Factor-augmented Bayesian treatment effects models for panel outcomes

Helga Wagner\textsuperscript{a}, Sylvia Frühwirth-Schnatter\textsuperscript{b}, Liana Jacobi\textsuperscript{c}

\textsuperscript{a}Department of Applied Statistics, Johannes Kepler University Linz, Altenberger Straße 69, 4040 Linz, Austria
\textsuperscript{b}Department of Finance, Accounting and Statistics, Vienna University of Economics and Business, Gebüude D4, 4. Stock, Welthandelsplatz 1, 1020 Vienna, Austria
\textsuperscript{c}Department of Economics, FBE Building, Level 4, 111 Barry Street, The University of Melbourne, VIC 3010, Australia

\textbf{A R T I C L E I N F O}

Article history:
Received 25 February 2021
Revised 15 April 2022
Accepted 16 April 2022
Available online xxx

Keywords:
Endogeneity
Factor-augmented model
Switching regression model
Shared factor model
Dynamic treatment effects

\textbf{A B S T R A C T}

A new, flexible model for inference of the effect of a binary treatment on a continuous outcome observed over subsequent time periods is proposed. The model allows to separate the associations due to endogeneity under treatment selection and additional longitudinal association of the outcomes, thus yielding unbiased estimates of dynamic treatment effects if both sources of association are present. The performance of the proposed method is investigated on simulated data and employed to re-analyze data on the longitudinal effects of a long maternity leave on mothers’ earnings after their return to the labour market.

© 2022 The Authors. Published by Elsevier B.V. on behalf of EcoSta Econometrics and Statistics.

This is an open access article under the CC BY-NC-ND license
(http://creativecommons.org/licenses/by-nc-nd/4.0/)

\section{1. Introduction}

Identification and estimation of treatment effects is an important issue in many fields to evaluate, for instance, the effectiveness of social programs, government policies or medical interventions. As each subject is observed only either under control or treatment conditions, individual outcome differences which would allow straightforward estimation of treatment effects are not available. Additionally, for data from observational studies, endogeneity of treatment selection can cause unobserved confounding and bias of treatment effects estimates if not adequately accounted for.

Bayesian approaches to inference on treatment effects rely on specifying a joint model for treatment selection and the two potential outcome sequences under either control or treatment conditions, of which only one is observed for each subject. To estimate the dynamic effect of a binary treatment on a continuous outcome observed over subsequent time periods, two models have been suggested so far, namely the switching regression model (Chib and Jacobi, 2007) and the shared factor model (Carneiro et al., 2003; Jacobi et al. (2016) build on both approaches to analyse the effects of a longer maternity leave on the subsequent earnings of mothers.

Both approaches combine a binary regression model for selection into treatment with two multivariate regression models for the outcome sequences under, respectively, control and treatment conditions. They differ, however, with respect to modeling the dependence across these regression models. Carneiro et al. (2003) model the association between treatment

\textsuperscript{E-mail addresses:} Helga.Wagner@jku.at (H. Wagner), Sylvia.Fruehwirth-Schnatter@wu.ac.at (S. Frühwirth-Schnatter), ljacobi@unimelb.edu.au (L. Jacobi).
selection and both potential outcome sequences via shared latent factors. Chib and Jacobi (2007) only specify marginal models for selection into treatment and, respectively, one sequence of potential outcomes, leaving the joint distribution of the two potential outcome sequences unspecified. Jacobi et al. (2016) show that both frameworks impose restrictions on the joint correlation structure of treatment selection and the two outcome sequences that can result in biased treatment effects estimates if the assumptions on the correlation structure of the model used for data analysis are violated in the data generating process.

To increase flexibility in modelling the dependence structure of treatment selection and potential outcomes this paper proposes a factor-augmented treatment effects model which extends the factor structure of the joint distribution to a bi-factor model. The bi-factor model was introduced in Holzinger and Swineford (1937) and recently gained popularity in item response analyses, see for example Reise (2012). Its basic assumption is that the covariance structure of multiple responses can be modelled by orthogonal factors where one common (or general) factor is shared by all responses and one or more further group (or specific) factors model the additional correlation among clusters of responses. This is appealing for the joint modeling of treatment selection and the two potential outcome sequences as it allows separate modelling of two possible sources of association, namely endogeneity of treatment selection and longitudinal correlation of the outcome sequences. The general factor shared by the binary selection and both potential outcome sequences accounts for unobserved confounding whereas outcome specific factors capture longitudinal correlation that cannot be attributed to the unobserved confounders.

