these quadratic moment conditions (Pua, Fritsch, and Schnurbus 2019a,b). These SA are the basis of the Arellano and Bond (1991) estimator, arguably the most popular default routine in dynamic panel data estimation.

Since the moment conditions employed in GMM estimation of linear dynamic panel data models are derived from model assumptions, a basic understanding of these assumptions is vital for setting up a plausible estimation routine. The methodological part of this paper briefly reviews the assumptions implied when using particular moment conditions in estimation. As a result, we add to the exposition in the **plm** vignette, where one of its functions **pgmm** is used to estimate linear dynamic panel data models as well. For further reading on the methodology, we suggest Fritsch (2019) and the book-length treatment in Croissant and Millo (2018).

2. Main contributions

Even though quadratic moment conditions were proposed more than 20 years ago, standard estimation routines are generally not available across statistical software. To the best of our knowledge, there is currently only the implementation provided by (Kripfganz 2019), called **xtdpdgmm**, for the commercial statistical software Stata (StataCorp 2015) that is explicitly designed to incorporate nonlinear moment conditions into GMM estimation. Furthermore, the current implementation in Stata requires version 13 or higher.

Our package **pdynmc** provides an implementation of GMM estimation of linear dynamic panel data models based on Holtz-Eakin *et al.* (1988), Arellano and Bover (1995), and Ahn and Schmidt (1995) moment conditions in the open source statistical software R (R Core Team 2019). In contrast to routines in the existing packages **plm** (Croissant and Millo 2008), and **panelvar** (Sigmund and Ferstl 2019), we allow the inclusion of quadratic moment conditions into the analysis. Nevertheless, the package **panelvar** allows the user to perform lag selection based on information criteria and structural analysis based on impulse response functions, while the package **plm** provides a variety of functions for the estimation of static and dynamic linear panel models, along with some random effects estimators and a number of different specification tests. Yet another package that tries to exploit the panel data structure by generalized estimating equations is **panelr** (Long 2019).

Furthermore, just like the function **pgmm** in the **plm** package, **pdynmc** allows the computation of one-step and two-step closed form GMM estimators and standard specification testing such as overidentifying restrictions tests, serial correlation tests, and Wald tests. These features are shared by other packages implemented in Gauss, Ox (Doornik, Arellano, and Bond 2012), R, and Stata.

More importantly, our contributed package already includes iterated GMM estimation as a default. Recent work by Hansen and Lee (2019) have shown the merits of the iterated GMM estimator and develop the theory under potential misspecification of moment conditions. Unfortunately, we have not incorporated their theory into the package but would eventually incorporate the theory to enable other researchers to use our package for simulation studies. For the moment, this feature is still "bleeding-edge" for the general audience. Furthermore,

having iterated GMM for dynamic panels available may help in further work to extend results found in, for example, Hwang and Sun (2018).

Regrettably, **pdynmc** is not as fast as **xtdpdgmm**. Some alternative R codes that have substantial speed improvements are available from Phillips and Han (2019) but lack the functionality of a full package. As a check of the current functionality of **pdynmc**, we estimated the model they used as a DGP for their simulations. We simulated only one realization of their DGP, which is a linear AR(1) dynamic panel model with one exogenous covariate with 1000 cross-sectional units, 40 time periods. We were able to estimate the model for one realization of their DGP in 1.5 to 2 minutes. In contrast, **pgmm** fails to even start. Although our computational time does not match Phillips and Han (2019), we believe this is a good start and will continue working on speed improvements of **pdynmc** in future releases of the package.

The structure of the paper is as follows. Section 3 briefly sketches the linear dynamic panel data model, states the underlying assumptions frequently used in the literature, and describes the moment conditions arising from the model assumptions. Section 4 covers GMM estimation of linear dynamic panel data models and illustrates the minimization criterion, and estimation in one, two, or multiple steps. Section 5 outlines the computation of standard errors, specification testing and overidentifying restrictions testing, and the testing of general linear hypotheses. Section 6 illustrates the estimation of linear dynamic panel data models with **pdynmc** for the data set of Arellano and Bond (1991) on adjustments of employment of firms located in the United Kingdom. Section 7 concludes and sketches functionality we plan to add to future releases of the package.

3. Linear dynamic panel data model

3.1. Model and standard assumptions

For a given dataset with cross section dimension n and time series dimension T, consider a typical linear dynamic panel data model of the following form:

$$y_{i,t} = \alpha y_{i,t-1} + \beta x_{i,t} + u_{i,t}, \quad i = 1, \dots, n; \ t = 2, \dots, T, \tag{1}$$

$$u_{i,t} = \eta_i + \varepsilon_{i,t}. \tag{2}$$

Here $y_{i,t}$ and $y_{i,t-1}$ denote the dependent variable and its lag, α is the lag parameter, and $x_{i,t}$ is a single covariate with corresponding slope coefficient β. The second equation shows that the model allows for a composite error term $u_{i,t}$ which can be separated into an unobserved individual-specific effect η_i and an idiosyncratic remainder component $\varepsilon_{i,t}$.

Combining the Equations (1) and (2) yields the single equation form of the model

$$y_{i,t} = \alpha y_{i,t-1} + \beta x_{i,t} + \eta_i + \varepsilon_{i,t}, \quad i = 1, \dots, n; \ t = 2, \dots, T. \tag{3}$$

For illustration purposes in this section, we only include one lag of the dependent variable, one covariate, and omit unobserved time-specific effects for simplicity of exposition and notational convenience. Extending the representation is straightforward and **pdynmc** can also accommodate AR(p) models. The initial time period is denoted by $t = 1$.

In the linear dynamic panel data literature, researchers have focsued on the following standard assumptions, henceforth SA, (see Ahn and Schmidt 1995):

$$\text{The data are independently distributed only across } i, \tag{4}$$

$$E(\eta_i) = 0, \quad i = 1, ..., n,$$

$$E(\varepsilon_{i,t}) = 0, \quad i = 1, ..., n; \ t = 2, ..., T,$$

$$E(\varepsilon_{i,t} \cdot \eta_i) = 0, \quad i = 1, ..., n; \ t = 2, ..., T,$$

$$E(\varepsilon_{i,t} \cdot \varepsilon_{i,s}) = 0, \quad i = 1, ..., n; \ t \neq s,$$

$$E(y_{i,1} \cdot \varepsilon_{i,t}) = 0, \quad i = 1, ..., n; \ t = 2, ..., T,$$

$$n \to \infty, \text{ while } T \text{ is fixed, such that } \frac{T}{n} \to 0.$$

We will assume that the SA holds for the rest of the paper. Next, we discuss how these SA help in the estimation of the parameters of the linear dynamic panel data model.

3.2. Moment conditions from standard assumptions

The unobserved individual-specific effects η_i may be eliminated from Equation (3) by taking first differences, i.e.,

$$\Delta y_{i,t} = \alpha \Delta y_{i,t-1} + \beta \Delta x_{i,t} + \Delta \varepsilon_{i,t}, \quad i = 1, \ldots, n; \ t = 2, \ldots, T. \tag{5}$$

Because the first difference of the lagged dependent variable $\Delta y_{i,t-1} = y_{i,t-1} - y_{i,t-2}$ and the first difference of the idiosyncratic remainder component $\Delta \varepsilon_{i,t} = \varepsilon_{i,t} - \varepsilon_{i,t-1}$ are not orthogonal, using least squares to estimate Equation (5) produces inconsistent estimators. The SA stated in Equation (4) provide a remedy.

In particular, Holtz-Eakin _et al._ (1988) (henceforth HNR) propose linear (in parameters) moment conditions

$$E(y_{i,s} \cdot \Delta u_{i,t}) = 0, \qquad t = 3, \ldots, T; \ s = 1, \ldots, t - 2. \tag{6}$$

Equation (6) provides $0.5(T-1)(T-2)$ moment conditions. Similar moment conditions can be derived from the covariate $x_{i,t}$, depending on its exogeneity classification. Endogenous, predetermined, and (strictly) exogenous covariates provide the following linear moment conditions, respectively:

$$E(x_{i,s} \cdot \Delta u_{i,t}) = 0, \qquad t = 3, \ldots, T, \qquad \text{where} \tag{7}$$

$$s = 1, \ldots, t - 2, \qquad \text{for} \quad x \ \text{endogenous},$$

$$s = 1, \ldots, t - 1, \qquad \text{for} \quad x \ \text{predetermined},$$

$$s = 1, \ldots, T, \qquad \text{for} \quad x \ \text{strictly exogenous}.$$

Ahn and Schmidt (1995) (henceforth AS) introduced the following $T - 3$ additional moment conditions, which are also implied by SA in Equation (4):

$$E(u_{i,t} \cdot \Delta u_{i,t-1}) = 0, \qquad t = 4, \ldots, T. \tag{8}$$

Rewriting the equation and expressing the moment conditions in terms of parameters and observables reveals that the AS moment conditions are quadratic in parameters.

