Identifying Heterogeneity in Dynamic Panel Models with Individual Parameter Contribution Regression

Manuel Arnold
Humboldt University of Berlin, Germany
Max Planck UCL Centre for Computational Psychiatry and Ageing Research, Berlin, Germany

Daniel L. Oberski
Utrecht University, the Netherlands

Andreas M. Brandmaier
Max Planck Institute for Human Development, Berlin, Germany
Max Planck UCL Centre for Computational Psychiatry and Ageing Research, Berlin, Germany

Manuel C. Voelkle
Humboldt University of Berlin, Germany
Max Planck Institute for Human Development, Berlin, Germany

Abstract

Dynamic panel models are a popular approach to study interrelationships between repeatedly measured variables. Often, dynamic panel models are specified and estimated within a structural equation modeling (SEM) framework. An endemic problem threatening the validity of such models is unmodelled heterogeneity. Recently, individual parameter contribution (IPC) regression was proposed as a flexible method to study heterogeneity in SEM parameters as a function of observed covariates. In the present paper, we derive how IPCs can be calculated for general maximum likelihood estimates and evaluate the performance of IPC regression to estimate group differences in dynamic panel models in discrete and continuous time. We show that IPC regression can be slightly biased in samples with large group differences and present a bias correction procedure. IPC regression showed generally promising results for discrete time models. However, due to highly nonlinear parameter constraints, caution is indicated when applying IPC regression to continuous time models.

Keywords: autoregressive cross-lagged model, continuous time modeling, heterogeneity, structural equation modeling

Identifying Heterogeneity in Dynamic Panel Models with Individual Parameter Contribution Regression

Introduction

Dynamic panel models (Hsiao, 2014) are routinely used in econometrics, psychology, and sociology to model the coupling between several repeatedly measured variables. Building upon the idea of Granger causality (Granger, 1969), dynamic models allow answering questions concerned with the direction and strength of reciprocal relationships. Especially in psychological research, it is common practice to specify and estimate dynamic panel models within the structural equation modeling framework (e.g., Bollen & Brand, 2010; Allison, Williams, & Moral-Benito, 2017). An endemic problem that complicates the analysis of longitudinal panel data is potential heterogeneity in processes of interest between individuals. For instance, the subjects may show stable or trait-like differences in the levels of the observed variables; a random shock might have a long-lasting effect on one person, while its effect vanishes quickly for another individual; or the coupling of the variables may differ between subjects. By overlooking such differences between individuals or groups, researchers may risk obtaining meaningless parameter estimates and incorrect standard errors (Halaby, 2004).

Various approaches are used to identify and account for heterogeneity in dynamic panel models. A popular way is the use of multilevel models with random effects (e.g., Singer & Willet, 2003). Multilevel models are often applied when heterogeneity is considered to be unobserved, that is in situations without a priori knowledge about the source of potential individual or group differences in the population. For instance, dynamic panel models are often specified with random intercepts to account for trait-like differences in the levels of the observed variables (e.g., Hamaker, Kuiper, & Grasman, 2015). By regressing the random effects on additionally observed covariates, multilevel models can also be used to explore correlates and predictors of heterogeneity. Another popular approach to investigate heterogeneity are multi-group structural equation models (MGSEM, Sörbom, 1974) which allow the specification of panel models with different parameter values across groups. MGSEMs are particularly useful for looking at the effect of one grouping variable at a time. However, using MGSEMs to disentangle the effects of several grouping variables can become tedious as multiple MGSEMs needs to be specified and estimated. Moreover, MGSEM requires continuous covariates to be split into meaningful discrete grouping variables which often prove to be a difficult and time-consuming task. Fortunately, there exist approaches to perform such testing automatically, which become feasible with large sample sizes: Brandmaier, von Oertzen, McArdle, and Lindenberger (2013) and Brandmaier, Prindle, McArdle, and Lindenberger (2016) proposed a combination of MGSEMs and recursive partitioning methods to recover groups with similar parameter values of a hypothesized SEM. These so-called SEM trees or SEM forests fit a large number of MGSEMs to identify which grouping variables are important and should be considered to improve the fit of a model while categorizing continuous covariates in the process. Recently, Brandmaier, Driver, and Voelkle (2018) demonstrated the use of these methods for dynamic panel models. Although SEM trees or forest solve many shortcomings of MGSEMs, their computational burden that increases with sample size and constitute a major impediment to implementing this approach in practice.

While the above methods are all well-understood and are able to detect and qualify heterogeneity in a wide range of situations, they also share a certain drawback: each method requires the user to formulate a complete joint model that not only include the original observed variables but also the putative predictors of heterogeneity. Much of the specification of this joint model merely forms a set of so to speak nuisance parameters for investigating differences in a target parameter of interests. For example, in random-effects modeling with covariates, aside from the panel model itself, not only the effect of the covariate of interest on the random effect of interest needs to be specified, but also its relation to all other random effects, and their mutual covariance structure. SEM trees and MGSEM compare the group-wise likelihoods, which necessarily consider differences in all parameters across levels of the covariates jointly, apart from the SEM specification and distributional assumptions. Such joint approaches, apart from the aforementioned increased computational complexity, may also experience difficulties when there is a clear set of target parameters of interest. If the main parameter differences across levels of a covariate are to be found in parameters other than the parameter of interest, the differences of interest may be obscured – an effect that is well-known in the regression mixture literature (George et al., 2013). This may occur particularly in the case of distributional misspecification (e.g., Usami, Hayes, & McArdle, 2017). In other words, joint methods, while asymptotically efficient under the model, may be less robust, and are, by definition, not targeted to differences in the parameter of interest.

As an alternative approach to identify and estimate heterogeneity in dynamic panel models, we propose the use of individual parameter contribution (IPC) regression (Oberski, 2013). IPC regression is a simple exploratory tool to test if the validity of a theorized model is threatened by substantial heterogeneity and can be used to guide researchers in re-formulating their original models. The method proceeds in three steps. First, a theory-driven (confirmatory) SEM is specified and estimated. Second, parameter values are calculated for every individual in the sample. Third, these individual parameter contributions (IPCs) are regressed on a set of additional covariates that were originally not hypothesized to be part of the model. Thus, IPC regression allows modeling SEM parameters as a function of additional covariates. For instance, a researcher could regress the IPCs to one parameter estimate on individuals' age in the sample to test whether this parameter is invariant to age differences and to estimate how the parameter might change as a function of age. The primary advantages of IPC regression over other approaches to heterogeneity outlined above are its simplicity and low computational demand. IPC regression separates the estimation of the theory-driven model from the investigation of individual or group differences. This separation is especially useful if the theory-driven model is complex, that is with many observed variables and parameters. IPC regression allows testing every type of SEM parameter for individual or group differences without the need of specifying random effects. Moreover, the method allows studying the effect of multiple discrete as well as continuous covariates and their interactions simultaneously. Furthermore, on the side of the applied researcher, only knowledge of linear regression analysis is required to successfully apply IPCs in practice.

IPCs are not limited to SEMs and can be derived for every type of maximum likelihood estimate. The contributions are calculated by using a Taylor series expansion of the casewise derivative of the log-likelihood function around the maximum likelihood estimates. The casewise derivative of the log-likelihood function, also known as score function, has long been used to investigate the plausibility of statistical models (e.g., Zeileis, 2005; Zeileis & Hornik, 2007). Recently, score-based tests became popular in the exploration of measurement invariance in SEM (Merkle, Fan, & Zeileis, 2014; Merkle & Zeileis, 2013; Wang, Merkle, & Zeileis, 2014; Wang, Strobl, Zeileis, & Merkle, 2018). These score-based tests are used to test measurement invariance with respect to a continuous or ordinal auxiliary variable. IPC regression is different to these tests by providing estimates of how a model parameter varies as a function of covariates. Other frequently applied score-based approaches to identify misspecification in SEMs are the modification index (Sörbom, 1989) and the expected parameter change (Saris, Satorra, & Sörbom, 1987), which both test the validity of certain parameter restrictions but do not address the problem of parameter heterogeneity even though they are closely related (Oberski, 2013).

As of now, IPC regression has only been evaluated for confirmatory factor analysis models (CFA; Brown, 2006) on cross-sectional data. In a Monte Carlo simulation, Oberski (2013) reported excellent finite sample performance. IPC regression provided nearly unbiased estimates of group differences in a factor loading and a latent covariance parameter. Furthermore, the power of the method to detect heterogeneity was comparable to the power of MGSEM and the type I error rate was close to the nominal rate for all parameters that were equal across subgroups.

The present study goes beyond this earlier research by evaluating the performance of IPC regression to uncover heterogeneity in contemporary dynamic panel models. We show that large individual or group differences can lead to biased IPC regression estimates. To solve this problem, we propose a bias correction procedure. At the formal level, we explicitly derive how IPCs can be calculated for general maximum likelihood estimates that are not limited to SEMs. The explicit derivation is important because it allows expressing the IPCs in terms of derivatives of the likelihood function. In addition, we briefly present a software implementation of IPC regression for the statistical programing language R (R Core Team, 2018). The remainder of this article is organized as follows: first, we will briefly present bivariate dynamic panel models in discrete and continuous time. Second, IPCs regression is formally introduced and its asymptotic properties are discussed. Third, we evaluate the finite-sample properties of IPC regression for dynamic panel models in two simulation studies.

Autoregressive and Cross-Lagged Models for Panel Data

Figure 1 shows a path diagram for a bivariate dynamic panel model for three waves of data. This structural model can be described with the following two equations:

$$x_{i,t} = \beta_{xx} x_{i,t-1} + \beta_{xy} y_{i,t-1} + u_{i,t} \tag{1}$$

$$y_{i,t} = \beta_{yy} y_{i,t-1} + \beta_{yx} x_{i,t-1} + v_{i,t}, \quad i = 1, \ldots, n, \quad t = 2, 3 \tag{2}$$

Here, $x_{i,t}$ and $y_{i,t}$ are the measurements of two different variables of individual i at time point t. For sake of simplicity, we assume that x and y are free of measurement error as well as mean-centered and we do not model any stable differences between individuals.

The regression coefficients β_{xx} and β_{yy} are called autoregressive parameters and they describe the stability in each x and y from one measurement occasion to the next. The regression coefficients β_{xy} and β_{yx} are referred to as cross-lagged or reciprocal effects. The initial assessments of x and y are treated as exogenous variables with zero mean and variance ϕ_{xx}, ϕ_{yy} respectively, and covariance ϕ_{yx}. For the remaining measurement occasions, u and v denote the dynamic error terms. The variance and covariance parameters of the dynamic error terms are symbolized by ψ_{xx}, ψ_{yy}, and ψ_{yx} respectively.