The paper is structured as follows. Section 2 discusses Bayesian treatment effects models for panel outcomes and reviews the switching regression and the shared factor model. Section 3 introduces the factor-augmented treatment effects model, discusses identification issues and describes posterior inference using Markov chain Monte Carlo (MCMC) methods. In Section 4, the flexibility of the factor-augmented treatment effects model is illustrated on simulated data. Section 5 provides a reanalysis of longitudinal effects of a long maternity leave on mothers' earnings after their return to the labour market and Section 6 concludes.

2. Bayesian modelling of panel treatment effects

Assessing the effect of a treatment on an outcome sequence of interest requires a comparison of these outcomes under two conditions, namely with and without treatment (i.e. under control conditions). However, the outcome sequence is observed for each individual subject only under either treatment or control conditions. Given outcome sequences from a collection of subjects, where some underwent treatment and others did not, estimation of treatment effects may rely on the potential outcomes framework (Rubin, 1981). This framework allows to define treatment effects based on separate models for the outcomes under, respectively, treatment and control conditions.

However, inference on treatment effects from such observational data is demanding. In addition to the fundamental problem that only one potential outcome is observed for each subject, treatment is typically not randomized, but self selected. Hence, treatment might be associated with the outcome implying the presence of observed and, even more importantly, unobserved confounders (endogeneity).

To take such endogeneity of treatment selection into account, Bayesian approaches to modelling treatment effects rely on specifying a joint model for treatment selection and the potential outcomes, often in the spirit of Roy’s switching regression model (Roy, 1951; Lee, 1978). For longitudinally observed outcomes, two approaches have been suggested sofar. Chib and Jacobi (2007) specify two models for selection into treatment and one potential outcome sequence respectively, whereas Carneiro et al. (2003) specify a joint model for selection into treatment and the two potential outcomes models. Both the switching regression model and the shared factor model specify a probit model for treatment selection and two multivariate normal regression models for the potential outcome sequences, but differ with respect to modelling their joint distributions. To discuss these differences in more detail, we introduce the marginal models for treatment selection and the potential outcome sequences in Section 2.1 and describe modelling of the dependence structure in both approaches in Section 2.2.

2.1. Marginal models for treatment selection and outcome sequences

Let \(x_i, i = 1, \ldots, n\), denote the treatment status selected by subject \(i\) at time \(t = 0\), with \(x_i = 0\) and \(x_i = 1\) referring, respectively, to control and treatment. Further, we will denote the vectors of potential outcomes by \(y_{ij} = (y_{i1j}, \ldots, y_{i1T})'\) for \(j = 0, 1\) and the vector of observed outcomes for a subject \(i\) with treatment status \(x_i = j\) by \(y_i'(x_i = j) = (y_{i1}, \ldots, y_{iT})'\). In randomized studies, treatment selection \(x_i\) is independent from the observed outcomes. Hence, given the treatment status \(x_i = j\), the observed outcomes \(y_i'(x_i = j)\) have the same distribution as the potential outcomes \(y_{ij}\) which allows a straightforward estimation of the average treatment effect from the observed outcomes. However, this independence is lost in observational studies where subjects choose a treatment based on their expectations regarding the outcomes (endogenous treatment). For such data, joint modelling of the treatment selection and outcomes, i.e. modelling the association between treatment selection and observed outcomes, is essential and an exclusion restriction (instrument) is required to identify the causal effect of the endogenous treatment within a potential outcomes framework.
We start with the specification of the treatment selection and the potential outcome models. Treatment selection depends on covariates (selection on observables) via a probit model for choosing treatment $x_{it}$, which can be specified in terms of a latent Gaussian random variable $x_{i}^*$ as

$$x_{i}^* = \mathbf{v}_i \alpha + \epsilon_{x_{i}}, \quad \epsilon_{x_{i}} \sim \mathcal{N}(0, \sigma_{x}^2).$$

(1)

$$x_{i} = I(x_{i}^* > 0),$$

(2)

where $\mathbf{v}_i$ denotes the $1 \times d_{a}$ vector of baseline covariates that control selection into treatment and the $d_{a} \times 1$ vector $\alpha$ captures their effect on treatment selection. $x_{i}^*$ can be interpreted as subject $i$'s latent utility of selection into treatment.