Parameter estimates can be obtained by stacking together the sample analogues of the previously discussed moment conditions. As a result, we have m sample moment conditions $\overline{\mathbf{M}} = \frac{1}{n} \sum_{i=1}^{n} \mathbf{M}_i$. Now, suppose x_{it} is a predetermined covariate and we want to use all the corresponding HNR moment conditions in (6) and (7), and the AS moment conditions in (8). Now, consider the following moment conditions on the left panel and the corresponding vector of individual moment condition contributions on the right panel $\mathbf{M}_i$ to be available for estimation:

$$
\underbrace{\begin{pmatrix}
E(y_{i,1} \cdot \Delta u_{i,3}) \\
E(y_{i,1} \cdot \Delta u_{i,4}) \\
E(y_{i,2} \cdot \Delta u_{i,4}) \\
E(y_{i,1} \cdot \Delta u_{i,5}) \\
\vdots \\
E(y_{i,3} \cdot \Delta u_{i,5}) \\
\vdots \\
E(y_{i,T-2} \cdot \Delta u_{i,T}) \\
\hline
E(x_{i,1} \cdot \Delta u_{i,3}) \\
E(x_{i,2} \cdot \Delta u_{i,3}) \\
E(x_{i,1} \cdot \Delta u_{i,4}) \\
\vdots \\
E(x_{i,3} \cdot \Delta u_{i,4}) \\
\vdots \\
E(x_{i,T-1} \cdot \Delta u_{i,T}) \\
\hline
E(u_{i,4} \cdot \Delta u_{i,3}) \\
\vdots \\
E(u_{i,T} \cdot \Delta u_{i,T-1})
\end{pmatrix}}_{m \times 1}
=
\begin{pmatrix}
0 \\
0 \\
\vdots \\
0 \\
0 \\
\vdots \\
0 \\
0 \\
\vdots \\
0
\end{pmatrix},
\qquad
\underbrace{\mathbf{M}_i}_{m \times 1}
=
\begin{pmatrix}
y_{i,1} \cdot \widetilde{\Delta u}_{i,3} \\
y_{i,1} \cdot \widetilde{\Delta u}_{i,4} \\
y_{i,2} \cdot \widetilde{\Delta u}_{i,4} \\
y_{i,1} \cdot \widetilde{\Delta u}_{i,5} \\
\vdots \\
y_{i,3} \cdot \widetilde{\Delta u}_{i,5} \\
\vdots \\
y_{i,T-2} \cdot \widetilde{\Delta u}_{i,T} \\
\hline
x_{i,1} \cdot \widetilde{\Delta u}_{i,3} \\
x_{i,2} \cdot \widetilde{\Delta u}_{i,3} \\
x_{i,1} \cdot \widetilde{\Delta u}_{i,4} \\
\vdots \\
x_{i,3} \cdot \widetilde{\Delta u}_{i,4} \\
\vdots \\
x_{i,T-1} \cdot \widetilde{\Delta u}_{i,T} \\
\hline
\widetilde{u}_{i,4} \cdot \widetilde{\Delta u}_{i,3} \\
\vdots \\
\widetilde{u}_{i,T} \cdot \widetilde{\Delta u}_{i,T-1}
\end{pmatrix}.
\tag{9}
$$

The dashed lines separate the three different sets of moment conditions. Note that we also distinguish the unobservable $u_{i,t}$ and $\Delta u_{i,t}$ from what can be computed from the data, namely $\widetilde{u}_{i,t}$ and $\widetilde{\Delta u}_{i,t}$, in both panels of (9) for emphasis. Further consider decomposing the individual moment condition contributions (seen in the right panel of (9)) into $\boldsymbol{M}_i = \boldsymbol{Z}_i' \cdot \tilde{\boldsymbol{s}}_i$, where $\boldsymbol{Z}_i'$ denotes the transpose of a matrix that does not depend on parameter estimates, while the column vector $\tilde{\boldsymbol{s}}_i$ does. From the right panel of (9), we have:

$$
\boldsymbol{Z}'_i =
\left(
\begin{array}{ccccc|ccc}
y_{i,1} & 0 & \cdots & 0 & 0 & \cdots & & 0 \\
0 & y_{i,1} & & & & & & \\
& y_{i,2} & & & & & & \\
& 0 & & & & & & \\
\vdots & \vdots & & \vdots & \vdots & & & \vdots \\
& & & 0 & & & & \\
& & & y_{i,1} & & & & \\
& & & \vdots & & & & \\
0 & 0 & \cdots & y_{i,T-2} & 0 & \cdots & & 0 \\ \hline
x_{i,1} & 0 & \cdots & 0 & 0 & \cdots & & 0 \\
x_{i,2} & 0 & & & & & & \\
0 & x_{i,1} & & & & & & \\
& x_{i,2} & & & & & & \\
& x_{i,3} & & & & & & \\
& 0 & & & & & & \\
\vdots & \vdots & & \vdots & \vdots & & & \vdots \\
& & & 0 & & & & \\
& & & x_{i,1} & & & & \\
& & & \vdots & & & & \\
0 & 0 & \cdots & x_{i,T-1} & 0 & \cdots & & 0 \\ \hline
0 & & \cdots & 0 & 1 & 0 & \cdots & 0 \\
\vdots & & & \vdots & \vdots & \ddots & & \vdots \\
& & & & & & & 0 \\
0 & & \cdots & 0 & 0 & \cdots & 0 & 1
\end{array}
\right),
\qquad
\tilde{\boldsymbol{s}}_i =
\left(
\begin{array}{c}
\widetilde{\Delta u}_{i,3} \\
\widetilde{\Delta u}_{i,4} \\
\vdots \\
\widetilde{\Delta u}_{i,T} \\ \hline
\widetilde{u}_{i,4} \cdot \widetilde{\Delta u}_{i,3} \\
\widetilde{u}_{i,5} \cdot \widetilde{\Delta u}_{i,4} \\
\vdots \\
\widetilde{u}_{i,T} \cdot \widetilde{\Delta u}_{i,T-1}
\end{array}
\right)_{(2T-5)\times 1} .
$$

$$\underbrace{}_{m \times (2T-5)}$$

3.3. Moment conditions from extended assumptions

Another set of moment conditions beyond those implied by SA that is popular in theoretical and applied research are those conditions that may be derived from the assumption that

$$E(\Delta y_{i,t} \cdot \eta_i) = 0, \quad i = 1, \ldots, n. \tag{10}$$

This expression requires that the dependent variable and the unobserved individual-specific effects are constantly correlated over time for each individual. Deviations from the assumption are required to be unsystematic over both, the cross section and the time series dimension (see Section 6.5 in Arellano 2003, which also provides an empirically relevant example). Blundell, Bond, and Windmeijer (2001) state that if $\Delta y_{i,t}$ and η_i are uncorrelated, then it places restrictions on the relationship between $\Delta x_{i,t}$ and η_i. The latter is problematic from the perspective of the SA because researchers prefer not to restrict a priori the relationship between $x_{i,t}$ and η_i – which is the essence of the fixed-effects approach in dynamic panel data settings.

The assumption in (10) has been called 'constant correlated effects' by Bun and Sarafidis (2015), effect stationarity by (Kiviet 2007a), or mean stationarity by (Arellano 2003). From this assumption, we can derive $T-2$ additional Arellano and Bover (1995) (henceforth ABov):

$$E(\Delta y_{i,t-1} \cdot u_{i,t}) = 0, \quad t = 3, \ldots, T. \tag{11}$$

By rewriting these moment conditions, it can be shown that the ABov moment conditions encompass the nonlinear AS moment conditions and render the latter redundant for estimation (for a derivation see Fritsch 2019).

Additional ABov moment conditions can be derived from the covariate $x_{i,t}$, depending on the nature of the $x_{i,t}$:

$$E(\Delta x_{i,v} \cdot u_{i,t}) = 0, \qquad \text{where}$$
$$v = t - 1; \ t = 3, \ldots T, \quad \text{for} \quad x \quad \text{endogenous},$$
$$v = t; \ t = 2, \ldots, T, \quad \text{for} \quad x \quad \text{strictly exogenous} \quad \text{or} \quad x \quad \text{predetermined}.$$