Equations 3a to 3c show the SEM specification of the model in Figure 1.

$$\mathbf{y}_i = \mathbf{B}\mathbf{y}_i + \boldsymbol{\zeta}_i \tag{3a}$$

$$\begin{bmatrix} x_{i,1} \\ y_{i,1} \\ x_{i,2} \\ y_{i,2} \\ x_{i,3} \\ y_{i,3} \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ \beta_{xx} & \beta_{xy} & 0 & 0 & 0 & 0 \\ \beta_{yx} & \beta_{yy} & 0 & 0 & 0 & 0 \\ 0 & 0 & \beta_{xx} & \beta_{xy} & 0 & 0 \\ 0 & 0 & \beta_{yx} & \beta_{yx} & 0 & 0 \end{bmatrix} \begin{bmatrix} x_{i,1} \\ y_{i,1} \\ x_{i,2} \\ y_{i,2} \\ x_{i,3} \\ y_{i,3} \end{bmatrix} + \begin{bmatrix} x_{i,1} \\ y_{i,1} \\ u_{i,2} \\ v_{i,2} \\ u_{i,3} \\ v_{i,3} \end{bmatrix} \tag{3b}$$

$$\mathrm{Cov}(\boldsymbol{\zeta}_i, \boldsymbol{\zeta}_i) = \boldsymbol{\Phi} = \begin{bmatrix} \phi_{xx} & \phi_{yx} & 0 & 0 & 0 & 0 \\ \phi_{yx} & \phi_{yy} & 0 & 0 & 0 & 0 \\ 0 & 0 & \psi_{xx} & \psi_{yx} & 0 & 0 \\ 0 & 0 & \psi_{yx} & \psi_{yy} & 0 & 0 \\ 0 & 0 & 0 & 0 & \psi_{xx} & \psi_{yx} \\ 0 & 0 & 0 & 0 & \psi_{yx} & \psi_{yy} \end{bmatrix} \tag{3c}$$

The resulting model-implied covariance matrix of x and y is given by

$$\mathrm{Cov}(\mathbf{y}_i, \mathbf{y}_i) = \boldsymbol{\Sigma}(\boldsymbol{\theta}) = (\mathbf{I}_6 - \mathbf{B})^{-1}\, \boldsymbol{\Phi} \left((\mathbf{I}_6 - \mathbf{B})^{-1} \right)^{\top}, \tag{4}$$

where $\boldsymbol{\theta}$ is a vector with the model parameters and $\mathbf{I}_6$ denotes an identity matrix of order six.

Although not explicitly stated, the temporal spacing between assessments plays an important role in the model as presented in Figure 1. The model treats time as a discrete variable that indicates the temporally ordering of the assessments and is therefore also referred to as *discrete-time* dynamic panel model. As pointed out elsewhere (e.g., Oud, 2007; Oud & Delsing, 2010; Voelkle, Oud, von Oertzen, & Lindenberger, 2012), treating time as a discrete variable complicates comparing estimates from models with different sample schemes and can bias estimates if assessments are not equally spaced. A solution to these problems is treating time as a continuous variable using stochastic differential equation models (Oud & Jansen, 2000;

for a recent overview of continuous-time modeling in the behavioral and related sciences, see van Montfort, Oud, & Voelkle, 2018). These *continuous-time* dynamic panel models allow estimating continuous-time parameters which can be used to extrapolate to any arbitrary time point.

Following Voelkle, Oud, Davidov, and Schmidt (2012), we specify a continuous-time model by constraining the discrete-time model parameters from Figure 1 to functions of underlying continuous-time parameters $\mathbf{A}$ and $\mathbf{Q}$, and the time intervals Δt_j. The new parameter matrix $\mathbf{A}$ corresponds to the continuous-time version of auto- and cross-lagged effects, the drift parameters, while $\mathbf{Q}$ contains the continuous-time version of dynamic error term variance parameters, or diffusion parameters:

$$\mathbf{A} = \begin{bmatrix} a_{xx} & a_{xy} \\ a_{yx} & a_{yy} \end{bmatrix} \qquad \mathbf{Q} = \begin{bmatrix} q_{xx} & q_{yx} \\ q_{yx} & q_{yy} \end{bmatrix} \tag{5}$$

Let Δt_j be the time interval between the assessments j and $j+1$; then the discrete-time regression coefficients are constrained as a function of $\mathbf{A}$:

$$\begin{bmatrix} b_{xx} & b_{xy} \\ b_{yx} & b_{yy} \end{bmatrix} = \exp(\mathbf{A} \cdot \Delta t_j), \tag{6}$$

where exp denotes the matrix exponential function. The corresponding constraint for the variance of the dynamic error term is

$$\begin{bmatrix} \psi_{xx} & \psi_{yx} \\ \psi_{yx} & \psi_{yy} \end{bmatrix} = \text{irow}\left\{\mathbf{A}_{\#}^{-1}\left[\exp(\mathbf{A}_{\#} \cdot \Delta t_j) - \mathbf{I}_4\right]\text{row}(\mathbf{Q})\right\}, \tag{7}$$

where $\mathbf{A}_{\#} := \mathbf{A} \otimes \mathbf{I}_2 + \mathbf{I}_2 \otimes \mathbf{A}$. The operator *row* puts the elements of $\mathbf{Q}$ into a column vector and the operator *irow* stacks the elements of a vector row-wise into a matrix.

The interpretation of the continuous-time model parameters can be facilitated by transforming them into the discrete-time parameters for an arbitrary time interval Δt_j. For example, plugging $\Delta t_j = 1$ into the estimated drift parameters on the right-hand side of Equation 6 gives the discrete-time regression coefficients for a time interval of one between assessments.

Individual Parameter Contribution Regression

In the following, we will show how heterogeneity in the parameters of dynamic panel models in discrete or continuous time can be identified and explained by IPC regression. To this end, we first motivate the derivation of IPCs for general maximum likelihood estimation. Next, we show how the contributions of SEM parameter estimates can be obtained. Then, we demonstrate that IPC regression can be biased in samples with large individual or group differences and, as a solution to this problem, present a bias correction procedure.

Individual Parameter Contributions to Maximum Likelihood Estimates

Let $\mathbf{y}_1, \ldots, \mathbf{y}_n$ be a sample of independently distributed p-variate random variables with corresponding density functions $f(\boldsymbol{\theta}_1; \mathbf{y}_1), \ldots, f(\boldsymbol{\theta}_n; \mathbf{y}_n)$. IPC regression is applicable in situations where individual differences between the values of the q-variate parameter vector $\boldsymbol{\theta}_i$ can be expressed as a function of a vector of covariates $\mathbf{z}_i$. For instance, if z_i is a single dummy-coded grouping variable, the sample would be generated from a two-group population with group-specific parameter values $\boldsymbol{\theta}_{g1}$ and $\boldsymbol{\theta}_{g2}$.

For sake of illustration, we will assume that f is a multivariate normal density. The associated log-likelihood function for a single individual i is given by

$$\ln \mathcal{L}(\boldsymbol{\theta}; \mathbf{y}_i) = -\frac{1}{2}\left\{\left[\mathbf{y}_i - \boldsymbol{\mu}(\boldsymbol{\theta})\right]^\top \boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\left[\mathbf{y}_i - \boldsymbol{\mu}(\boldsymbol{\theta})\right] + \ln\left[\det(\boldsymbol{\Sigma}(\boldsymbol{\theta}))\right] + p\ln(2\pi)\right\}, \quad (8)$$

with model-implied mean vector $\boldsymbol{\mu}(\boldsymbol{\theta})$ and model-implied covariance matrix $\boldsymbol{\Sigma}(\boldsymbol{\theta})$. In the following, we will use $\boldsymbol{\theta}$ to denote parameter values. True values of the parameters will be marked by a subscript, for instance $\boldsymbol{\theta}_i$, and the maximum likelihood estimate will be denoted by $\widehat{\boldsymbol{\theta}}$.

The first and second derivatives of the log-likelihood function for a given person are important for computing IPCs. The first-order partial derivative of the individual log-likelihood function with respect to the parameters is the score function

$$\mathcal{S}(\boldsymbol{\theta}; \mathbf{y}_i) = \left[\frac{\partial \ln \mathcal{L}(\boldsymbol{\theta}; \mathbf{y}_i)}{\partial \theta^{(1)}} \quad \ldots \quad \frac{\partial \ln \mathcal{L}(\boldsymbol{\theta}; \mathbf{y}_i)}{\partial \theta^{(q)}}\right]^\top, \quad (9)$$

where $\theta^{(j)}$ denotes the j-th element of the parameter vector $\boldsymbol{\theta}$. Evaluation of the score function at specific parameter values measures to which extent an individual's log-likelihood is maximized. Note that the expected values of the score function at the true parameter values are zero, that is $\mathrm{E}\left[\mathcal{S}(\boldsymbol{\theta}_i; \mathbf{y}_i)\right] = \mathbf{0}$ holds for all individuals in the sample. The second-order partial derivative is known as Hessian matrix and will be denoted by

$$\mathcal{H}(\boldsymbol{\theta}; \mathbf{y}_i) = \frac{\partial^2 \ln \mathcal{L}(\boldsymbol{\theta}; \mathbf{y}_i)}{\partial \boldsymbol{\theta}\, \partial \boldsymbol{\theta}^\top}. \tag{10}$$

The expected value of the negative Hessian matrix evaluated at the true individual specific parameter values

$$\mathcal{I}(\boldsymbol{\theta}_i) = \mathrm{E}\left[-\frac{\partial^2 \ln \mathcal{L}(\boldsymbol{\theta}; \mathbf{y}_i)}{\partial \boldsymbol{\theta}\, \partial \boldsymbol{\theta}^\top}\bigg|_{\boldsymbol{\theta}=\boldsymbol{\theta}_i} \right] \tag{11}$$

is called the Fisher information matrix and plays a key role in determining standard errors and asymptotic sampling variance of the maximum likelihood estimates.

The maximum likelihood parameter estimate $\widehat{\boldsymbol{\theta}}$ can be obtained by solving the first order conditions

$$\sum_{i=1}^{n} \mathcal{S}(\widehat{\boldsymbol{\theta}}; \mathbf{y}_i) = \mathbf{0}, \tag{12}$$

such that $\widehat{\boldsymbol{\theta}}$ is an extremum. In homogeneous samples, where $\boldsymbol{\theta}_i = \boldsymbol{\theta}_0$ for $i = 1, \ldots, n$, the resulting parameter estimate $\widehat{\boldsymbol{\theta}}$ is a consistent estimate of true parameter values $\boldsymbol{\theta}_0$. In heterogeneous samples, $\widehat{\boldsymbol{\theta}}$ lies roughly near the mean of the individuals' true parameter values $\boldsymbol{\theta}_1, \ldots, \boldsymbol{\theta}_n$.