As opposed to the usual specification of a probit model, (1) leaves the variance $\sigma_{x}^2$ of the error term $\epsilon_{x_{i}}$ unspecified, whereas usually $\sigma_{x}^2 = 1$ is fixed, as the regression effects $\alpha$ are only identified up to a scale factor. However, in a factor model where the error term $\epsilon_{x_{i}}$ is modelled by a latent factor plus an idiosyncratic error it is more convenient to fix the variance of the idiosyncratic error to one, as will be explained below.

The selection model given in Equations (1) and (2) is combined with a model for the potential outcomes for subject $i$ at time points $t = 1, \ldots, T$ which are denoted by $y_{0_{it}}$ and $y_{1_{it}}$ for the outcomes under control and treatment conditions, respectively. The potential outcomes are modelled as

$$y_{0_{it}} = \eta_{0_{it}}(\mathbf{w}_{it}) + \epsilon_{0_{it}}, \quad \epsilon_{0_{it}} \sim \mathcal{N}(0, \sigma_{0}^2).$$

(3)

$$y_{1_{it}} = \eta_{1_{it}}(\mathbf{w}_{it}) + \epsilon_{1_{it}}, \quad \epsilon_{1_{it}} \sim \mathcal{N}(0, \sigma_{1}^2).$$

(4)

with structural means $\eta_{j_{it}}(\mathbf{w}_{it})$ for $j = 0, 1$ depending on a $1 \times d$ vector of covariates $\mathbf{w}_{it}$:

$$\eta_{0_{it}}(\mathbf{w}_{it}) = \mu_{t} + \mathbf{w}_{it}^\top \gamma,$n

(5)

$$\eta_{1_{it}}(\mathbf{w}_{it}) = (\mu_{t} + \kappa_{t}) + \mathbf{w}_{it}^\top (\gamma + \theta).$$

(6)

$\mathbf{w}_{it}$ contains covariates that are not directly affected by the treatment. These include a subset of the baseline covariates from the selection equation, $\mathbf{v}_i$, except for at least one exclusion restriction (such as a policy instrument) that only affects treatment selection. The latter is key to identifying the correlation in unobservables between selection and treatment modelled in the next section, and hence to obtain unbiased parameter and treatment effects estimates.

Parameters of interest are the intercepts $\mu_{t}$ and $\mu_{t} + \kappa_{t}$ and the $d \times 1$ vectors of covariate effects $\gamma$ and $\gamma + \theta$ under control and under treatment conditions, respectively. Assuming that we control for baseline characteristics in both the treatment and selection equations as in our application in Section 5, this specification captures the dynamic effects of the treatment at $t = 0$ on the outcomes at times $t = 1, \ldots, T$ via the time-varying intercept $\kappa_{t}$ as well as through additional treatment effects in $\theta$. Hence, the average dynamic treatment effect $\text{ATE}_t$ of a subject with covariate values $\mathbf{w}_{it}$ in panel period $t$ takes the form

$$\text{ATE}_t = E(y_{1_{it}} - y_{0_{it}}|\mathbf{w}_{it}) = \kappa_{t} + \mathbf{w}_{it}^\top \theta.$$

(2.2. Modelling the dependence structure)

As noted above, endogeneity of treatment selection (selection on unobservables) can be taken into account by allowing for correlation of treatment selection and the potential outcomes. This approach is followed by both the switching regression as well as the shared factor model. However, the impossibility to observe both outcomes $y_{0_{it}}$ and $y_{1_{it}}$ for the same subject $i$ makes it impossible to observe the joint distribution of the errors $\epsilon_{x_{i}}$ and $\epsilon_{j_{it}}$, where $\epsilon_{j_{it}} = (\epsilon_{j_{i1}}, \ldots, \epsilon_{j_{iT}})^\top$ in (3) and (4) and the two models differ with respect to specifying the error distribution.

In the switching regression model, the joint distribution of $(\epsilon_{0_{it}}, \epsilon_{1_{it}})$ is left unspecified and only the two joint $(T + 1)$-variate distributions of the latent utility error $\epsilon_{x_{i}}$ and the errors $\epsilon_{j_{it}}$ in each outcome equation, i.e. the marginal distributions of $(\epsilon_{x_{i}}, \epsilon_{0_{it}})$ and $(\epsilon_{x_{i}}, \epsilon_{1_{it}})$, are specified as multivariate normal distributions.