Notice that no additional ABov moment conditions arise if $x_{i,t}$ is strictly exogenous. When using the HNR and ABov moment conditions to estimate the linear dynamic panel data model in Equation (3) with a predetermined explanatory variable, $\boldsymbol{M}_i$ is as follows:

$$
\boldsymbol{Z}'_i = \underbrace{\left(\begin{array}{cccc|ccc|ccc}
y_{i,1} & 0 & \cdots & 0 & 0 & \cdots & 0 & 0 & \cdots & 0 \\
0 & y_{i,1} & & & & & & & & \\
& y_{i,2} & & & & & & & & \\
\vdots & 0 & & \vdots & \vdots & & \vdots & \vdots & & \vdots \\
& & & 0 & & & & & & \\
& & & y_{i,1} & & & & & & \\
& & & \vdots & & & & & & \\
0 & \cdots & 0 & y_{i,T-2} & 0 & \cdots & 0 & 0 & \cdots & 0 \\
\hline
x_{i,1} & 0 & \cdots & 0 & 0 & \cdots & 0 & 0 & \cdots & 0 \\
x_{i,2} & 0 & & & & & & & & \\
0 & x_{i,1} & & & & & & & & \\
& x_{i,2} & & & & & & & & \\
& x_{i,3} & & & & & & & & \\
\vdots & 0 & & \vdots & \vdots & & \vdots & \vdots & & \vdots \\
& & & 0 & & & & & & \\
& & & x_{i,1} & & & & & & \\
& & & \vdots & & & & & & \\
0 & \cdots & 0 & x_{i,T-1} & 0 & \cdots & 0 & 0 & \cdots & 0 \\
\hline
0 & & \cdots & 0 & \Delta y_{i,2} & 0 & \cdots & 0 & 0 & \cdots & 0 \\
\vdots & & & \vdots & 0 & \Delta y_{i,3} & & \vdots & \vdots & & \vdots \\
& & & & \vdots & & \ddots & 0 & & & \\
0 & & \cdots & 0 & 0 & \cdots & 0 & \Delta y_{i,T-1} & 0 & \cdots & 0 \\
\hline
0 & & \cdots & 0 & 0 & \cdots & & 0 & \Delta x_{i,2} & 0 & \cdots & 0 \\
& & & & & & & & 0 & \Delta x_{i,3} & & \vdots \\
\vdots & & \vdots & \vdots & \vdots & & & \vdots & \vdots & & \ddots & 0 \\
0 & & \cdots & 0 & 0 & \cdots & & 0 & 0 & \cdots & 0 & \Delta x_{i,T}
\end{array}\right)}_{m \times (3T-5)},
$$

$$
\tilde{\boldsymbol{s}}'_i = \underbrace{\left(\widetilde{\Delta u}_{i,3}, \quad \widetilde{\Delta u}_{i,4}, \quad \cdots, \quad \widetilde{\Delta u}_{i,T} \mid \tilde{u}_{i,3}, \quad \tilde{u}_{i,4}, \quad \cdots, \quad \tilde{u}_{i,T}, \mid \tilde{u}_{i,2}, \quad \tilde{u}_{i,3}, \quad \cdots, \quad \tilde{u}_{i,T}\right)}_{1 \times (3T-5)}.
$$

4. GMM estimation

4.1. Minimization criterion

For a given sample, GMM estimation minimizes the aggregated squared distance of the sample moment conditions from zero and can be represented as

$$L_{\boldsymbol{W}}^2 = \overline{\boldsymbol{M}}' \cdot \boldsymbol{W} \cdot \overline{\boldsymbol{M}}. \tag{12}$$

The index of the $L_{\boldsymbol{W}}^2$-norm expresses that the norm depends on the weighting matrix $\boldsymbol{W}$ and the superscript indicates that the norm is a quadratic form. The $m \times m$ weighting matrix $\boldsymbol{W}$ guides the aggregation of the m sample moment conditions.

Stacking the $\boldsymbol{Z}_i'$ for all cross sectional observations horizontally yields $\boldsymbol{Z}' = (\boldsymbol{Z}_1', \dots, \boldsymbol{Z}_n')$ and concatenating the column vectors $\tilde{\boldsymbol{s}}_i$ yields the vector $\tilde{\boldsymbol{s}}$. From the notation developed in Section 3, where the sample moment conditions are decomposed into a vector $\tilde{\boldsymbol{s}}$ that depends on the parameter estimates and a matrix $\boldsymbol{Z}$ that does not. Plugging these two terms into Equation (12) gives:

$$L_{\boldsymbol{W}}^2 = \frac{1}{n^2} \cdot \tilde{\boldsymbol{s}}' \boldsymbol{Z} \cdot \boldsymbol{W} \cdot \boldsymbol{Z}' \tilde{\boldsymbol{s}}.$$

Minimizing $L_{\boldsymbol{W}}^2$ yields a GMM estimator $\widetilde{\boldsymbol{\theta}}$ which depends on a choice of $\boldsymbol{W}$. When estimating the linear dynamic panel data model in Equation (3) by GMM based on linear moment conditions only, numerical optimization is not required because closed form linear estimators are available and are programmed in **pdynmc**. In this situation, the closed form results are computed and stored along with the optimization results.

When nonlinear moment conditions are used, nonlinear optimization techniques are required to obtain coefficient estimates. By default, GMM estimation by **pdynmc** is based on numerical optimization. To initialize the optimization procedure, starting values are drawn for all parameter estimates from a uniform distribution over the interval $[-1, 1]$.

For the optimization procedure, we rely on the R-package **optimx** (Nash and Varadhan 2011; Nash 2014). All optimization routines implemented in **optimx** are available in **pdynmc**. Based on our experience, the Variable Metric method (Fletcher 1970; Nash 1990) seems to work satisfactorily in the estimation of linear dynamic panel data models. The Variable Metric method is named BFGS in **optimx** and serves as the default procedure in **pdynmc**. In contrast, the numerical optimization of the GMM objective function used in **xtdpdgmm** is based on a Gauss-Newton technique.

4.2. One-step, two-step, and iterated GMM estimation

In practice, GMM estimation is frequently carried out iteratively. In order to start the estimation process, an initial estimate of the weighting matrix $\widehat{\boldsymbol{W}}$ is required. Obviously, plugging in different weighting matrices into Equation (12) yields varying objective function values and different estimates for the model parameters.

Different proposals for the initial weighting matrix – with varying degrees of asymptotic efficiency – exist in the literature (see Blundell *et al.* 2001) for the various types of moment conditions which can be employed in the estimation of the linear dynamic panel data model

in Equation (3). Common examples involve identity or tridiagonal matrices. Generally, the proposed weighting matrices are based on the expected variances and covariances of the moment conditions and are derived from the underlying model assumptions. The optimal $\boldsymbol{W}$ is proportional (up to a multiplicative constant) to the inverse of the covariance matrix of the moment conditions (see, e.g., Arellano 2003).

A popular proposal for the initial weighting matrix $\widehat{\boldsymbol{W}}_1$ of the one-step GMM estimator (GMM1S) is

$$\widehat{\boldsymbol{W}}_1 = \left(\frac{1}{n} \cdot \boldsymbol{Z}'\boldsymbol{H}\boldsymbol{Z}\right)^{-1}. \tag{13}$$

The structure of the matrix $\boldsymbol{H}$ varies depending on the types of moment conditions employed in estimation. When only the HNR moment conditions are used, Arellano and Bond (1991) propose to set the matrix to

$$\boldsymbol{H}_{HNR} = \begin{pmatrix} 2 & -1 & 0 & 0 & \ldots & 0 \\ -1 & 2 & -1 & 0 & & \\ 0 & -1 & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & & & 0 \\ & & & & 2 & -1 \\ 0 & & \ldots & 0 & -1 & 2 \end{pmatrix}.$$

The tridiagonal matrix $\boldsymbol{H}_{HNR}$ accounts for the serial correlation in the idiosyncratic remainder components induced after taking first differences of Equation (3) to eliminate η_i.

When using only the ABov moment conditions in estimation, a choice for $\boldsymbol{H}$ often encountered in practice is the identity matrix with $T-2$ diagonal elements, i.e. $\boldsymbol{H}_{ABov} = I_{T-2}$. For the AS moment conditions, $\boldsymbol{H}_{AS} = I_{T-2}$ (see, e.g., Blundell *et al.* 2001; Kripfganz 2019) is the default choice. Finally, when two different sets of moment conditions are employed, a general representation of $\boldsymbol{H}$ is

$$\boldsymbol{H} = \begin{pmatrix} \boldsymbol{A} & \boldsymbol{B} \\ \boldsymbol{B}' & \boldsymbol{C} \end{pmatrix},$$

where the matrices $\boldsymbol{A}$, $\boldsymbol{B}$, and $\boldsymbol{C}$ are chosen depending on the particular moment conditions employed in GMM estimation.

When only linear moment conditions are used, we can obtain a closed form for GMM1S $\hat{\boldsymbol{\theta}}_1$ as:

$$\hat{\boldsymbol{\theta}}_1 = (\boldsymbol{X}'\boldsymbol{Z}\widehat{\boldsymbol{W}}_1\boldsymbol{Z}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{Z}\widehat{\boldsymbol{W}}_1\boldsymbol{Z}'\boldsymbol{y}. \tag{14}$$

The matrix $\boldsymbol{X}$ contains all the first differences of the right-hand side variables, along with the corresponding levels depending on the set of moment conditions used. Similarly, $\boldsymbol{y}$ contains all the first differences of the left-hand side variables, along with the corresponding levels.

An estimate for the weighting matrix $\widehat{\boldsymbol{W}}_2$ of the two-step GMM estimator (GMM2S) is

$$\widehat{\boldsymbol{W}}_2 = \left(\frac{1}{n} \cdot \boldsymbol{Z}'\hat{\boldsymbol{s}}_1\hat{\boldsymbol{s}}_1'\boldsymbol{Z}\right)^{-1}, \tag{15}$$

where $\hat{\boldsymbol{s}}_1$ denotes the residuals from one-step estimation. In order to calculate the two-step coefficient estimates $\hat{\boldsymbol{\theta}}_2$, $\widehat{\boldsymbol{W}}_1$ needs to be replaced by the estimated two-step weighting matrix $\widehat{\boldsymbol{W}}_2$:

$$\hat{\boldsymbol{\theta}}_2 = (\boldsymbol{X}'\boldsymbol{Z}\widehat{\boldsymbol{W}}_2\boldsymbol{Z}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{Z}\widehat{\boldsymbol{W}}_2\boldsymbol{Z}'\boldsymbol{y}. \tag{16}$$

We can continue by obtaining residuals from two-step estimation, computing the third-step weighting matrix using those residuals, and obtaining the third-step GMM estimator. If we update the residuals until either the change in coefficient estimates from one estimation step to the next does not exceed a certain pre-specified threshold z_{tol} or after a pre-specified number of maximum iterations h_{iter}, then the result is an iterated GMM estimator.