The underlying idea behind the derivation of the IPCs is to find the individual roots of the score function instead of finding the roots of the sum of all individual score values as shown in Equation 12. Hypothetically, solving $\mathcal{S}(\widehat{\boldsymbol{\theta}}_i; \mathbf{y}_i) = \mathbf{0}$ for every individual in the sample would yield individual parameter estimates $\widehat{\boldsymbol{\theta}}_1, \ldots, \widehat{\boldsymbol{\theta}}_n$. Unfortunately, for many probability distribution such as the normal distribution, the system of equations $\mathcal{S}(\widehat{\boldsymbol{\theta}}_i; \mathbf{y}_i) = \mathbf{0}$ does not have a unique solution for a single data point. However, we can approximate the individual scores by linearizing the mean of all scores around the maximum likelihood estimate and then disaggregate the resulting expression:

$$\frac{1}{n}\sum_{i=1}^{n}\mathcal{S}(\boldsymbol{\theta};\mathbf{y}_i) \approx \frac{1}{n}\sum_{i=1}^{n}\mathcal{S}(\widehat{\boldsymbol{\theta}};\mathbf{y}_i) + \frac{1}{n}\sum_{i=1}^{n}\mathcal{H}(\widehat{\boldsymbol{\theta}};\mathbf{y}_i)\left(\boldsymbol{\theta} - \widehat{\boldsymbol{\theta}}\right) \tag{13}$$

Without changing the right-hand side of Equation 13, the Hessian matrix can be replaced by the estimated negative Fisher information matrix.

$$\frac{1}{n}\sum_{i=1}^{n}\mathcal{S}(\widehat{\boldsymbol{\theta}};\mathbf{y}_i) - \mathcal{I}(\widehat{\boldsymbol{\theta}})\left(\boldsymbol{\theta} - \widehat{\boldsymbol{\theta}}\right) \tag{14}$$

In geometric terms, Equation 14 approximates the mean of scores with a tangent line at the maximum likelihood estimate. Now, we disaggregate this averaged tangent into n individuals tangents by replacing the mean of scores evaluated at the maximum likelihood estimate with the individual score values evaluated at the maximum likelihood estimate:

$$\mathcal{S}(\widehat{\boldsymbol{\theta}};\mathbf{y}_i) - \mathcal{I}(\widehat{\boldsymbol{\theta}})\left(\boldsymbol{\theta} - \widehat{\boldsymbol{\theta}}\right) \tag{15}$$

Finally, setting Equation 15 to zero and solving for $\boldsymbol{\theta}$ yields the individual parameter contributions which approximate individual parameter values:

$$\begin{aligned} \mathbf{0} &= \mathcal{S}(\widehat{\boldsymbol{\theta}};\mathbf{y}_i) - \mathcal{I}(\widehat{\boldsymbol{\theta}})\left(\mathcal{IPC}(\widehat{\boldsymbol{\theta}};\mathbf{y}_i) - \widehat{\boldsymbol{\theta}}\right) \\ \mathcal{IPC}(\widehat{\boldsymbol{\theta}};\mathbf{y}_i) &= \widehat{\boldsymbol{\theta}} + \mathcal{I}(\widehat{\boldsymbol{\theta}})^{-1}\mathcal{S}(\widehat{\boldsymbol{\theta}};\mathbf{y}_i) \end{aligned} \tag{16}$$

Figure 2 illustrates the derivation of the IPCs using the rate parameter of the exponential distribution as an arbitrary example.

The interpretation or meaning of the IPCs, and all averages or statistics based on them, follows from the interpretation of the pooled parameter estimates. This property is particularly important for dynamic panel models. The IPCs of autoregressive or cross-lagged parameter will only approximate the individual within-person relationship if the pooled dynamic model separates the within-person process from stable between-person differences (Hamaker et al., 2015).

Individual Parameter Contributions to Structural Equation Model Parameter Estimates

Instead of the sum of individual log-likelihoods in Equation 8, it is common to use the aggregated log-likelihood function (also called fitting function) in structural equation modeling (Voelkle, Oud, von Oertzen, & Lindenberger, 2012). The maximum likelihood fitting function for multivariate normally distributed variables is

$$\mathcal{F}\left(\bar{\mathbf{y}}, \mathbf{S}, \boldsymbol{\mu}(\boldsymbol{\theta}), \boldsymbol{\Sigma}(\boldsymbol{\theta})\right) = \left[\bar{\mathbf{y}} - \boldsymbol{\mu}(\boldsymbol{\theta})\right]^\top \boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1} \left[\bar{\mathbf{y}} - \boldsymbol{\mu}(\boldsymbol{\theta})\right] + \operatorname{tr}\left(\mathbf{S}\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\right) - \ln\left(|\mathbf{S}\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}|\right) - p, \tag{17}$$

with sample means $\bar{\mathbf{y}}$ and sample covariance matrix $\mathbf{S}$ (Yuan & Bentler, 2007). Optimizing either the sum of individual log-likelihood functions or an aggregated fitting function yields equivalent parameter estimates (Bollen, 1989).

Using the aggregated fitting function, IPCs to SEM parameter estimates are a function of the individual's data and two matrices $\boldsymbol{\Delta}$ and $\mathbf{V}$ that are provided by most standard structural equation modeling software packages. The first matrix $\boldsymbol{\Delta}$ is the following Jacobian matrix

$$\boldsymbol{\Delta} = \frac{\partial[\boldsymbol{\mu}(\boldsymbol{\theta}), \boldsymbol{\sigma}(\boldsymbol{\theta})]^\top}{\partial\boldsymbol{\theta}}, \tag{18}$$

where $\boldsymbol{\sigma}(\boldsymbol{\theta})$ denotes the half-vectorized model-implied covariance matrix. $\boldsymbol{\Delta}$ indicates the sensitivity of the model-implied mean vector and covariance matrix to changes in the parameters. The second matrix is the weight matrix $\mathbf{V}$ which depends on the chosen estimator (e.g., Savalei, 2014). In SEMs estimated with normal theory maximum likelihood, $\mathbf{V}$ is

$$\mathbf{V} = \begin{bmatrix} \boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1} & \mathbf{0} \\ \mathbf{0} & \frac{1}{2}\mathbf{D}_p^\top \left[\boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1} \otimes \boldsymbol{\Sigma}(\boldsymbol{\theta})^{-1}\right] \mathbf{D}_p \end{bmatrix}, \tag{19}$$

with duplication matrix $\mathbf{D}_p$ (Magnus & Neudecker, 2007). Sample estimates of $\boldsymbol{\Delta}$ and $\mathbf{V}$ can be obtained by replacing $\boldsymbol{\theta}$ with $\widehat{\boldsymbol{\theta}}$.

Following Satorra (1989) and Neudecker and Satorra (1991), the Fisher information matrix can be expressed as $\mathcal{I}(\boldsymbol{\theta}) = \boldsymbol{\Delta}^\top\mathbf{V}\boldsymbol{\Delta}$ and a partial derivative of the fitting function is given by

$$-\frac{1}{2}\frac{\partial \mathcal{F}\left(\bar{\mathbf{y}}, \mathbf{S}, \boldsymbol{\mu}(\boldsymbol{\theta}), \boldsymbol{\Sigma}(\boldsymbol{\theta})\right)}{\partial \boldsymbol{\theta}} = \boldsymbol{\Delta}^\top \mathbf{V}\left(\begin{bmatrix}\bar{\mathbf{y}}\\ \mathbf{s}\end{bmatrix} - \begin{bmatrix}\boldsymbol{\mu}(\boldsymbol{\theta})\\ \boldsymbol{\sigma}(\boldsymbol{\theta})\end{bmatrix}\right). \tag{20}$$

Individual score values can be obtained by replacing the aggregated mean vector and covariance matrix in Equation 20 by the individual contributions to these sample moments. To this end, we define n vectors

$$\mathbf{d}_i := \begin{bmatrix}\mathbf{y}_i\\ \text{vech}\left(\begin{bmatrix}\mathbf{y}_i - \bar{\mathbf{y}}\end{bmatrix}\begin{bmatrix}\mathbf{y}_i - \bar{\mathbf{y}}\end{bmatrix}^\top\right)\end{bmatrix} \tag{21}$$

(Satorra, 1992), where the operator vech half-vectorizes a symmetric matrix. Note that the averaged individual contributions to the sample moments are identical to the observed sample moments, that is $\frac{1}{n}\sum_{i=1}^{n}\mathbf{d}_i = \begin{bmatrix}\bar{\mathbf{y}} & \mathbf{s}\end{bmatrix}^\top$ [1]. Thus, analogous to Equation 16, the individual contributions to SEM parameter estimates can be estimated by

$$\mathcal{IPC}(\widehat{\boldsymbol{\theta}}; \mathbf{y}_i) = \widehat{\boldsymbol{\theta}} + \left(\widehat{\boldsymbol{\Delta}}^\top \widehat{\mathbf{V}}\widehat{\boldsymbol{\Delta}}\right)^{-1}\widehat{\boldsymbol{\Delta}}^\top \widehat{\mathbf{V}}\left(\mathbf{d}_i - \begin{bmatrix}\boldsymbol{\mu}(\widehat{\boldsymbol{\theta}})\\ \boldsymbol{\sigma}(\widehat{\boldsymbol{\theta}})\end{bmatrix}\right). \tag{22}$$

The above definition of the IPCs should replace that given by Oberski (2013), which yields incorrect means of the IPCs to factor loading and regression parameters.

Predicting Heterogeneity in Panel Models with IPC Regression

The IPCs of a single individual are usually plagued by random fluctuation and will most likely be poor estimates of the true individual parameter values. However, studying the IPCs of groups of individuals or jointly modeling the IPCs of the whole sample can average out this noise. One obvious method for revealing meaningful differences in the parameters is linear regression. Regressing the IPCs on a set of additional covariates $\mathbf{z}$ allows to test and estimate if and how individual parameter values vary as a function of $\mathbf{z}$.

For instance, we could investigate via IPC regression whether the cross-lagged estimated effect $\widehat{\beta}_{yx}$ from x on y in the model shown in Figure 1 differs between women

[1] The biased estimate of the sample covariance is used.

and men. To this end, the IPCs to $\widehat{\beta}_{yx}$ are regressed on a dummy variable z representing gender. Using women as a baseline group, the following IPC regression equation is estimated

$$\mathcal{IPC}(\widehat{\beta}_{yx}; \mathbf{y}_i) = \widehat{\gamma}_0 + \widehat{\gamma}_1 z_i + \nu_i, \tag{23}$$

where ν_i is a random residual with mean zero. In the above equation, the IPC regression intercept $\widehat{\gamma}_0$ is the estimated value of β_{yx} for women and $\widehat{\gamma}_1$ denotes the estimated difference between between women and men in β_{yx}. In other words, the IPC regression slope estimate $\widehat{\gamma}_1$ is a measure of heterogeneity in the cross-lagged effect $\widehat{\beta}_{yx}$ with respect to the covariate gender. As in standard regression analysis, a t-test could be applied to test $\widehat{\gamma}_1$, that is, to infer whether the estimated subgroup difference between women and men in $\widehat{\beta}_{yx}$ is significantly different from zero. In this setup, Oberski (2013) showed that $\widehat{\gamma}_1$ and its Wald statistic are equivalent to the robust expected parameter change and robust modification index familar from MGSEM (Satorra, 1989). Based on the size of the estimate and the test result, an informed decision can be made to modify the original model or not. An obvious choice of modification would be to use age as a grouping variable in a MGSEM. Note that the estimated cross-lagged effect of men can be obtained by summing $\widehat{\gamma}_0$ and $\widehat{\gamma}_1$.

The partial effects of several covariates on the parameters can be investigated using multiple linear regression analysis. To investigate parameter heterogeneity in the complete model presented in Figure 1, an IPC regression equation needs to be estimated for each of the ten model parameters:

$$\mathcal{IPC}(\widehat{\beta}_{xx}; \mathbf{y}_i) = \widehat{\boldsymbol{\gamma}}_{\beta_{xx}}^{\top} \mathbf{z}_i + \nu_{i,\beta_{xx}} \tag{24a}$$

$$\mathcal{IPC}(\widehat{\beta}_{yx}; \mathbf{y}_i) = \widehat{\boldsymbol{\gamma}}_{\beta_{yx}}^{\top} \mathbf{z}_i + \nu_{i,\beta_{yx}} \tag{24b}$$

$$\vdots$$

$$\mathcal{IPC}(\widehat{\psi}_{yy}; \mathbf{y}_i) = \widehat{\boldsymbol{\gamma}}_{\psi_{yy}}^{\top} \mathbf{z}_i + \nu_{i,\psi_{yy}} \tag{24j}$$

In Equation 24a to 24j, the IPC regression estimates $\widehat{\boldsymbol{\gamma}}$ indicates the estimated effects from multiple covariates $\mathbf{z}$ on a certain parameter estimate.