In contrast, in the shared factor model the joint $(2T + 1)$-dimensional distribution of all error terms $(\epsilon_{x_{i}}, \epsilon_{0_{it}}, \epsilon_{1_{it}})$ is modelled in terms of latent factors and independent idiosyncratic errors. Carneiro et al. (2003) specify a multi-factor model and exploit additional measurements from psychological tests to identify factors and factor loadings. Such additional measurements are not required in the analysis of Jacobi et al. (2016) and are available in many applications. Jacobi et al. (2016) propose a simpler factor structure with a single subject specific latent factor that accounts for both sources of dependence, namely endogeneity and within subject dependence. The error terms are specified as

$$\epsilon_{x_{i}} = \lambda_{x} f_{i} + \epsilon_{x_{i}}, \quad \epsilon_{x_{i}} \sim \mathcal{N}(0, 1),$$

(7)

$$\epsilon_{j_{it}} = \lambda_{j} f_{i} + \epsilon_{j_{it}}, \quad \epsilon_{j_{it}} \sim \mathcal{N}(0, S_{j}) \quad S_{j} = \text{Diag}(\sigma_{j_{1}}^2, \ldots, \sigma_{j_{T}}^2) \quad j = 0, 1,$$

(8)

where $f_{i} \sim \mathcal{N}(0, 1)$ is an unobserved subject specific factor, $\lambda_{x}$ is its loading on the latent utility and $\lambda_{j} = (\lambda_{j_{1}}, \ldots, \lambda_{j_{T}})^\top$, $j = 0, 1$ denote the $T \times 1$ vectors of factor loadings for the potential outcomes. $\epsilon_{x_{i}}$ and $\epsilon_{j_{it}}$, $j = 0, 1$ are the idiosyncratic errors.
of the latent utility and the potential outcome vectors, respectively. Both the factor loadings \( \lambda_{ji} \) as well as the variances \( \sigma^2_{ji} \) of the idiosyncratic errors are allowed to vary over time. The joint covariance matrix of the vector \( \varepsilon_i = (\varepsilon_{xi}, \varepsilon'_{0i}, \varepsilon'_{1i})' \) is then given as

\[
\text{Cov}(\varepsilon_i) = \Sigma = \begin{pmatrix}
\sigma^2_x & \sigma'_{x0} & \sigma'_{x1} \\
\sigma_{x0} & \Sigma_0 & \Sigma_{01} \\
\sigma_{x1} & \Sigma_{01} & \Sigma_1
\end{pmatrix} = \begin{pmatrix}
1 + \lambda^2_x & \lambda_x\lambda_0 & \lambda_x\lambda_1 \\
\lambda_x\lambda_0 & \lambda_0^2 + \Sigma_0 & \lambda_0\lambda_1 \\
\lambda_x\lambda_1 & \lambda_0\lambda_1 & \lambda_1^2 + \Sigma_1
\end{pmatrix}.
\]

Hence, for fixed covariates, \( \sigma_{xj} = \text{Cov}(x_j, y_{ji}) \) denotes the \( T \times 1 \) vector of covariances between the latent utility \( x_j \) and the potential outcome vector \( y_{ji} \). \( \Sigma_j = \text{Cov}(y_{ji}) \) is the \( T \times T \) covariance matrix of the potential outcome vector \( y_{ji} \) for \( j = 0, 1 \) and \( \Sigma_{01} = \text{Cov}(y_{0i}, y_{1i}) \) is the \( T \times T \) covariance matrix of the two potential outcome vectors \( y_{0i} \) and \( y_{1i} \).

The assumption that the latent factor \( f_j \) is shared by the latent utility and all potential outcomes implies that the vectors of time-varying factor loadings \( \lambda_{ji}, j = 0, 1 \) determine not only \( \Sigma_j \) (and thus the correlation within each potential outcome vector) and \( \sigma_{xj} = \lambda_x\lambda_j \) (and thus the correlation between latent utility and each potential outcome vector), but also \( \Sigma_{01} \) and thus the correlation between the potential outcome vectors. Note, however, that this assumption is not testable from empirical data where only one potential outcome is observed for each subject.