5. Standard errors and inference

5.1. Standard errors

Asymptotic one-step standard errors for the estimated coefficients can be obtained by taking the square root of the main diagonal elements of the estimated one-step variance covariance matrix

$$\widehat{\boldsymbol{\Omega}}(\hat{\boldsymbol{\theta}}_1) = n \cdot (\boldsymbol{X}'\boldsymbol{Z}\widehat{\boldsymbol{W}}_1\boldsymbol{Z}'\boldsymbol{X})^{-1}\hat{\sigma}_1^2, \qquad \text{with} \qquad \hat{\sigma}_1^2 = \hat{\boldsymbol{s}}_1'\hat{\boldsymbol{s}}_1 \cdot \frac{1}{N-p}. \tag{17}$$

In the formula, N is the number of observations available for estimation (i.e., n times T minus the number of missing observations), p denotes the number of estimated coefficients, and $\hat{\boldsymbol{s}}_1$ are residuals from GMM1S (see Doornik *et al.* 2012). As stated in Windmeijer (2005), robust one-step standard errors are available from

$$\widehat{\boldsymbol{\Omega}}_r(\hat{\boldsymbol{\theta}}_1) = n \cdot (\boldsymbol{X}'\boldsymbol{Z}\widehat{\boldsymbol{W}}_1\boldsymbol{Z}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{Z}\widehat{\boldsymbol{W}}_1\widehat{\boldsymbol{W}}_2^{-1}\widehat{\boldsymbol{W}}_1\boldsymbol{Z}'\boldsymbol{X}(\boldsymbol{X}'\boldsymbol{Z}\widehat{\boldsymbol{W}}_1\boldsymbol{Z}'\boldsymbol{X})^{-1}, \tag{18}$$

while asymptotic two-step standard errors can be computed from

$$\widehat{\boldsymbol{\Omega}}(\hat{\boldsymbol{\theta}}_2) = n \cdot (\boldsymbol{X}'\boldsymbol{Z}\widehat{\boldsymbol{W}}_2\boldsymbol{Z}'\boldsymbol{X})^{-1}. \tag{19}$$

Since asymptotic two-step GMM standard errors for the estimated coefficients exhibit a downward bias in small samples, they can, however, be substantially lower than one-step GMM standard errors (see, e.g., Arellano and Bond 1991). Windmeijer (2005) relates the bias to the dependence of the two-step weighting matrix on parameter estimates (the one-step estimates) and proposes an analytic correction of the two-step standard errors based on a first order Taylor-series expansion:

$$\begin{aligned}\widehat{\boldsymbol{\Omega}}_c(\hat{\boldsymbol{\theta}}_2) = &\widehat{\boldsymbol{\Omega}}(\hat{\boldsymbol{\theta}}_2) + \boldsymbol{D}_{\hat{\boldsymbol{\theta}}_2,\widehat{\boldsymbol{W}}_2}\widehat{\boldsymbol{\Omega}}(\hat{\boldsymbol{\theta}}_2) + \widehat{\boldsymbol{\Omega}}(\hat{\boldsymbol{\theta}}_2)\boldsymbol{D}'_{\hat{\boldsymbol{\theta}}_2,\widehat{\boldsymbol{W}}_2} \\ &+ \boldsymbol{D}_{\hat{\boldsymbol{\theta}}_2,\widehat{\boldsymbol{W}}_2}\widehat{\boldsymbol{\Omega}}_r(\hat{\boldsymbol{\theta}}_1)\boldsymbol{D}'_{\hat{\boldsymbol{\theta}}_2,\widehat{\boldsymbol{W}}_2},\end{aligned} \tag{20}$$

Notice that the first term is the estimated uncorrected two-step variance covariance matrix of the coefficient estimates. The computation of the correction $\boldsymbol{D}_{\hat{\boldsymbol{\theta}}_2,\widehat{\boldsymbol{W}}_2}$ is involved when multiple parameters are estimated. For a single parameter, it equals

$$\boldsymbol{D}_{\hat{\boldsymbol{\theta}}_2,\widehat{\boldsymbol{W}}_2} = -\frac{1}{n} \cdot \widehat{\boldsymbol{\Omega}}(\hat{\boldsymbol{\theta}}_2)\boldsymbol{X}'\boldsymbol{Z}\widehat{\boldsymbol{W}}_2 \left.\frac{\partial \widehat{\boldsymbol{W}}^{-1}(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}}\right|_{\boldsymbol{\theta}=\hat{\boldsymbol{\theta}}_1} \widehat{\boldsymbol{W}}_2\boldsymbol{Z}'\hat{\boldsymbol{s}}_2.$$

The vector $\hat{\boldsymbol{s}}_2$ denotes the two-step residuals and the first derivative of the weighting matrix for two-step GMM estimation evaluated at $\hat{\boldsymbol{\theta}}_1$ can be calculated from

$$\left.\frac{\partial \widehat{\boldsymbol{W}}^{-1}(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}}\right|_{\boldsymbol{\theta}=\hat{\boldsymbol{\theta}}_1} = -\frac{1}{n} \cdot \boldsymbol{Z}'(\boldsymbol{X}\hat{\boldsymbol{s}}_1' + \hat{\boldsymbol{s}}_1\boldsymbol{X}')\boldsymbol{Z}.$$

5.2. Specification testing

Arellano and Bond (1991) suggest a test for second order serial correlation in the idiosyncratic remainder components. The test is generalized to higher orders j by Arellano (2003) and can be used as a specification test in the estimation of linear dynamic panel data models. The reasoning is that, although, first order serial correlation is present in the idiosyncratic remainder components for GMM estimation based on first-differenced equations, no higher-order autocorrelation beyond the first-order should prevail. The serial correlation test of Arellano and Bond (1991) boils down to checking if the deviation of the covariance of the residuals of period t with the residuals of period $t-j$ from zero is large enough to indicate that j-th order serial correlation might be present in the idiosyncratic remainder components. The null hypothesis of the test is that there is no serial correlation in the $\varepsilon_{i,t}$. The corresponding test statistics are defined as

$$T_{m_j} = \frac{\hat{r}_j}{\hat{\sigma}_{\hat{r}_j}}, \qquad \text{with} \qquad T_{m_j} \overset{a}{\sim} \mathcal{N}(0,1),$$

where $\hat{\sigma}_{\hat{r}_j}$ is the standard error of the j-th order autocovariance of the residuals $\hat{r}_j$. For the linear dynamic panel data model specified in Equation (3), this autocovariance of the residuals is the sample equivalent to

$$r_j = \frac{1}{T-3-j} \cdot \sum_{t=4+j}^{T} r_{t,j}, \qquad \text{with} \qquad r_{t,j} = E(\Delta s_{i,t}\Delta s_{i,t-j}),$$

the average j-th order autocovariance of the idiosyncratic remainder components (see Arellano 2003). As detailed by Arellano and Bond (1991) and Doornik *et al.* (2012), the corresponding scaled autocovariance of the residuals can be calculated by

$$\hat{r}_{t,j} = \frac{1}{\sqrt{n}} \cdot \hat{\boldsymbol{s}}_t' \, \hat{\boldsymbol{s}}_{t-j},$$

where $\hat{\boldsymbol{s}}_t$ and $\hat{\boldsymbol{s}}_{t-j}$ are column vectors which may contain the residuals from one-step, two-step, or iterated GMM estimation for all cross sectional units at the respective time period; the index at $\hat{\boldsymbol{s}}_{t-j}$ indicates that the corresponding residuals are lagged j time periods. According to Arellano and Bond (1991), the estimated variance of the j-th order autocovariance of the residuals is available from

$$\hat{\sigma}_{\hat{r}_j}^2 = \frac{1}{n} \cdot \hat{\boldsymbol{s}}_{t-j}' \, \widehat{\boldsymbol{\Omega}}(\hat{s})\hat{\boldsymbol{s}}_{t-j} - 2 \cdot \hat{\boldsymbol{s}}_{t-j}' \, \boldsymbol{X}(\boldsymbol{X}'\boldsymbol{Z}\widehat{\boldsymbol{W}}\boldsymbol{Z}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{Z}\widehat{\boldsymbol{W}}\,\boldsymbol{Z}'\widehat{\boldsymbol{\Omega}}(\hat{s})\hat{\boldsymbol{s}}_{t-j} +$$
$$\hat{\boldsymbol{s}}_{t-j}' \, \boldsymbol{X}\widehat{\boldsymbol{\Omega}}(\hat{\boldsymbol{\theta}})\boldsymbol{X}'\hat{\boldsymbol{s}}_{t-j}.$$

Note that the vectors of residuals $\hat{\boldsymbol{s}}_t$, $\hat{\boldsymbol{s}}_{t-j}$ and the matrices $\widehat{\boldsymbol{W}}$, $\widehat{\boldsymbol{\Omega}}(\hat{s})$, and $\widehat{\boldsymbol{\Omega}}(\hat{\boldsymbol{\theta}})$ depend on the actual estimation step and the latter two matrices also depend on the estimated type of

variance covariance matrix (i.e., robust or asymptotic for one-step estimation; Windmeijer-corrected or asymptotic for two-step estimation).