Due to its flexibility and computational efficiency, the linear regression framework offers researchers many possibilities to investigate heterogeneity by means of IPC regression. The interplay of the covariates could be studied by adding interactions to Equation 24a to 24j. Furthermore, higher order polynomial terms, such as quadratic or cubic terms can be easily specified to test for nonlinear relationships. If the number of covariates is large, regularization techniques like lasso (Tibshirani, 1996) could be used to aid the selection of important covariates. Finally, latent variables could be included by replacing the regression equations above with SEMs.

Bias of Individual Parameter Contributions

IPCs as estimators of individual- or group-specific parameter values can be systematically biased under certain circumstances. As a rule of thumb, IPCs to a parameter estimate will be biased if they depend on a pooled estimate of a heterogeneous parameter. The size of this bias depends on the distance between the maximum likelihood estimate and the true individual- or group-specific value of the parameter. It is important to note that this bias is unaffected by sample size. As a result, IPC regression estimates based on biased IPCs will also be systematically biased. In the following, we illustrate the bias with the help of a simple example.

Consider the individuals $y_1, \ldots, y_n$ sampled from the corresponding exponential distributions $f(\lambda_1; y_1), \ldots, f(\lambda_n; y_n)$. In order to determine the bias of the IPCs, we first derive the conditional expected value of the IPC from individual i given the pooled maximum likelihood estimate of λ. Using $\mathrm{E}[y_i] = \lambda_i^{-1}$, it follows that

$$\begin{aligned}
\mathrm{E}\left[\mathcal{IPC}(\widehat{\lambda}; y_i) \,\middle|\, \widehat{\lambda}\right] &= \mathrm{E}\left[\widehat{\lambda} + \mathcal{I}(\widehat{\lambda})^{-1}\mathcal{S}(\widehat{\lambda}; y_i) \,\middle|\, \widehat{\lambda}\right] \\
&= \mathrm{E}\left[\widehat{\lambda} + \widehat{\lambda}^2\left(\frac{1}{\widehat{\lambda}} - y_i\right) \,\middle|\, \widehat{\lambda}\right] \\
&= 2\widehat{\lambda} - \frac{\widehat{\lambda}^2}{\lambda_i}. \qquad (25)
\end{aligned}$$

Next, we define the distance between the true individual parameter value and the pooled maximum likelihood estimate as $\delta_i := \widehat{\lambda} - \lambda_i$. By substituting $\widehat{\lambda} = \lambda_i + \delta_i$ in

Equation 25, we obtain:

$$\mathrm{E}\left[\mathcal{IPC}(\widehat{\lambda}; y_i) \,\middle|\, \widehat{\lambda}\right] = \lambda_i - \frac{\delta_i^2}{\lambda_i} \tag{26}$$

Thus, the bias of the IPCs is $-\delta_i^2/\lambda_i$, ignoring the sampling error of $\widehat{\lambda}$. In homogeneous samples, this bias vanishes since $\widehat{\lambda}$ approaches the true parameter value λ_0 as sample size increases. In heterogeneous samples, however, the bias persists as $\widehat{\lambda}$ does not approach any of the true individual parameter values. Moreover, the bias grows with the distance between the true individual- or group-specific parameter value from the pooled maximum likelihood estimate.

It should be noted that the expected biases will typically be small and only large differences in the true parameters will substantially bias the IPCs. For instance, consider a sample drawn in equal shares from two different populations with $\lambda_{g1} = 0.5$ and $\lambda_{g2} = 1.5$. These parameter values imply that the expected value of the pooled maximum likelihood estimate $\widehat{\lambda}$ is 0.75. The expected value of IPCs from the first group is 0.375 and 1.125 for IPCs from the second. Notice that both, the group-specific estimates and the size of the group difference, are underestimated.

Deriving the bias of IPC regression for more complex models such as SEMs is challenging. However, later in the manuscript we will demonstrate by means of Monte Carlo simulations that the results derived above generalize to dynamic panel models.

Iterative IPC Regression: Bias Correction Procedure

Here, we propose an iterative bias correction procedure similar to the Fisher scoring algorithm (e.g., Demidenko, 2013). As discussed in the previous section, IPCs are systemically biased if they depend on estimates of heterogeneous parameters. However, this bias can be removed by replacing the pooled maximum likelihood estimates with individual- or group-specific parameter estimates. Such estimates can be obtained in a stepwise fashion. First, IPC regression is conducted as outlined above. Second, the IPC regression estimates are used to predict individual- or group-specific parameter estimates which are then used to re-calculate the IPCs. By alternately re-calculating the IPCs and re-estimating the IPC regression estimates, the bias of the

IPCs is filtered out. A graphical demonstration for this bias correction procedure is presented in Figure 3.

The resulting updating procedure, which we call iterative IPC regression, can be summarized as follows:

1. Fit a SEM and obtain pooled parameter estimates $\widehat{\boldsymbol{\theta}}$.
2. Calculate IPCs as a function of the data $\mathbf{y}_1, \ldots, \mathbf{y}_n$ and the parameter estimates $\widehat{\boldsymbol{\theta}}$.
3. Estimate the initial IPC regression estimates $\widehat{\boldsymbol{\gamma}}$.
4. Predict individual- or group-specific parameter estimates $\widehat{\boldsymbol{\theta}}_g$ as a function of the covariates $\mathbf{z}$ and the IPC regression estimates $\widehat{\boldsymbol{\gamma}}$.
5. Use those individual- or group-specific parameter estimates $\widehat{\boldsymbol{\theta}}_g$ to re-calculate the IPCs.
6. Pool the re-calculated IPCs and re-estimate the IPC regression estimates $\widehat{\boldsymbol{\gamma}}$.
7. Repeat step 4 to 6 until $\widehat{\boldsymbol{\gamma}}$ converges.

The algorithm converges if the change in $\widehat{\boldsymbol{\gamma}}$ becomes negligibly smaller than a pre-specified numerical threshold. Unfortunately, the iterated IPC algorithm does not always converge. Especially, if the true individual- or group-specific value of a parameter lies close to (or at) the border of its parameter space, the algorithm might go awry. However, given strong heterogeneity in a sample, we observed in simulation across various models that the iterations often yield substantial improvement over the initial IPC regression estimates before breaking down. Therefore, we recommended to compute the log-likelihood value for every iteration. Even if the algorithm fails to converge, the iteration with the largest log-likelihood might be preferred to the initial results.

We would like to note two more observations on the behavior of IPC. First, IPC regression estimates are unbiased in homogeneous samples and therefore cannot be further improved by updating the IPCs. If iterated IPC regression is used in a homogeneous sample, the algorithm will overfit the estimates to random fluctuation of the data. In this case, the resulting estimates can be marginally worse than the initial estimates, but the difference will be inconsequential for most practical purposes. Second, updating the IPCs comes at the cost of additional computational demands. In our experience, however, the algorithm usually converges quickly with few iterations needed. Even for samples with a few thousand individuals and models with more than thirty parameters, updating the IPCs took less than a minute with a standard desktop PC.

Software Implementation

IPC regression is implemented as a package for the statistical programming language R (R Core Team, 2018), termed *ipcr*. The *ipcr* package makes it easy for researchers to study heterogeneity in the parameter estimates of a SEM fitted with the *OpenMx* package (Neale et al., 2015). The *ipcr* package performs "vanilla" IPC regression as introduced by Oberski (2013) as well as iterated IPC regression. More information of how the *ipcr* package can be installed and used can be found under `https://github.com/manuelarnold/ipcr/`.

Monte Carlo Simulations

To evaluate the performance of vanilla and iterated IPC regression to detect and estimate heterogeneity in dynamic panel models in discrete and continuous time we conducted the following two Monte Carlo simulations. The first simulation aims to substantiate our theoretical considerations regarding the bias for bivariate dynamic panel models. The second simulation investigates whether IPC regression provides valid inferences and compares the power of the method with MGSEM.

Simulation I: Demonstration of the Bias

In the following simulation studies, we used the discrete-time dynamic panel model depicted in Figure 1 with five measurement waves (instead of three) as a simulation model. The data were sampled from a multivariate normal distribution with two distinct sets of parameter values. 125 observations were generated per group, resulting in a pooled sample with 250 observations in total. A discrete-time and a continuous-time dynamic panel model were fitted to the same data, ignoring the group differences. Then, we used vanilla and iterated IPC regression with a dummy variable to recover the group differences in the parameter values of the dynamic panel models. We repeated this procedure 10,000 times.

The discrete-time population parameter values used to generate the data are shown in the upper half of Table 1, separated for both groups. For easy reference, we transformed these parameter values into continuous time and printed them in the lower half of the table. As clearly apparent from Table 1, group 1 and 2 differ substantively. The first group is characterized by strong autoregressive coefficients and no cross-lagged effects, whereas the second group exhibits substantial cross-lagged effects and smaller autoregressive coefficients. In addition, the variance of x and y was chosen twice as high for the second group as compared to the first.

We will first discuss the results for the discrete-time dynamic panel model. As expected from the theoretical example, both IPC regression methods provided accurate estimates of heterogeneity in the initial variance and covariance parameters. Further, IPC regression estimates for regression coefficients and dynamic error term variance parameters were slightly distorted. Figure 4 depicts boxplots visualizing the bias of the IPC methods for regression coefficients (top graph) and dynamic error term variance parameters (lower graph). The estimates of vanilla IPC regression are printed in red and estimates after updating the IPCs are depicted in blue. Boxplots whose median lines lie close to the dotted black line indicate that the corresponding IPC regression estimates were approximately unbiased. Using the vanilla method, the intercepts (marked with the subscript 0) of the IPC regression equations were more biased than the slopes (subscript 1). Our updated IPC method erased the bias in the intercepts and provided accurate estimates for all type of model parameters. Averaged over all parameters, the root mean squared error of iterated IPC regression (RMSE = 0.089) was slightly smaller than the one of the vanilla procedure without updating (RMSE = 0.094).

The performance of the IPC regression methods for the continuous-time dynamic model was similar to the findings for the discrete-time parameters above. The estimates for the initial variance and covariance parameters provided by both IPC regression methods were near the true values, whereas estimates for the remaining model parameters were biased. Figure 5 presents the bias in the IPC regression estimates for drift and diffusion parameters. Overall, the IPC regression estimates showed more variability for the continuous-time parameters than for the discrete-time parameters. As for the discrete-time model, vanilla IPC regression exhibited a slight bias. Updating the IPCs reduced this bias at the cost of increased variability of the IPC regression estimates. Moreover, the updating algorithm converged only in 53.78% of the trials and fell back to the starting values or an intermediate solution in the remaining trials. Nevertheless, in terms of the RMSE averaged over all parameters, iterated IPC regression (RMSE = 0.168) slightly outperformed vanilla IPC regression (RMSE = 0.174).