The switching regression model avoids the specification of the joint distribution of the potential outcome vectors and assumes outcome specific latent factors \( f_{ji} \) and factor loadings \( \xi_j \) for \( j = 0, 1 \) to model the potential outcomes vector by specifying the error terms \( \varepsilon_{ji} \) as

\[
\varepsilon_{ji} = \xi_j f_{ji} + \varepsilon_{ji} \sim \mathcal{N}(0, \Sigma_j), \quad j = 0, 1,
\]

where (as in the shared factor model) \( \Sigma_j = \text{Diag}(\sigma^2_{ji}, \ldots, \sigma^2_{ji}) \) is the covariance matrix of the idiosyncratic errors. The outcome specific latent factors \( f_{ji} \) and \( \xi_j \) are assumed to arise marginally from standard normal distributions, \( f_{ji} \sim \mathcal{N}(0,1) \), but no assumption is made on their joint distribution. While the error term of the latent utility, \( \varepsilon_{xi} = \varepsilon_{xi} \), is independent of the factors, it is allowed to be correlated with the idiosyncratic errors \( \varepsilon_{ji} \) of each potential outcome equation, to capture selection on unobservables. Hence, the joint \( (T+1) \)-variate distributions of the idiosyncratic errors \( (\varepsilon_{xi}, \varepsilon_{ji})' \) are given as

\[
(\varepsilon_{xi}, \varepsilon_{ji})' \sim \mathcal{N}(0, \Omega_j), \quad \Omega_j = \begin{pmatrix}
1 & \omega_j' \\
\omega_j & \Sigma_j
\end{pmatrix}, \quad j = 0, 1.
\]

where the vector \( \omega_j \) of covariances between the latent utility and the potential outcome vector \( y_{ji} \) is unstructured, but restricted to guarantee positive definiteness of the joint covariance matrix \( \Omega_j \).

Both models have limitations. The shared factor model relies on the untenable assumption that the latent factor is shared by the latent utility and both potential outcomes. Hence, as the factor loadings determine all correlations in the multivariate normal distribution of the errors \( \varepsilon_{ji} \), this correlation structure is not fully flexible. On the other hand, in the switching regression model no joint model for the error terms is specified and therefore the variance of the outcome difference is not available. Additionally, the assumption that each latent factor affects only the corresponding potential outcomes vector but not the latent utility, implies that the covariance matrix of \( \varepsilon_{ji} \) conditional on latent utility error \( \varepsilon_{xi} \) is given as

\[
\text{Cov}(\varepsilon_{ji}|\varepsilon_{xi}) = \Sigma_j - \omega_j\omega'_j.
\]

Thus, conditional on \( \varepsilon_{xi} \), the idiosyncratic errors \( \varepsilon_{ji} \) in the potential outcomes model are correlated (unless \( \omega_j = 0 \)), which is an unusual assumption for idiosyncratic errors. Moreover, to guarantee positive definiteness of the joint covariance matrices of the idiosyncratic errors \( \Omega_j \), the correlations between the latent utility and the potential outcomes are restricted, see Chib and Jacobi (2007), and cannot capture severe endogeneity.

None of the models encompasses the other, but the positive semi-definiteness of the specified covariance matrices can result in restrictions on their elements in one model which cannot be recovered under the other. Thus, as illustrated for simulated data in Jacobi et al. (2016), treatment effects can be biased, when data generated under the shared factor model are analysed using the switching regression model and vice versa. Avoiding such a bias is our main motivation for introducing the factor-augmented treatment effects model in Section 3. In a simulation study in Section 4, we analyse data generated from the factor-augmented treatment effects model with the shared factor model and the switching regression model to discuss restrictions and implications of these two models in more detail.