5.3. Overidentifying restrictions testing

When the system of equations from which the model parameters are estimated by GMM is overidentified (i.e., when the number of moment conditions exceeds the number of parameters to be estimated), it is possible to assess the validity of the overidentifying restrictions by the Sargan test (Sargan 1958). The presumed null hypothesis is that the overidentifying restrictions are valid. According to Arellano and Bond (1991) and Doornik *et al.* (2012), the test statistic of the Sargan test can be computed from

$$T_S = n \cdot \hat{\boldsymbol{s}}_1' \boldsymbol{Z} \widehat{\boldsymbol{W}}_1 \boldsymbol{Z}' \hat{\boldsymbol{s}}_1 \cdot \hat{\sigma}_1^{-2}, \quad \text{with} \quad T_S \overset{a}{\sim} \chi^2(m-p).$$

Under suitable conditions, which ensure asymptotic normality of the GMM estimator and the additional assumption of conditional homoscedasticity, the test statistic is asymptotically χ^2-distributed with $m - p$ degrees of freedom; m equals the number of moment conditions employed in estimation (see, e.g., Hayashi 2000).

An alternative test statistic, where a finite fourth moments assumption is imposed instead of conditional homoscedasticity, is the J-test (Hansen 1982). The J-test statistic results from replacing $\widehat{\boldsymbol{W}}_1$ in the above formula by $\widehat{\boldsymbol{W}}_2$, the one-step residuals by the two-step residuals,

$$T_J = n \cdot \hat{\boldsymbol{s}}_2' \boldsymbol{Z} \widehat{\boldsymbol{W}}_2 \boldsymbol{Z}' \hat{\boldsymbol{s}}_2, \quad \text{with} \quad T_J \overset{a}{\sim} \chi^2(m-p).$$

The idea underlying the test statistics T_S and T_J is, that when the moment conditions are valid, the sample analogues of these conditions should be close to zero. A large value of the test statistic indicates that some of the moment conditions may be invalid, that some of the model assumptions may be incorrect, or both (see, e.g., Hayashi 2000).

These two tests also allow us to check the validity of nested subsets of moment conditions. These tests are referred to as 'difference-in-Hansen'/'difference-in-Sargan' tests (see, e.g., Roodman 2009), 'incremental Hansen'/'incremental Sargan' tests (see, e.g., Arellano 2003), or C-statistics (see, e.g., Hayashi 2000). The test statistic is obtained by first estimating the null model with a restricted set of moment conditions and the full model with more moment conditions than in the null model. Then, one computes either T_S or T_J for both models, and then take the difference of these two test statistics. This difference is asymptotically χ^2-distributed with degrees of freedom equal to the difference in the number of moment conditions under the full model and the null model (see Hayashi 2000).

5.4. Testing linear hypotheses

The Wald test is one possibility to test general linear hypotheses of the form $H_0 : \boldsymbol{R\theta} = \boldsymbol{r}$, where the matrix $\boldsymbol{R}$ is a $c \times p$ matrix, which selects the elements of the $p \times 1$ vector of population parameters $\boldsymbol{\theta}$ required to express the left-hand side of the c equations of the null hypothesis (i.e., the restrictions under the null) and the vector $\boldsymbol{r}$ is a $c \times 1$ vector that states the right-hand side of the equations. The Wald statistic can be obtained from

$$T_W = n \cdot (\boldsymbol{R\hat{\theta}} - \boldsymbol{r})' \left(\boldsymbol{R} \, \widehat{\boldsymbol{\Omega}}(\hat{\boldsymbol{\theta}}) \, \boldsymbol{R}' \right)^{-1} (\boldsymbol{R\hat{\theta}} - \boldsymbol{r}), \quad \text{with} \quad T_W \overset{a}{\sim} \chi^2(c).$$

Note that the covariance matrix $\widehat{\boldsymbol{\Omega}}(\hat{\boldsymbol{\theta}})$ and the estimator $\boldsymbol{\theta}$ depend on the actual estimation step and the former depends on the type of covariance matrix (i.e., robust or asymptotic for one-step estimation; Windmeijer-corrected or asymptotic for two-step estimation). As usual, a large value of the Wald statistic casts doubt on the null hypothesis.

Tests of three different standard null hypotheses are currently available in **pdynmc**: (a) all coefficients corresponding to the lagged-dependent and covariates are zero jointly, (b) all coefficients corresponding to the time dummies are zero jointly, or both (a) and (b).

6. Sample session

The functionality of **pdynmc** is illustrated by replicating some of the empirical results in Arellano and Bond (1991). Additionally, we show how to incorporate the linear AB and the nonlinear AS moment conditions into the analysis. We explain all arguments which need to be set to reproduce the results and point out some alternative options. We also draw comparisons between **pdynmc**, the Stata implementations **xtabond2** (Roodman 2018), **xtdpdgmm** (Kripfganz 2019), and the **pgmm** (Croissant, Millo, and Tappe 2019) function in the R-package **plm** – where we are aware of differences between the implementations.

The data set employed in Arellano and Bond (1991) is an unbalanced panel of $n = 140$ firms located in the UK observed over $T = 9$ time periods and is available from R package **plm**:

```
data(EmplUK, package = "plm")
dat                    <- EmplUK
dat[,c(4:7)]           <- log(dat[,c(4:7)])
names(dat[,c(4:7)])    <- c("n", "w", "k", "ys")
```

The authors investigate employment equations and consider the dynamic specification

$$
\begin{aligned}
n_{i,t} =& \alpha_1 n_{i,t-1} + \alpha_2 n_{i,t-2} + \\
& \beta_1 w_{i,t} + \beta_2 w_{i,t-1} + \beta_3 k_{i,t} + \beta_4 k_{i,t-1} + \beta_5 k_{i,t-2} + \beta_6 ys_{i,t} + \beta_7 ys_{i,t-1} + \beta_8 ys_{i,t-2} + \\
& \gamma_3 d_3 + \cdots + \gamma_T d_T + \eta_i + \varepsilon_{i,t}, \qquad i = 1, ..., n; \ t = 3, ..., T.
\end{aligned}
\tag{21}
$$

In the equation, i denotes the firm and t is the time series dimension. The natural logarithm of employment n is explained by its first two lags and the further explanatory variables natural logarithm of wage w, natural logarithm of capital k, natural logarithm of output ys, and their lags of order up to one (for w) or two (for k and ys). The variables $d_3, \ldots, d_T$ are time dummies with corresponding coefficients $\gamma_3, \ldots, \gamma_T$; the unobserved individual-specific effect is represented by η, and ε is an idiosyncratic remainder component. The goal is to estimate the lag parameters α_1 and α_2 and the coefficients of the further explanatory variables β_j, with $j = 1, \ldots, 8$, while controlling for (unobserved) time effects and accounting for unobserved individual-specific heterogeneity.

6.1. GMM estimation with HNR moment conditions

When reproducing the results in Table 4 on p.290 of Arellano and Bond (1991) with **pdynmc**, the model structure underlying Equation (21) can be specified by:

```
m1 <- pdynmc(
```

```
  dat = dat, varname.i = "firm", varname.t = "year",
  use.mc.diff = TRUE, use.mc.lev = FALSE, use.mc.nonlin = FALSE,
  include.y = TRUE, varname.y = "emp", lagTerms.y = 2,
  fur.con = TRUE, fur.con.diff = TRUE, fur.con.lev = FALSE,
  varname.reg.fur = c("wage", "capital", "output"),
  lagTerms.reg.fur = c(1,2,2),
  include.dum = TRUE, dum.diff = TRUE, dum.lev = FALSE, varname.dum = "year",
  w.mat = "iid.err", std.err = "corrected",
  estimation = "onestep", opt.meth = "none")
```

The first arguments relate to the data set (`dat`), the cross-section identifier (`varname.i`), and time series dimension (`varname.t`). Next, the moment conditions are defined – here, only moment conditions from equations in differences are used. The moment conditions are derived for the dependent variable (`varname.y`), and for the corresponding number of lags of the dependent variable to be included as explanatory variables (`lagTerms.y`). Further explanatory variables aside from the lags of the dependent variable (`fur.con`) are included in the equations in differences, not in the equations in levels (compare `fur.con.diff = TRUE`, `fur.con.lev = FALSE`), their variable names are stated in (`varname.reg.fur`), their lag structure is specified in (`lagTerms.reg.fur`). Note that the first element of the vector denoting the lag structure corresponds to the first element of the vector with the variable names, the second element to the second, and so on. Also note that all names given in the vectors that refer to variables in the data set need to have the same names as in the data set. The time dummies are included by (`include.dum`) into the equations in differences not in the equations in levels (`dum.diff = TRUE`, `dum.lev = FALSE`), the dummy indicator variable is (`varname.dum`). Note that time dummies can be constructed from one or multiple variables by **pdynmc** by simply passing a scalar or vector with the respective variable names in the data to `varname.dum`. Specifying the matrix H in Equation (13), which governs the structure of the one-step weighting matrix, and carrying out one-step estimation can be achieved by the commands in the last two codelines. Choosing the option `iid.err` uses the matrix H_{HNR} proposed by Arellano and Bond (1991). Alternatively, an identity matrix can be employed for H by the option `identity`. Code `std.err = "corrected"` yields robust standard errors for one-step estimation (in case of a two-step estimation, the correction of Windmeijer 2005 was employed by this argument).