Simulation II: Statistical Power and False Positive Rate

In the second simulation, we investigated the power to detect a difference in a parameter value and the false positive rate for homogeneous parameters of the IPC regression methods. We generated multivariate normal data from bivariate dynamic panel models with five measurement occasions. We specified the population models in a way that only the cross-lagged effects from the variable x on y differed slightly between two groups. All other parameters were equal. In contrast to the previous simulation, we used different population models for the discrete- and continuous-time models. The corresponding parameter values for both population models (shown in Table 2) resulted in similar but not identical population covariance matrices. After a data set was generated, a pooled dynamic panel model was fitted, and parameter heterogeneity was tested with IPC regression (vanilla and iterated) using a dummy variable. We investigated power and false positive rate for group sizes of 100, 125, 150, 175, and 200 resulting in total sizes of 200, 250, 300, 350, and 400. For each sample size, we replicated this process 10,000 times.

As a reference, we compared the power of the IPC regression methods to the power of MGSEM. Although MGSEM lacks the flexibility and computational simplicity of IPC regression, in simple (single-variable) group comparisons with correctly specified models, standard maximum-likelihood theory suggests it should provide the uniformly most powerful test. MGSEM therefore presents a good gold standard reference for these cases. The MGSEMs were specified by letting only the cross-lagged effects of x on y differ between groups. We computed the power of the MGSEMs by conducting likelihood ratio tests that compared the fit of the MGSEMs to the fit of the pooled models.

Figure 6 shows the power of IPC regression for the discrete-time model. Depicted is the rejection rate of the null hypothesis that the cross-lagged effects from x on y are equal in both groups, plotted against the number of individuals for a significance level of 5%. Red lines refer to the power of vanilla IPC regression, blue lines to iterated IPC regression, and black lines mark the power of MGSEM. For the discrete-time model, the IPC regression methods appeared to be on average 3.97 percentage points (range: [3.03, 5.30]) percentage points less powerful than MGSEM. Iterated IPC regression achieved a marginally larger power with a difference of 0.66 percentage points (range: [0.35, 0.94]). The power for the continuous-time model is presented in Figure 7. We found that the difference in power between the IPC regression methods and MGSEM were substantively larger for the continuous-time model than for the discrete-time model. On average, the power of the IPC regression was 20.68 percentage points (range: [14.25, 27.47]) smaller than the power of MGSEM. In addition, the power of IPC regression appeared to grow more slowly as a function of sample size. Again, iterated IPC appeared slightly more powerful than vanilla IPC regression (average difference: 0.28, range: [0.17, 0.44]).

Besides power, the false detection rate of the IPC regression methods is of great importance for drawing correct conclusions from the data. We assessed the type I error rate for population parameters that are identical in the two groups for a significance level of 5%. We summarize the results by averaging the type I error rate for the three parameter types in the models (initial variance, regression coefficients/drift, dynamic error term variance/diffusion). Table 3 shows the proportions of type I errors for the discrete-time model and Table 4 for the continuous-time model. As suggested by the theoretical results in (Oberski, 2013), the type I error rates were close to 5% for most parameters. Iterated IPC regression committed slightly more type I errors for regression and drift parameters. These findings imply that the standard errors of iterated IPC regression for regression/drift parameters were slightly too small and could explain why iterated IPC regression appeared marginally more powerful to detect heterogeneity.

In contrast to Simulation I, there was not a single case of non-convergence of the iterated IPC regression algorithm in Simulation II. This finding suggests that the convergence problems for the continuous-time dynamic panel model were mainly driven by the larger group differences used in the previous simulation.

Discussion

The present study investigated the performance of IPC regression (Oberski, 2013) to identify and estimate heterogeneity in dynamic panel models. Overall, we found that IPC regression is a promising method to identify and estimate heterogeneity in model parameters. In comparison to other contemporary approaches formally addressing heterogeneity, IPC regression fills an important niche for applied researchers: IPC regression offers a general framework that encompasses all kind of model parameter (e.g., regression coefficients, intercepts, variances, covariances) and makes identifying and explaining heterogeneity as simple and fast as linear regression.

IPC regression was evaluated in terms of bias in the recovery of true group differences, the power to detect parameter heterogeneity, and the type I error rate for homogeneous parameters. By means of a theoretical example and through Monte Carlo simulations, we demonstrated that original, “vanilla” IPC regression estimates can be slightly biased due to large differences in regression parameters. Additional heterogeneity in variance parameters may amplify this bias. As a rule of thumb, the bias seems to affect mainly parameters connected to endogenous variables like regression and residual variance parameters, whereas the IPC regression estimates for parameters associated with exogenous variables such as the initial variance parameters remain comparatively unbiased. Hence, IPC regression may perform worse for SEMs with many directed paths such as dynamic panel models than for models with few directed paths such as CFA models. This argument would also explain why Oberski (2013) found nearly unbiased estimates of group differences in a CFA model.

To correct the bias, in vanilla IPC regression, we introduced a novel updating procedure, which we termed iterated IPC regression. Iterated IPC regression produced approximately unbiased estimates of group differences in the parameters of a discrete-time dynamic panel model and outperformed vanilla IPC regression in terms of the RMSE. For the continuous-time dynamic panel model, however, iterated IPC regression corrected the bias but at the cost of adding additional variability to the estimates. Nevertheless, updating the IPCs still improved the estimates on average as indicated by a smaller RMSE.

In situations in which MGSEM could be applied as an alternative to IPC regression, we compared the power of IPC regression to that of MGSEM, which theory suggests is uniformly most-powerful in these cases. IPC regression yielded power only slightly below that of this theoretically optimal method to detect group differences in the cross-lagged effect of a discrete-time dynamic panel model. For the continuous-time model, however, IPC regression was no more than half as powerful as MGSEM. It should be noted that MGSEM cannot be applied to all scenarios allowed by IPC regression; for example, MGSEM does not investigate partial effects of multiple covariates of model parameters. In agreement with earlier theoretical findings, both IPC regression methods did control the Type I error rate accurately.

In summary, our findings demonstrate that (iterated) IPC regression is a useful tool to study heterogeneity in discrete-time dynamic panel model. For continuous-time dynamic panel models, however, our findings were mixed: high variance caused by the bias correction procedure and a small power make (iterated) IPC regression unappealing especially in smaller data sets. We believe that these problems are caused by non-linear parameter constraints and high correlation between parameter estimates of the continuous-time dynamic panel model. Considering these difficulties, IPC regression seems more appropriate for models that can be parameterized without non-linear constraints such as the discrete-time dynamic panel model or many other contemporaneous models for longitudinal data such as latent growth curve models (Bollen & Curran, 2006) or latent change score models (McArdle, 2001), if these models are applicable.

Although IPC regression is a general, easy to use, and flexible approach to heterogeneity, we want to stress that it is not always the most appropriate one. Depending on a study's objective, other methods for addressing heterogeneity should be preferred to IPC regression. For example, multilevel models appear like an obvious choice in situations where it is sufficient to allow for varying parameter values between individuals and there is no interest in explaining these differences. In contrast, if a study aims to test differences between few known groups in the data (e.g., in variance parameters), MGSEM will often be the better choice. If a study's goal is to determine homogeneous groups in the data with help of additionally observed covariates, partitioning methods like SEM trees or forests often are better suited for the task, in particular if computation time is not an issue.

Moreover, we will briefly touch upon some limitations of IPC regression that researchers should consider. First, the usefulness of IPC regression depends on the covariates available. If none of the additional covariates is related to individual or group differences in the data set, IPC regression will fail to detect the source of heterogeneity. In cases of unobserved group membership, researchers may resort to methods like finite mixture models (Jedidi, Jagpal, & DeSarbo, 1997; Lubke & Muthén, 2005; Muthén & Shedden, 1999). Second, IPC regression is a data-driven or exploratory procedure and therefore susceptible to capitalize on chance characteristics of the data set (MacCallum, Roznowski, & Necowitz, 1992). Modifying models by blindly following the advice of IPC regression may lead to a model that works well in the observed sample but does not generalize to others. We thus recommend paying not only close attention to the p-value provided by IPC regression, but also to the size of the estimated individual or group difference. See also Saris, Satorra, and van der Veld (2009), for a related discussion about model modification using the modification index and expected parameter change. Third, using IPC regression to investigate the effect of a large number of covariates on complex models with many parameters will yield large number of IPC regression estimates that can be challenging to interpret. Regularization techniques such as lasso (Tibshirani, 1996) could be used to find a subset of the most important covariates.

In summary, however, we believe that IPC regression is a useful tool to investigate parameter heterogeneity in SEMs for longitudinal data such as dynamic panel models that combines flexibility with its unique computational simplicity.

References

Allison, P. D., Williams, R., & Moral-Benito, E. (2017). Maximum likelihood for cross-lagged panel models with fixed effects. *Socius: Sociological Research for a Dynamic World*, *3*, 1–17.

Bollen, K. A. (1989). *Structural equation with latent variables.* New York: Wiley.

Bollen, K. A., & Brand, J. E. (2010). A general panel model with random and fixed effects: A structural equations approach. *Social Forces; a Scientific Medium of Social Study and Interpretation*, *89*(1), 1–34.

Bollen, K. A., & Curran, P. J. (2006). *Latent curve models: A structural equation perspective.* Hoboken, NJ: John Wiley & Sons.

Brandmaier, A. M., Driver, C. C., & Voelkle, M. C. (2018). Recursive partitioning in continuous time analysis. In K. van Montfort, J. H. Oud, & M. C. Voelkle (Eds.), *Continuous time modeling in the behavioral and related sciences* (pp. 259–282). New York: Springer.

Brandmaier, A. M., Prindle, J. J., McArdle, J. J., & Lindenberger, U. (2016). Theory-guided exploration with structural equation model forests. *Psychological Methods*, *21*(4), 566–582.

Brandmaier, A. M., von Oertzen, T., McArdle, J. J., & Lindenberger, U. (2013). Structural equation model trees. *Psychological Methods*, *18*(1), 71–86.

Brown, T. A. (2006). *Confirmatory factor analysis for applied research.* New York: Guilford.

Demidenko, E. (2013). *Mixed models: Theory and applications with R* (2nd ed.). Hoboken, NJ: Wiley.

George, M. R. W., Yang, N., Jaki, T., Feaster, D. J., Lamont, A. E., Wilson, D. K., & van Horn, M. L. (2013). Finite mixtures for simultaneously modelling differential effects and non-normal distributions. *Multivariate Behavioral Research*, *48*(6), 816–844.

Granger, C. W. J. (1969). Investigating causal relations by econometric models and cross-spectral methods. *Econometrica*, *37*(3), 424–438.

Halaby, C. N. (2004). Panel models in sociological research: Theory into practice. *Annual Review of Sociology*, *30*(1), 507–544.

Hamaker, E. L., Kuiper, R. M., & Grasman, R. P. P. P. (2015). A critique of the cross-lagged panel model. *Psychological Methods*, *20*(1), 102–116.

Hsiao, C. (2014). *Analysis of panel data* (3rd ed.). Cambridge: University Press.

Jedidi, K., Jagpal, H. S., & DeSarbo, W. S. (1997). Finite-mixture structural equation models for response-based segmentation and unobserved heterogeneity. *Marketing Science*, *16*(1), 39–59.

Lubke, G. H., & Muthén, B. (2005). Investigating population heterogeneity with factor mixture models. *Psychological Methods*, *10*(1), 21–39.

MacCallum, R. C., Roznowski, M., & Necowitz, L. B. (1992). Model modifications in covariance structure analysis: The problem of capitalization on chance. *Psychological Bulletin*, *111*(3), 490–504.