One-step, two-step, and iterated GMM are carried out by numerical optimization of the GMM objective function given in Equation (12). Since a closed form solution exists for the estimator when employing only linear moment conditions, numerical optimization is not required and can be switched off by setting `opt.meth = "none"`.

The standard output can be accessed via `summary(m1)` and could be found in panel (a) of Table 1. These reproduce the coefficient estimates in Table 4, column (a1) on p.290 of Arellano and Bond (1991), when one specifies all arguments as stated in this section.

Changing the argument `estimation` to `twostep` yields the two-step GMM coefficient estimates (the `pdynmc`-output object is assigned to `m2`) from Table 4, column (a2) on p.290 of Arellano and Bond (1991). These results may be found in panel (b) of Table 1.

Note that the standard errors presented in column (b) of Table 1 are based on the Windmeijer-correction and deviate from the conventional standard errors reported in Arellano and Bond (1991). The standard errors from the original analysis can be reproduced by setting `std.err`

```
= "unadjusted".
```

Table 1: Estimates in the spirit of Table 4 in Arellano and Bond (1991)

	(a) Estimate	(a) Std.Err.rob	(b) Estimate	(b) Std.Err.rob	(c) Estimate	(c) Std.Err.rob	(d) Estimate	(d) Std.Err.rob	(e) Estimate	(e) Std.Err.rob
L1.n	0.68623***	0.14459	0.62871**	0.19341	1.10335***	0.05002	1.08852***	0.07476	1.16847***	0.06239
L2.n	-0.08536	0.05602	-0.06519	0.04505	-0.10423*	0.04652	-0.05953	0.07776	-0.19362***	0.05663
w	-0.60782***	0.17821	-0.52576***	0.15461	-0.44828**	0.14872	-0.53856***	0.11410	-0.21777	0.12304
L1.w	0.39262*	0.16799	0.31129	0.20300	0.42262**	0.15564	0.52214***	0.11837	0.17753	0.12743
k	0.35685***	0.05902	0.27836***	0.07280	0.29027***	0.04999	0.34252***	0.05727	0.22181***	0.04912
L1.k	-0.05800	0.07318	0.01410	0.09246	-0.15286*	0.06704	-0.21398**	0.06679	-0.07124	0.06904
L2.k	-0.01995	0.03271	-0.04025	0.04327	-0.13719***	0.04127	-0.15273***	0.03866	-0.12912**	0.04568
ys	0.60851***	0.17253	0.59192***	0.17309	0.54843**	0.19379	0.59404**	0.18922	0.29825	0.17402
L1.ys	-0.71116**	0.23172	-0.56599*	0.26110	-0.66598**	0.22142	-0.77509***	0.22750	-0.39531	0.21439
L2.ys	0.10580	0.14120	0.10054	0.16110	0.12729	0.15561	0.17620	0.15027	0.12557	0.13491
1979	0.00955	0.01029	0.01122	0.01168	0.02354*	0.01097	0.02278***	0.00621	0.01567	0.00969
1980	0.02202	0.01771	0.02307	0.02006	0.04106*	0.01953	0.04178*	0.01800	0.02034	0.01929
1981	-0.01177	0.02951	-0.02136	0.03324	0.00215	0.03389	0.01271	0.02742	-0.04586	0.03014
1982	-0.02706	0.02928	-0.03112	0.03397	0.01762	0.02289	0.01554	0.02401	0.00703	0.02050
1983	-0.02132	0.03046	-0.01799	0.03693	0.04332*	0.01803	0.03006	0.01730	0.04808*	0.01886
1984	-0.00770	0.03141	-0.02337	0.03661	0.02912	0.02161	0.02610	0.03956	0.02236	0.02069

(a) one-step estimates; instruments for equations in first differences: $L(2/8).n, D.w, L.D.w, D.k, L.D.k, L2.D.k, D.ys, L.D.ys, L2.D.ys, D.1979 - D.1984$

(b) two-step estimates; instruments for equations in first differences: $L(2/8).n, D.w, L.D.w, D.k, L.D.k, L2.D.k, D.ys, L.D.ys, L2.D.ys, D.1979 - D.1984$

(c) two-step estimates; equations in first differences: $L(2/8).n, D.w, L.D.w, L2.D.w, D.k, L.D.k, L2.D.k, D.ys, L.D.ys, L2.D.ys, D.1979 - D.1984$
    equations in levels: $L(1/7).D.n, w, L.w, L2.w, k, L.k, L2.k, ys, L.ys, L2.ys$

(d) two-step estimates; equations in first differences: $L(2/8).n, u, D.w, L.D.w, L2.D.w, D.k, L.D.k, L2.D.k, D.ys, L.D.ys, L2.D.ys, D.1979 - D.1984$
    equations in levels: $w, L.w, L2.w, k, L.k, L2.k, ys, L.ys, L2.ys$

(e) 17 step estimates; equations in first differences: $L(2/8).n, D.w, L.D.w, L2.D.w, D.k, L.D.k, L2.D.k, D.ys, L.D.ys, L2.D.ys, D.1979 - D.1984$
    equations in levels: $L(1/7).D.n, w, L.w, L2.w, k, L.k, L2.k, ys, L.ys, L2.ys$

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$ (refers to t-test of the null that the coefficient is equal to zero)

Regarding the arguments

```
use.mc.diff = TRUE, include.y = TRUE, include.x = FALSE
```

it has to be noted that the HNR moment conditions (`use.mc.diff`) derived from the lagged dependent variable (`include.y`) are employed, while none are derived from other explanatory variables (by default `include.x = FALSE`). The latter argument implies that these other explanatory variables in the model are assumed to be exogenous and instrument themselves.

Different capabilities for testing hypotheses about the population parameters are available in **pdynmc**. Among them are the tests for serial correlation in the idiosyncratic remainder components proposed by Arellano and Bond (1991), Hansen tests, and Wald tests. In the following, carrying out these tests and interpreting the results is briefly illustrated based on the two-step GMM estimation results presented in column (b) of Table 1.

Employing the test for second order serial correlation of Arellano and Bond (1991) described in Section 5.2 by `m.test(m2, t.order = 2)` yields:

```
        Serial correlation test of degree 2

data:  GMM Estimation; H0: no serial correlation of order 2 in epsilon
normal = -0.36744, p-value = 0.7133
```

The test does not reject the null hypothesis at any plausible significance level and does not provide any indication that the model specification might be inadequate. The test statistic and p-value are similar to **xtabond2** and **pgmm**.

Computing the Hansen J-test of the overidentifying restrictions described in Section 5.3 by `j.test(m2)` yields:

```
        J-Test of Hansen

data:  GMM Estimation; H0: overidentifying restrictions valid
chisq = 31.381, df = 25, p-value = 0.1767
```

The test does not reject the overidentifying restrictions and does not provide any indications that the validity of the instruments employed in estimation may be in doubt. Comparing the results to **xtabond2** shows that the degrees of freedom and the p-value differ. We consider 25 degrees of freedom to be the appropriate number here, as 41 instruments are employed in estimation to obtain 16 coefficient estimates. It seems that the function **xtabond2** does not correct the degrees of freedom for the number of dummies dropped in estimation[1]. The difference in the p-value is due to the differences in the degrees of freedom. Our results are equivalent to the results of **pgmm** for the overidentifying restrictions test. In **pgmm**, the above test is referred to as 'Sargan test'.

For the Wald test illustrated in Section 5.4, consider the null hypothesis that the population parameters of all coefficients included in the model are zero jointly, which is tested by `wald.fct(param = "all", object = m2`:

---

[1]Dummies are dropped by the estimation routine in case of high collinearity.

```
        Wald test

data:  GMM Estimation; H0: beta = 0; tested model parameters: all
chisq = 1104.7, df = 16, p-value < 2.2e-16
```

The test rejects the null hypothesis. Comparing the test result to the implementation of the test in **xtabond2** – again – reveals differences concerning the degrees of freedom. We consider 16 to be the appropriate number of degrees of freedom here, since this corresponds to the number of estimated parameters. As noted previously, the differences seem to stem from **xtabond2** not adjusting the degrees of freedom for the dummies dropped in estimation. Alternative hypotheses that can be tested via the Wald test in **pdynmc** are that all slope parameters are zero jointly and that all parameters corresponding to the time dummies are zero jointly (`param = "time.dum"` only tests the time dummies, while `param = "slope"` only tests the slope parameters).

6.2. GMM estimation with HNR and ABov moment conditions

When the 'constant correlated effects' assumption stated in Equation (10) holds, the HNR moment conditions from equations in differences employed in Section 6.1 can be extended by the ABov moment conditions from equations in levels.

The ABov moment conditions are particularly useful for data generating processes, which are highly persistent (Blundell and Bond 1998). In this case, identification by the HNR moment conditions from equations in levels may fail and GMM estimation based on HNR moment conditions is documented to possess poor finite sample performance (see, e.g., Blundell and Bond 1998; Blundell *et al.* 2001; Bun and Sarafidis 2015).

In **pdynmc**, the ABov moment conditions from equations in levels can be (additionally) incorporated by :

```
use.mc.lev = TRUE
```

In principle, both, the time dummies and the further explanatory variables can be included in the equations in first differences and the level equations. It is recommended, though, to include the dummies only in one of the equations, as it can be shown that incorporating them in both equations renders one set of dummies redundant for estimation – while for the non-lagged dependent explanatory variables, this equivalence does not hold.[2] We accommodate non-lagged dependent explanatory variables in the levels equations by

```
fur.con.lev = TRUE
```

and use this argument together with the earlier specified ones.

In order to obtain coefficient estimates, a decision about the matrix H in the one-step weighting matrix is required. When using the HNR and ABov moment conditions, the decision about H effectively involves specifying the matrices A, B, and C in the general structure given

---

[2]Note that this is the case in balanced panels. The results may also not be numerically identical across function calls for different choices of the one-step weighting matrix. For a discussion, see https://www.statalist.org/forums/forum/general-stata-discussion/general/1357268-system-gmm-time-dummies.

in Section 4.2. As mentioned, the diagonal elements A and C reflect the expected variance covariance properties within a set of moment conditions, while B reflects the expected covariances across different sets of moment conditions. In the given setting, A corresponds to the variance covariance properties of the HNR moment conditions, C to those of the ABov moment conditions, and B to those across the HNR and ABov moment conditions. Three different options are currently available in **pdynmc** to set up the weighting matrix `w.mat`: `iid.err`, `identity`, and `zero.cov`. The first option leads to H_{HNR} being used for A, an identity for C, and a matrix B, such that $BB' = H_{HNR}$. Setting `w.mat` to `identity` leads to an identity matrix being used for the diagonal matrices A and C and an adequately dimensioned matrix B with 1 on the diagonal[3]. When using the option `zero.cov`, the matrices A and C are as for option `iid.err`, but B is set to a null matrix. In case nonlinear moment conditions are used, the part of H which corresponds to the nonlinear moment conditions is set to an identity for all choices of `w.mat`. All elements of the matrices containing the expected covariance properties of the nonlinear moment conditions with other moment conditions are always set to zero.

The results presented in column (c) of Table 1 are the two-step estimates of column (a2) of Table 4 in Arellano and Bond (1991) extended by the ABov moment conditions. All arguments are specified as described above. Including the ABov moment conditions into the analysis leads to substantial changes in the coefficient estimates of the first lag of the dependent variable. Note that the results indicate a markedly higher persistence of employment and render including two lags of the dependent variable questionable (Blundell and Bond 1998, e.g., estimate a version of the equation which contains only one lag of all explanatory variables). Note that the coefficient estimates of the explanatory variables, besides the first lag of the dependent variable, appear to be similar across estimations.

Equivalent results to column (c) of Table 1 can be obtained from the **pgmm** function in the **plm**-package – besides some minor numerical differences at the fifth digit. When replicating the results with **xtabond2**, differences in the implementations become obvious: The instrument set for the ABov moment conditions is extended in similar fashion to the HNR moment conditions in **xtabond2**, while this is not the case in **pgmm**. An argument is available in **pdynmc** to extend the instrument set as in **xtabond2**:

```
inst.stata = TRUE
```

Due to the reasons described in Section 3.3, this argument is set to `FALSE` per default. When setting the option to `TRUE`, the results from **xtabond2** and **pdynmc** are very close to our results.

6.3. GMM estimation with HNR and AS moment conditions

Recall that the linear ABov moment conditions from equations in levels comprise the nonlinear AS moment conditions and render them redundant for estimation (Blundell and Bond 1998; a derivation is provided in Fritsch 2019). Both sets of moment conditions may be useful in GMM estimation when the lag parameter is close to unity and it can be shown that extending the HNR moment conditions by either the ABov- or the AS moment conditions may identify the lag parameter – even when the individual moment conditions fail to do so (Bun and Kleibergen

---

[3]Note that the matrix B is not necessarily a quadratic matrix.

2019; Gørgens *et al.* 2019). The ABov moment conditions require the 'constant correlated effects' assumption, while the AS moment conditions only require standard assumptions to hold. Therefore, the latter may be useful in situations where the 'constant correlated effects' assumption is in doubt and the statistician aims to investigate a highly persistent dynamic process with a structure similar to Equation (3). In **pdynmc**, including nonlinear moment conditions into the analysis is available via:

```
use.mc.nonlin = TRUE
```

When extending the analysis of Arellano and Bond (1991) by the nonlinear AS moment conditions, the results differ substantially from column (b) of Table 1 and are very similar to the coefficient estimates shown in column (c) of Table 1. This casts doubt on the HNR moment conditions and may be a hint that there is high persistence in the employment process – as high persistence leads to the lag parameters not being identified by the HNR moment conditions (Bun and Kleibergen 2019; Gørgens *et al.* 2019). Note that since the unobservable error term u (which can be expressed in terms of observable model components and parameters) is included in the instrument set for the equations in first differences, nonlinear moment conditions are employed in estimation. Also note that the coefficient estimates in column (d) of Table 1 are very close to coefficient estimates obtained from **xtdpdgmm**.

6.4. Iterated GMM

Iterated GMM can be used, by specifying the following commands:

```
estimation = "iterative", max.iter = 100, iter.tol = 0.01,
```

When `estimation = "iterative"` is used, `max.iter` specifies the maximum number of iterations, `iter.tol` the search tolerance w.r.t. convergence. We employ `max.iter = 100` and `iter.tol = 0.01`. These values are the default values and can be adjusted by the user. Iterated GMM results are shown in column (e) of Table 1. The moment conditions employed are the same as in column (c) of the table. The parameter estimates obtained after 17 steps are relatively similar to those in columns (c)-(d).

6.5. Starting values

If numerical optimization techniques are used, the starting for all parameters are drawn from the uniform distribution on an interval [-1, 1] by the following commands (set as default):

```
start.val.lo = -1, start.val.up = 1, seed.input = 42
```

As usual, the `seed.input` ensures reproducibility. The starting values can be varied by the user via arguments `start.val.lo` and `start.val.up` and setting `custom.start.val` to TRUE.

7. Concluding remarks

The R-package **pdynmc** provides a function to estimate linear dynamic panel data models. The implementation allows for general lag structures of the explanatory variables, which may

encompass lags of the dependent variable and further explanatory variables. For estimation, linear and nonlinear moment conditions are derived from the model assumptions; further controls and external instruments (if available) may also be added. Estimation is carried out by numerical optimization of the GMM objective function. Corresponding closed form solutions are computed – where possible – and stored besides the results from numerical optimization. The estimation routine can handle balanced and unbalanced panel data sets and provides one-step-, two-step-, and continuously updating estimation. Accounting for (unobserved) time-specific effects is possible by including time dummies. Different choices for the weighting matrix, which guides the aggregation of moment conditions in one-step GMM estimation are available. Concerning the computation of standard errors for the coefficient estimates, the following options are currently available in **pdynmc**: non-robust one- and two-step standard errors and robust one-step- and Windmeijer-corrected two-step standard errors. Some standard hypothesis and specification tests are also available. Among them are Wald tests, overidentifying restrictions tests and a test for serial correlation in the idiosyncratic remainder components.

We plan to extend the package by the following features in the future:

- As mentioned in Section 2, speed improvements are a priority to enable the use of this package in Monte Carlo simulations. In addition, once the "bleeding-edge" features of iterated GMM become mainstream, having a package that provides some basic functionality will be of use.

- Incorporate further diagnostics and tests to assess the validity of the estimated specifications and the underlying moment conditions and assumptions (e.g., testing the 'constant correlated effects' assumption and testing for structural breaks).

- Facilitate choosing an adequate dynamic specification by lag selection techniques.

- Include moment selection capabilities based on an appropriate criterion into GMM estimation which allow to remove certain instruments/moment conditions.

- Expand the possible choices for the one-step weighting matrix by, e.g., the proposition in Kiviet (2007b) for GMM estimation based on linear HNR- and ABov moment conditions.

- Implement the IV estimator solely based on the nonlinear moment conditions proposed by Pua *et al.* (2019a,b).

8. Acknowledgements

References

Ahn SC, Schmidt P (1995). "Efficient estimation of models for dynamic panel data." *Journal of Econometrics*, **68**(1), 5–27. doi:https://doi.org/10.1016/0304-4076(94)01641-C.

Anderson T, Hsiao C (1982). "Formulation and estimation of dynamic models using panel data." *Journal of Econometrics*, **18**(1), 47–82. doi:https://doi.org/10.1016/0304-4076(82)90095-1.