Magnus, J. R., & Neudecker, H. (2007). *Matrix differential calculus with applications in statistics and econometrics* (3rd ed.). Chichester: Wiley.

McArdle, J. J. (2001). A latent difference score approach to longitudinal dynamic structural analysis. In R. Cudeck, S. Du Toit, & D. Sörbom (Eds.), *Structural equation modeling* (pp. 342–380). Lincolnwood: Scientific Software International.

Merkle, E. C., Fan, J., & Zeileis, A. (2014). Testing for measurement invariance with respect to an ordinal variable. *Psychometrika*, *79*(4), 569–584.

Merkle, E. C., & Zeileis, A. (2013). Tests of measurement invariance without subgroups: A generalization of classical methods. *Psychometrika*, *78*(1), 59–82.

Muthén, B. O., & Shedden, K. (1999). Finite mixture modeling with mixture outcomes using the EM algorithm. *Biometrics*, *55*(2), 463–469.

Neale, M. C., Hunter, M. D., Pritkin, J., Zahery, M., Brick, T. R., Kirkpatrick, R. M., ... Boker, S. M. (2015). Openmx 2.0: Extended structural equation and statistical modeling. *Psychometrika*, *81*(2), 535–549.

Neudecker, H., & Satorra, A. (1991). Linear structural relations: Gradient and hessian of the fitting function. *Statistics & Probability Letters*, *11*, 57–61.

Oberski, D. L. (2013). *A flexible method to explain differences in structural equation model parameters over subgroups.* Retrieved 14.12.2018, from http://daob.nl/wp-content/uploads/2013/06/SEM-IPC-manuscript-new.pdf

Oud, J. H. L. (2007). Continuous time modeling of reciprocal relationshiops in the cross-lagged panel design. In S. M. Boker & M. J. Wenger (Eds.), *Data analtic techniques for dynamic systems in the social and behavioral sciences* (pp. 87–129). Mahwah: Erlbaum.

Oud, J. H. L., & Delsing, M. J. M. H. (2010). Continuous time modeling of panel data by means of SEM. In K. van Montfort, J. H. L. Oud, & A. Satorra (Eds.), *Longitudinal research with latent variables* (pp. 201–244). New York: Springer.

Oud, J. H. L., & Jansen, R. A. R. G. (2000). Continuous time state space modeling of panel data by means of SEM. *Psychometrika*, *65*(2), 199–215.

R Core Team. (2018). *R: A language and environment for statistical computing.* Vienna, Austria): R Foundation for Statistical Computing. Retrieved from https://www.R-project.org/

Saris, W. E., Satorra, A., & Sörbom, D. (1987). The detection and correction of specification errors in structural equation models. *Sociological Methodology*, *17*, 105–129.

Saris, W. E., Satorra, A., & van der Veld, W. M. (2009). Testing structural equation models or detection of misspecifications? *Structural Equation Modeling: A Multidisciplinary Journal*, *16*(4), 561–582.

Satorra, A. (1989). Alternative test criteria in covariance structure analysis: A unified approach. *Psychometrika*, *54*(1), 131–151.

Satorra, A. (1992). Asymptotic robust inferences in the analysis of mean and covariance structures. *Sociological Methodology*, *22*, 249–278.

Savalei, V. (2014). Understanding robust corrections in structural equation modeling. *Structural Equation Modeling: A Multidisciplinary Journal*, *21*(1), 149–160.

Singer, J. D., & Willet, J. B. (2003). *Applied longitudinal data analysis.* Oxford:

University Press.

Sörbom, D. (1974). A general method for studying differences in factor means and factor structure between groups. *British Journal of Mathematics and Statistical Psychology*, *27*(2), 229–239.

Sörbom, D. (1989). Model modification. *Psychometrika*, *54*(3), 371–384.

Tibshirani, R. (1996). Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society. Series B (Methodological)*, *58*(1), 267–288.

Usami, S., Hayes, T., & McArdle, J. J. (2017). Fitting structural equation model trees and latent growth curve mixture models in longitudinal designs: The influence of model misspecification. *Structural Equation Modeling: A Multidisciplinary Journal*, *24*(4), 585–598.

van Montfort, K., Oud, J. H., & Voelkle, M. C. (2018). *Continuous time modeling in the behavioral and related sciences*. New York: Springer.

Voelkle, M. C., Oud, J. H. L., Davidov, E., & Schmidt, P. (2012). An SEM approach to continuous time modeling of panel data: Relating authoritarianism and anomia. *Psychological Methods*, *17*(2), 176–192.

Voelkle, M. C., Oud, J. H. L., von Oertzen, T., & Lindenberger, U. (2012). Maximum likelihood dynamic factor modeling for arbitrary N and T using SEM. *Structural Equation Modeling: A Multidisciplinary Journal*, *19*(3), 329–350.

Wang, T., Merkle, E. C., & Zeileis, A. (2014). Score-based tests of measurement invariance: Use in practice. *Frontiers in Psychology*, *5*, 438.

Wang, T., Strobl, C., Zeileis, A., & Merkle, E. C. (2018). Score-based tests of differential item functioning via pairwise maximum likelihood estimation. *Psychometrika*, *83*(1), 132–155.

Yuan, K.-H., & Bentler, P. M. (2007). Structural equation modeling. In C. R. Rao & S. Sinharay (Eds.), *Handbook of statistics* (pp. 297–358). Amsterdam: North Holland.

Zeileis, A. (2005). A unified approach to structural change tests based on ML scores, F statistics, and OLS residuals. *Econometric Reviews*, *24*(4), 445–466.

Zeileis, A., & Hornik, K. (2007). Generalized M-fluctuation tests for parameter instability. *Statistica Neerlandica*, *61*(4), 488–508.

Table 1

Group-specific population parameter values for the dynamic panel models in discrete and continuous time.

Time	θ	Group 1	Group 2	θ	Group 1	Group 2
	β_{xx}	0.700	0.450	ϕ_{yx}	0.300	1.000
	β_{yx}	0.000	0.300	ϕ_{yy}	1.000	2.000
Discrete	β_{xy}	0.000	0.300	ψ_{xx}	0.510	1.145
	β_{yy}	0.700	0.450	ψ_{yx}	0.153	0.168
	ϕ_{xx}	1.000	2.000	ψ_{yy}	0.510	1.145
	a_{xx}	-0.357	-1.092	ϕ_{yx}	0.300	1.000
	a_{yx}	0.000	0.805	ϕ_{yy}	1.000	2.000
Continuous	a_{xy}	0.000	0.805	q_{xx}	0.713	2.760
	a_{yy}	-0.357	-1.092	q_{yx}	0.214	-1.034
	ϕ_{xx}	1.000	2.000	q_{yy}	0.713	2.760

Table 2

Population parameter values for the dynamic panel models used in Simulation II. Note that the cross-lagged effects β_{yx} *and* a_{yx} *differ between the two groups.*

Discrete time				Continuous time			
θ	Value	θ	Value	θ	Value	θ	Value
β_{xx}	0.500	ϕ_{yx}	0.300	a_{xx}	−0.780	ϕ_{yx}	0.300
β_{yx}	0.200/0.300	ϕ_{yy}	1.000	a_{yx}	0.424/0.546	ϕ_{yy}	0.300
β_{xy}	0.200	ψ_{xx}	0.650	a_{xy}	0.424	q_{xx}	1.306
β_{yy}	0.500	ψ_{yx}	0.013	a_{yy}	−0.780	q_{yx}	−0.379
ϕ_{xx}	1.000	ψ_{yy}	0.650	ϕ_{xx}	1.000	q_{yy}	1.306

Table 3

Proportions of type I errors for the parameters estimates of the discrete-time dynamic panel model.

	Vanilla IPC			Iterated IPC		
Group size	β	ϕ	ψ	β	ϕ	ψ
100	5.483	5.047	5.353	6.097	5.047	5.437
125	5.167	5.163	5.133	5.663	5.163	5.277
150	5.037	5.143	5.010	5.500	5.143	5.207
175	5.173	4.940	5.150	5.600	4.940	5.090
200	5.093	4.900	4.850	5.443	4.900	4.983

Table 4

Proportions of type I errors for the parameters estimates of the continuous-time dynamic panel model.

	Vanilla IPC			Iterated IPC		
Group size	a	ϕ	q	a	ϕ	q
100	5.333	5.083	5.353	5.250	5.083	4.477
125	5.103	5.007	5.263	5.183	5.007	4.620
150	5.207	5.090	5.143	5.247	5.090	4.657
175	5.077	5.193	4.993	5.157	5.193	4.580
200	4.810	5.207	4.803	4.920	5.207	4.463

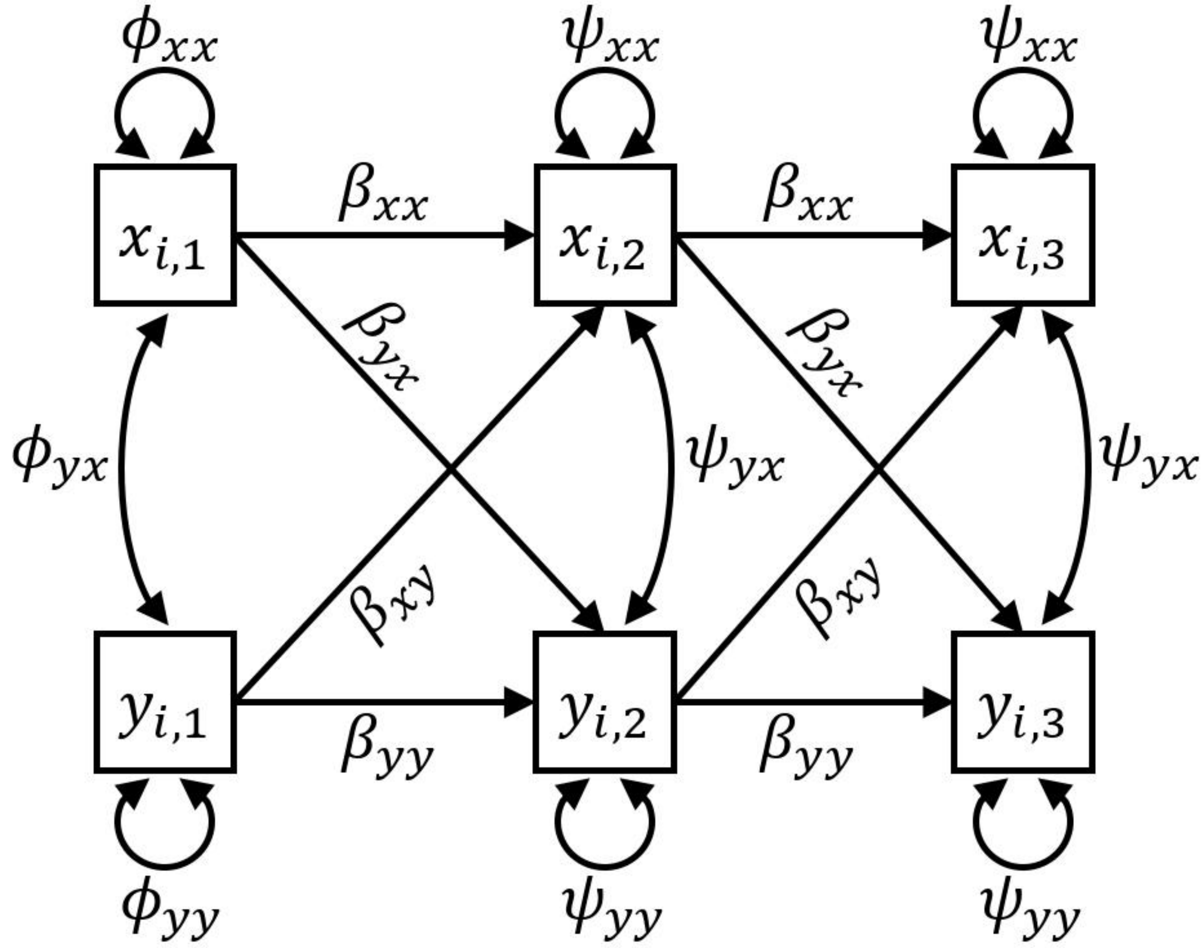

Figure 1. Path diagram of a bivariate autoregressive and cross-lagged panel model for three waves of data.