Arellano M (2003). *Panel Data Econometrics*. Oxford University Press. doi:https://doi.org/10.1093/0199245282.001.0001.

Arellano M, Bond S (1991). "Some Tests of Specification for Panel Data: Monte Carlo Evidence and an Application to Employment Equations." *The Review of Economic Studies*, **58**(2), 277–297. doi:https://doi.org/10.2307/2297968.

Arellano M, Bover O (1995). "Another look at the instrumental variable estimation of error-components models." *Journal of Econometrics*, **68**(1), 29–51. doi:https://doi.org/10.1016/0304-4076(94)01642-D.

Blundell R, Bond S (1998). "Initial conditions and moment restrictions in dynamic panel data models." *Journal of Econometrics*, **87**(1), 115–143. doi:http://dx.doi.org/10.1016/S0304-4076(98)00009-8.

Blundell R, Bond S, Windmeijer F (2001). "Estimation in dynamic panel data models: Improving on the performance of the standard GMM estimator." In BH Baltagi, TB Fomby, RC Hill (eds.), *Nonstationary Panels, Panel Cointegration, and Dynamic Panels*, volume 15 of *Advances in Econometrics*, pp. 53–91. Emerald Group Publishing. doi:https://doi.org/10.1016/S0731-9053(00)15003-0.

Bun MJG, Kleibergen F (2019). "Identification robust inference for moments based analysis of linear dynamic panel data models." *Technical report*, Universiteit van Amsterdam. URL https://drive.google.com/file/d/1UTdJvv0TFH2WbnCoo4FJ7u4tynw7ObmH/view.

Bun MJG, Sarafidis V (2015). "Chapter 3 – Dynamic Panel Data Models." In BH Baltagi (ed.), *The Oxford Handbook of Panel Data*, pp. 76–110. Oxford University Press. doi:https://doi.org/10.1093/oxfordhb/9780199940042.013.0003.

Croissant Y, Millo G (2008). "Panel Data Econometrics in R: The plm Package." *Journal of Statistical Software*, **27**(2), 1–43. doi:https://doi.org/10.18637/jss.v027.i02.

Croissant Y, Millo G (2018). *Panel Data Econometrics with R*. Wiley.

Croissant Y, Millo G, Tappe K (2019). "plm: Linear Models for Panel Data." URL https://cran.r-project.org/web/packages/plm.

Doornik JA, Arellano M, Bond S (2012). *Panel Data estimation using DPD for Ox*. URL http://www.doornik.com/download.html.

Fletcher R (1970). "A new approach to variable metric algorithms." *The Computer Journal*, **13**(3), 317–322. doi:https://doi.org/10.1093/comjnl/13.3.317.

Fritsch M (2019). "On GMM-estimation of linear dynamic panel data models." *University of Passau Working Papers in Business Administration B-36-19*, University of Passau. URL `https://www.wiwi.uni-passau.de/en/research/working-papers/`.

Gørgens T, Han C, Xue S (2019). "Moment Restrictions and Identification in Linear Dynamic Panel Data Models." *Annals of Economics and Statistics*, (134), 149–176. URL `https://www.jstor.org/stable/10.15609/annaeconstat2009.134.0149`.

Hansen BE, Lee S (2019). "Inference for Iterated GMM Under Misspecification." Unpublished Working Paper, URL `https://www.ssc.wisc.edu/~bhansen/papers/IteratedGMM.pdf`.

Hansen LP (1982). "Large Sample Properties of Generalized Method of Moments Estimators." *Econometrica*, **50**(4), 1029–1054. `doi:https://doi.org/10.2307/1912775`.

Hayashi F (2000). *Econometrics*. Princeton University Press.

Holtz-Eakin D, Newey Whitney K, Rosen HS (1988). "Estimating Vector Autoregressions with Panel Data." *Econometrica*, **56**(6), 1371–1395. `doi:https://doi.org/10.2307/1913103`.

Hsiao C (2014). *Analysis of Panel Data*. Cambridge University Press.

Hwang J, Sun Y (2018). "Should we go one step further? An accurate comparison of one-step and two-step procedures in a generalized method of moments framework." *Journal of Econometrics*, **207**(2), 381 – 405. ISSN 0304-4076. `doi:https://doi.org/10.1016/j.jeconom.2018.07.006`.

Kiviet JF (2007a). "Chapter 11 – Judging Contending Estimators by Simulation: Tournaments in Dynamic Panel Data Models." In GDA Phillips, E Tzavalis (eds.), *The Refinement of Econometric Estimation and Test Procedures: Finite Sample and Asymptotic Analysis*, pp. 282–318. Cambridge University Press. `doi:https://doi.org/10.1017/CBO9780511493157`.

Kiviet JF (2007b). "On the optimal weighting matrix for the GMM system estimator in dynamic panel data models." *UvA-Econometrics Discussion Paper 2007/08*, Universiteit van Amsterdam, Dept. of Econometrics. URL `https://ase.uva.nl/content/research-programmes/uva-econometrics/discussion-papers/discussion-papers-uva-econometrics.html#anker-discussion-papers-2007`.

Kripfganz S (2019). "XTDPDGMM: Stata module to perform generalized method of moments estimation of linear dynamic panel data models." URL `http://EconPapers.repec.org/RePEc:boc:bocode:s458395`.

Long JA (2019). *panelr: Regression Models and Utilities for Repeated Measures and Panel Data*. R package version 0.7.0, URL `https://cran.r-project.org/package=panelr`.

Nash JC (1990). *Compact Numerical Methods for Computers: Linear Algebra and Function Minimisation*. CRC Press.

Nash JC (2014). "On Best Practice Optimization Methods in R." *Journal of Statistical Software*, **60**(2), 1–14. `doi:https://doi.org/10.18637/jss.v060.i02`.

Nash JC, Varadhan R (2011). "Unifying Optimization Algorithms to Aid Software System Users: optimx for R." *Journal of Statistical Software*, **43**(9), 1–14. doi:https://doi.org/10.18637/jss.v043.i09.

Phillips PC, Han C (2019). "Chapter 5 - Dynamic panel GMM using R." In HD Vinod, C Rao (eds.), *Conceptual Econometrics Using R*, volume 41 of *Handbook of Statistics*, pp. 119 – 144. Elsevier. doi:https://doi.org/10.1016/bs.host.2019.01.002.

Pua A, Fritsch M, Schnurbus J (2019a). "Large sample properties of an IV estimator based on the Ahn and Schmidt moment conditions." *University of Passau Working Papers in Business Administration B-37-19*, University of Passau. URL https://www.wiwi.uni-passau.de/en/research/working-papers/.

Pua A, Fritsch M, Schnurbus J (2019b). "Practical aspects of using quadratic moment conditions in linear AR(1) panel data models." *University of Passau Working Papers in Business Administration B-38-19*, University of Passau. URL https://www.wiwi.uni-passau.de/en/research/working-papers/.

R Core Team (2019). *R: A Language and Environment for Statistical Computing.* R Foundation for Statistical Computing, Vienna, Austria. URL http://www.R-project.org.

Roodman D (2009). "A Note on the Theme of Too Many Instruments." *Oxford Bulletin of Economics and Statistics*, **71**(1), 135–158. doi:https://doi.org/10.1111/j.1468-0084.2008.00542.x.

Roodman D (2018). "xtabond2: Stata module to extend xtabond dynamic panel data estimator." URL https://econpapers.repec.org/software/bocbocode/s435901.htm.

Sargan JD (1958). "The Estimation of Economic Relationships Using Instrumental Variables." *Econometrica*, **26**(3), 393–415. doi:https://doi.org/10.2307/1907619.

Sigmund M, Ferstl R (2019). "Panel vector autoregression in R with the package panelvar." *The Quarterly Review of Economics and Finance*, p. forthcoming. doi:https://doi.org/10.1016/j.qref.2019.01.001.

StataCorp (2015). *Stata Longitudinal-Data/Panel-Data Reference Manual Release 14.* StataCorp LP, College Station, Texas.

Windmeijer F (2005). "A finite sample correction for the variance of linear efficient two-step GMM estimators." *Journal of Econometrics*, **126**(1), 25–51. doi:https://doi.org/10.1016/j.jeconom.2004.02.005.

Affiliation:

Markus Fritsch
Chair of Statistics and Data Analytics
School of Business, Economics, and Information Systems
University of Passau
94032 Passau, Germany
E-mail: markus.fritsch@uni-passau.de
URL: https://www.wiwi.uni-passau.de/en/statistics/team/assistant-professors/

Andrew Adrian Yu Pua
MOE Key Laboratory of Econometrics, The Wang Yanan Institute for Studies in Economics,
Department of Statistics, School of Economics, Xiamen University
Fujian Key Laboratory of Statistical Science, Xiamen University
E-mail: andrewypua@outlook.com
URL: http://andrew-pua.ghost.io

Joachim Schnurbus
Computational Statistics and Mathematics Teaching Unit
School of Business, Economics, and Information Systems
University of Passau
94032 Passau, Germany
E-mail: joachim.schnurbus@uni-passau.de
URL: https://www.wiwi.uni-passau.de/csm/team/