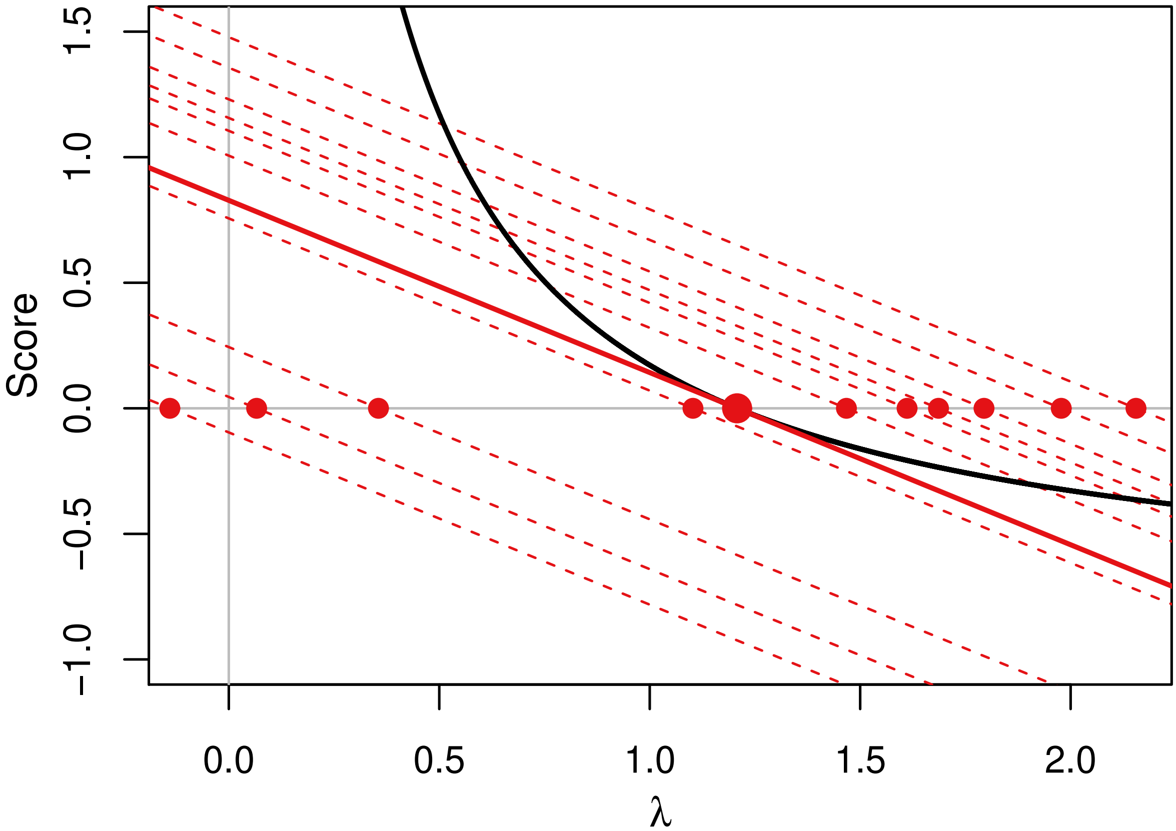

Figure 2. IPCs to the rate parameter λ of the exponential distribution. 10 observations were sampled from an exponential distribution with $\lambda = 1$. The solid black line indicates the mean of the 10 individual score values for different values of λ. The solid red line approximates the mean scores at the maximum likelihood estimate of the rate parameter (marked as a large red dot). Disaggregating the tangent of the mean scores yields 10 individual tangents depicted as dashed red lines. The x-intercepts of these individuals tangents (highlighted with small red dots) correspond to the IPCs to the rate parameter estimate as derived in Equation 16.

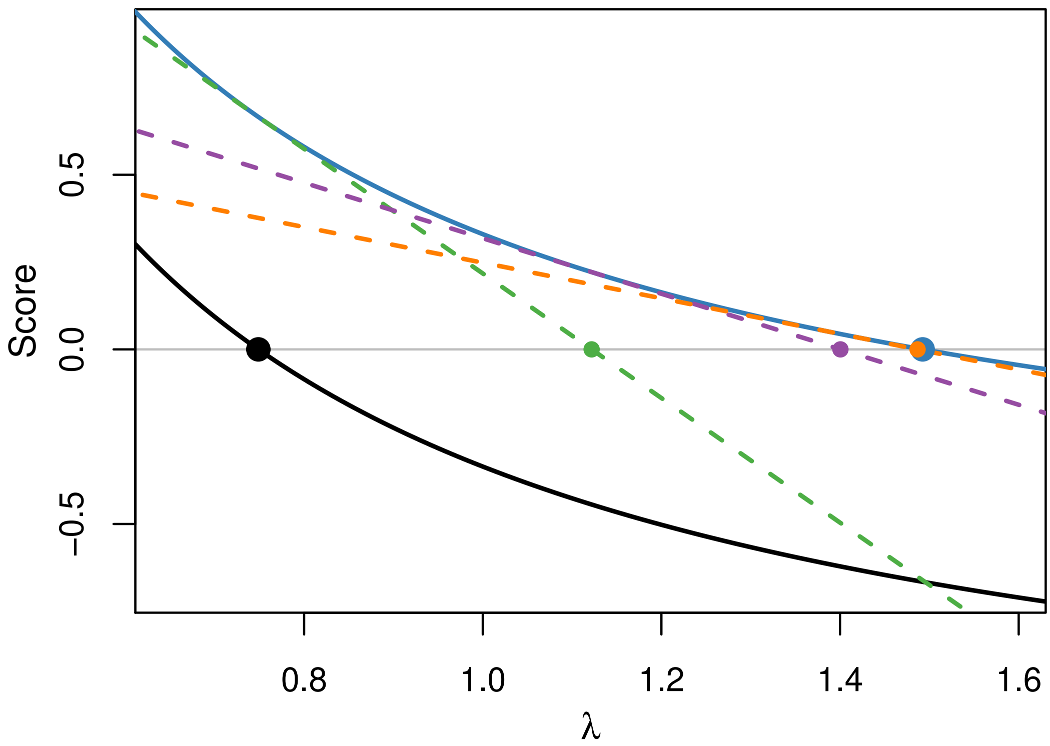

Figure 3. Bias correction procedure. 200,000 individuals were sampled in equal shares from two exponential distributions with group-specific rate parameters $\lambda_{g1} = 0.5$ and $\lambda_{g2} = 1.5$. The solid black line represents the mean scores of the pooled sample for different values of the rate parameter λ. The pooled maximum likelihood estimate is found at the x-intercept of the black line, marked as a large black dot. The solid blue line indicates the mean scores of the second group with $\lambda_{g2} = 1.5$. The group-specific maximum likelihood estimate is depicted as a large blue point. The dashed lines approximate the group-specific mean scores and their x-intercepts (marked with smaller dots) correspond to the group-specific mean IPCs. The dashed green line uses the pooled maximum likelihood estimate to approximate the group-specific mean scores. The orange dashed line uses the group-specific estimate provided by the green line instead of the pooled maximum likelihood estimate and the purple dashed line uses the estimate provided by the orange line. Note that the bias decreases with each iteration. After two more iterations, the bias correction procedure converges to the group-specific maximum likelihood estimate of λ_{g2}.

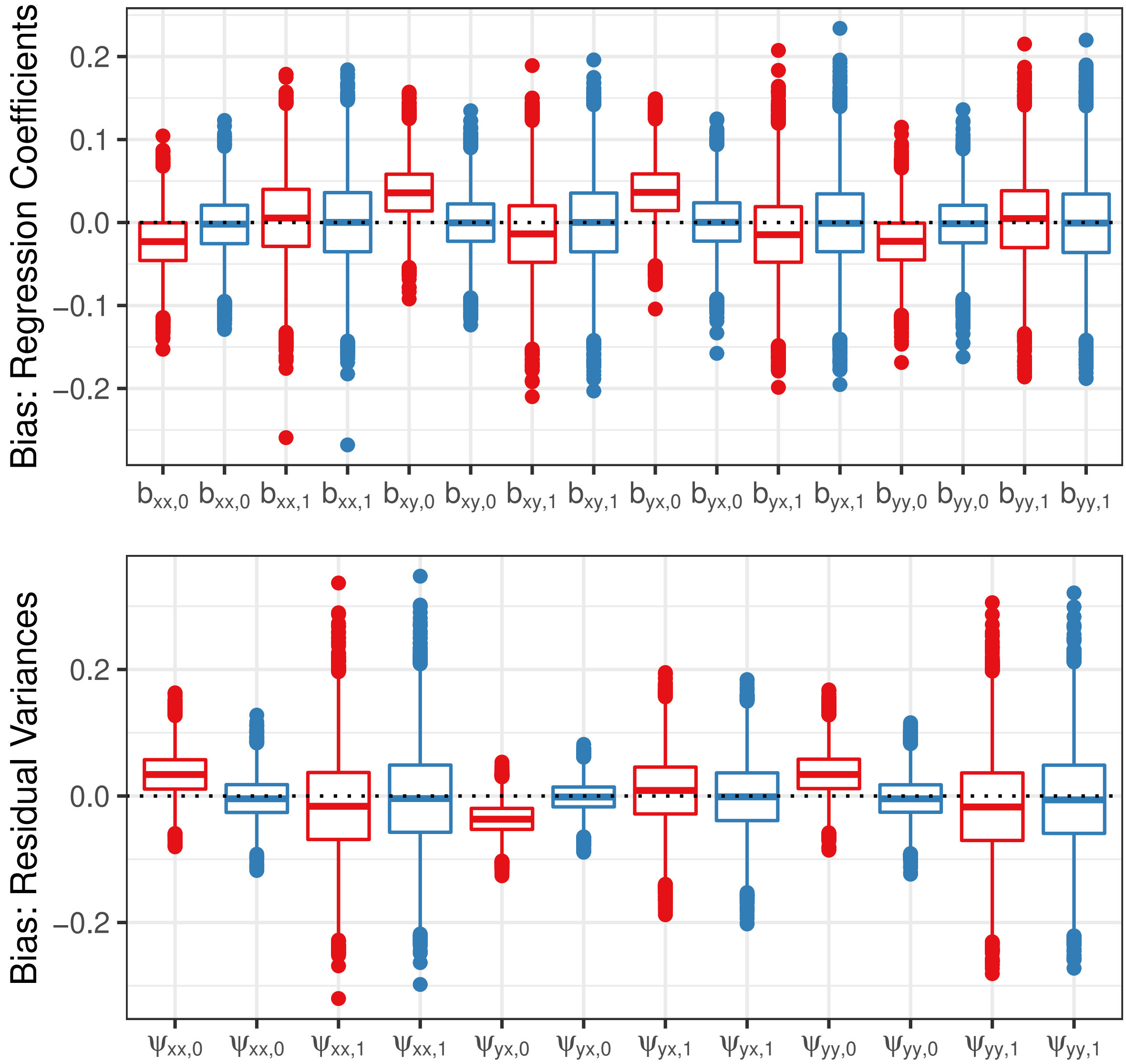

Figure 4. Boxplots of the bias of the IPC regression estimates for the discrete-time dynamic panel model.

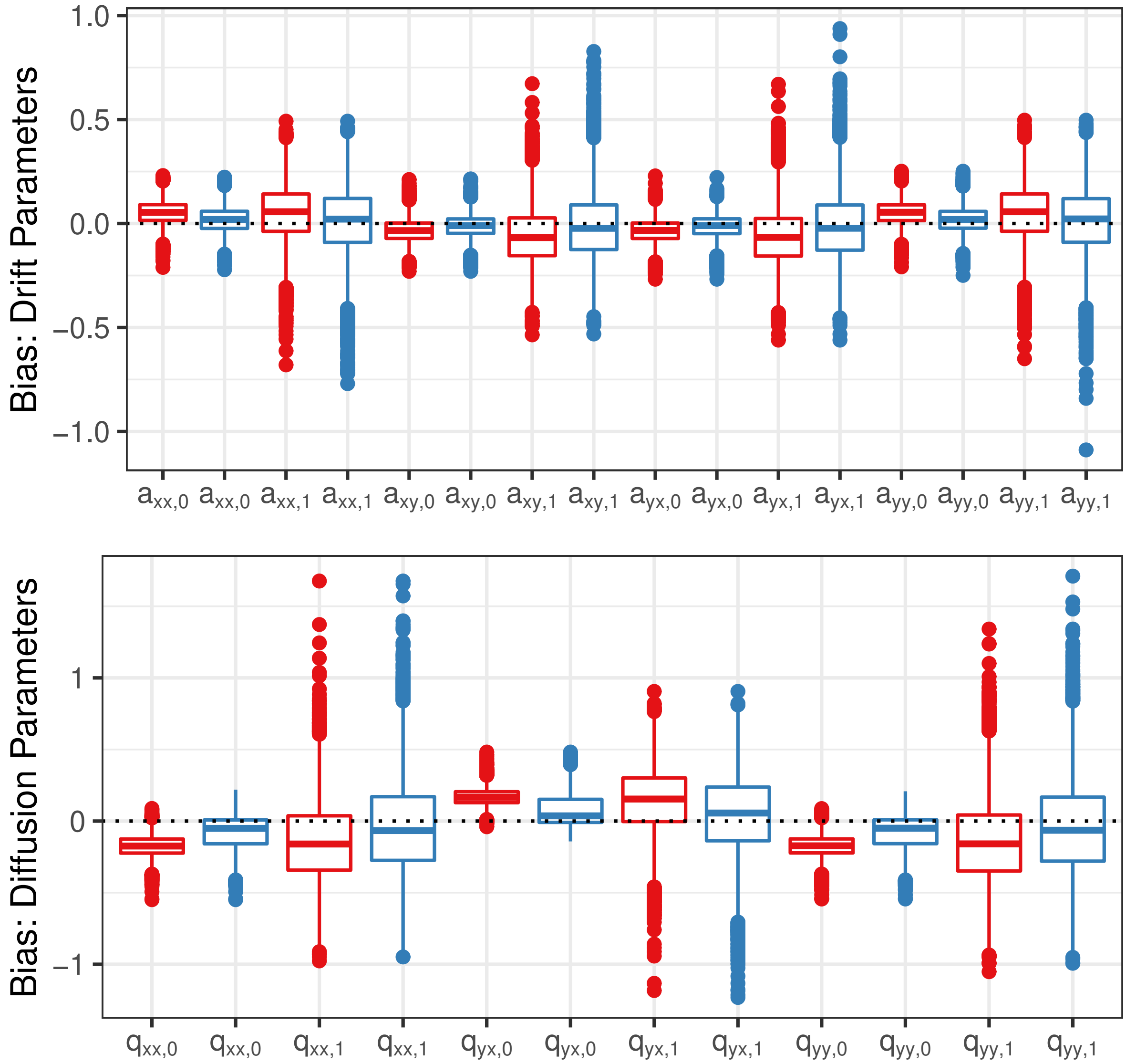

Figure 5. Boxplots of the bias of the IPC regression estimates for the continuous-time dynamic panel model.

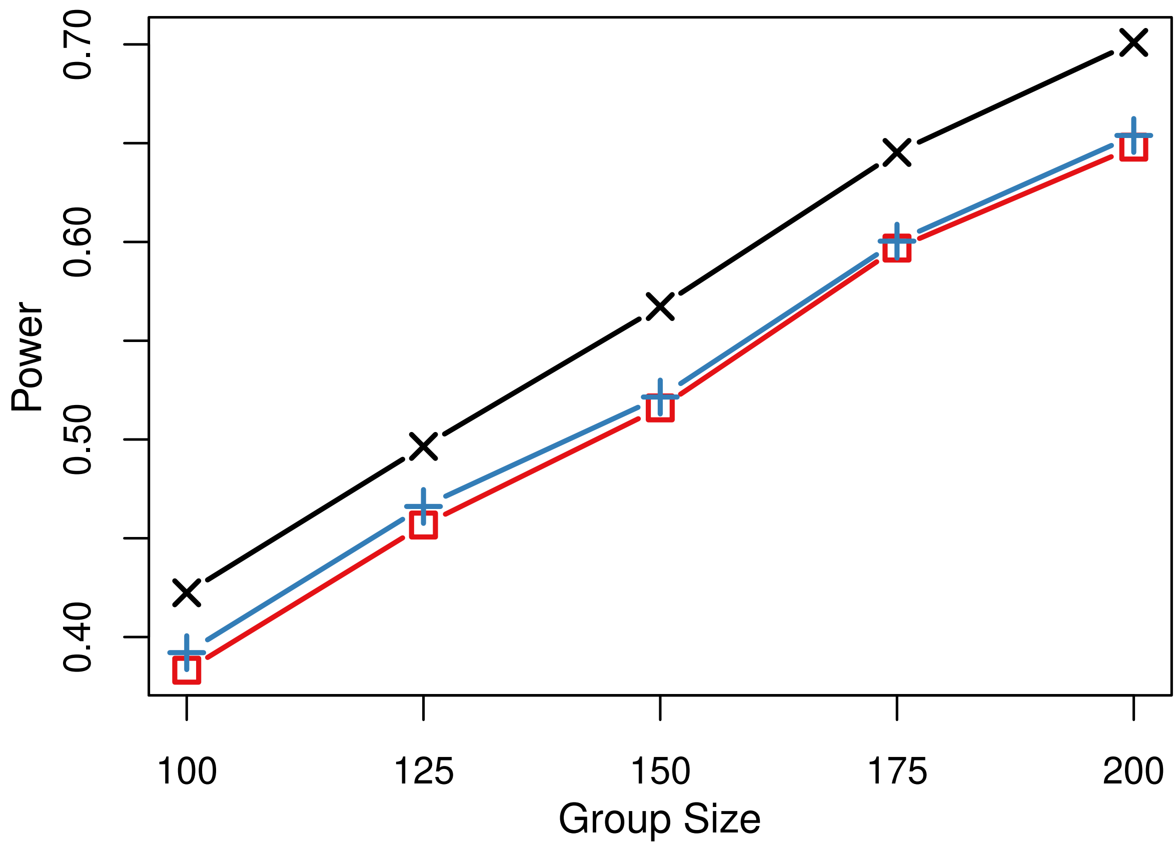

Figure 6. Power to detect that the population group difference in the cross-lagged effect β_{yx} of the discrete-time model is different from zero.

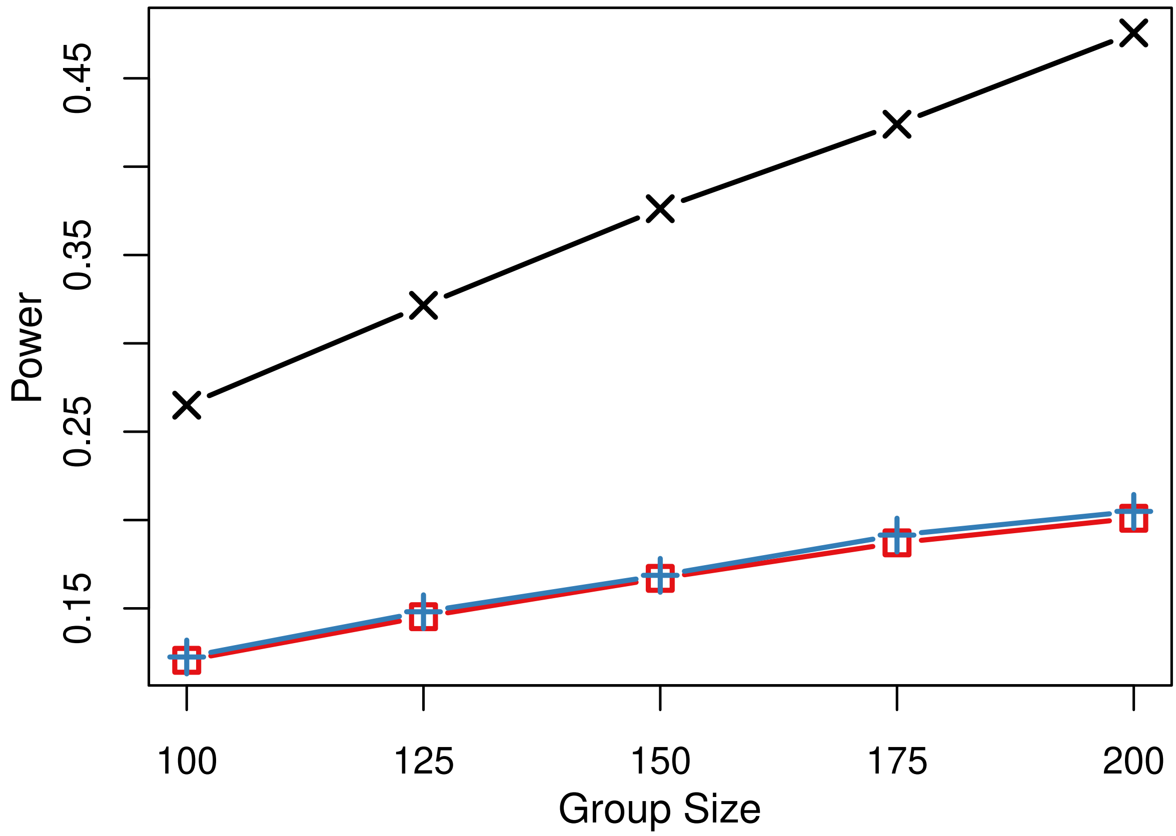

Figure 7. Power to detect that the population group difference in the drift parameter a_{yx} of the continuous-time model is different from zero.