### 3. Factor-augmented treatment effects model

In this section we specify a factor-augmented model for the joint distribution of the latent utility and the two potential outcome sequences. The model allows for a flexible dependence structure where the correlation within an outcome sequence is separated into correlation due to confounding and additional longitudinal correlation. For estimation, we will utilize a Bayesian approach. We introduce the model in Section 3.1 and discuss identification issues in Section 3.2. Section 3.3 describes specification of the prior distributions and Section 3.4 outlines posterior inference based on Markov chain Monte Carlo (MCMC).
3.1. Model specification

We consider the model specified in Section 2.1, with the probit model for treatment selection given in Equations (1) and (2) and the model for the potential outcome sequences \( \mathbf{y}_{0i} \) and \( \mathbf{y}_{1i} \) given as

\[
\mathbf{y}_{0i} = \mathbf{\mu} + \mathbf{W}_i \mathbf{y} + \mathbf{\epsilon}_{0i},
\]
\[
\mathbf{y}_{1i} = \mathbf{\mu} + \kappa + \mathbf{W}_i (\mathbf{y} + \mathbf{\theta}) + \mathbf{\epsilon}_{1i},
\]

where \( \mathbf{W}_i \) is the \( T \times d \) matrix of covariate values with rows \( \mathbf{w}_{i1}, \ldots, \mathbf{w}_{iT} \) and \( \mathbf{\mu} \) and \( \kappa \) are the \( T \times 1 \) vectors \( \mathbf{\mu} = (\mu_1, \ldots, \mu_T)' \) and \( \kappa = (\kappa_1, \ldots, \kappa_T)' \). We specify the time-varying intercepts and treatment effects as \( \mu_t = \mu \) and \( \kappa_t = \kappa \) for \( t = 1 \) and as \( \mu_t = \mu_t + v_{t-1} \) and \( \kappa_t = \kappa_t + \tau_{t-1} \) for \( t = 2, \ldots, T \) in terms of unknown parameters \((\mu, v_1, \ldots, v_{T-1})\) and \((\kappa, \tau_1, \ldots, \tau_{T-1})\). This specification is useful as Bayesian variable selection methods can be applied to learn whether the treatment effect \( \kappa \) is different from zero, and also whether \( v_{t-1} \) and \( \tau_{t-1} \) are different from zero implying that \( \mu_t \) and \( \kappa_t \), respectively, are indeed dynamic.

To achieve more flexibility in modelling the association of treatment selection and the potential outcome sequences, we assume that all dependencies in the \( (2T + 1) \times 1 \) error vector \( \mathbf{\epsilon}_i = (\epsilon_{xi}, \epsilon_{0i}, \epsilon_{1i})' \) are captured by three subject specific latent factors: one common factor \( f_{ci} \) which is shared by the error terms of the latent utility \( x_i \) and both potential outcome vectors \( \mathbf{y}_{0i} \) and \( \mathbf{y}_{1i} \) and two specific factors \( f_{0i} \) and \( f_{1i} \) which affect only the error vectors of the potential outcomes \( \mathbf{\epsilon}_{0i} \) and \( \mathbf{\epsilon}_{1i} \). respectively.

The common factor thus accounts for unobserved confounding whereas the two outcome specific factors \( f_{0i} \) and \( f_{1i} \) capture the additional longitudinal association in the outcome vectors that cannot be attributed to unobserved confounders. The joint model for the error terms is thus specified as

\[
\epsilon_{xi} = \lambda_x f_{ci} + \epsilon_{xi}, \quad \epsilon_{xi} \sim \mathcal{N}(0, 1),
\]
\[
\epsilon_{0i} = \lambda_0 f_{ci} + \zeta_0 f_{0i} + \epsilon_{0i}, \quad \epsilon_{0i} \sim \mathcal{N}(0, \sigma_{0i}^2),
\]
\[
\epsilon_{1i} = \lambda_1 f_{ci} + \zeta_1 f_{1i} + \epsilon_{1i}, \quad \epsilon_{1i} \sim \mathcal{N}(0, \sigma_{1i}^2),
\]

where the factors \( f_{ci}, f_{0i} \) and \( f_{1i} \) are assumed to be independent standard normals. Hence, the factor loadings \( \lambda_x, \lambda_j \) and \( \zeta_j, j = 0, 1 \) determine the joint covariance matrix of all error terms.

Assumptions on how factors are related to outcomes simplify the structure of the factor loadings matrix. First, treatment selection and both outcome panels depend on the common factor \( f_{ci} \) with loadings \( \lambda_x \) for the latent utility and \( \lambda_j, j = 0, 1 \) for the two potential outcomes. Second, the potential outcome vector \( y_{ji} \) depends only on one of the specific factors, \( f_{ji} \) with loadings \( \zeta_j \). Therefore, the factor model for the errors is given in matrix form as

\[
\mathbf{\epsilon}_i = \mathbf{\Lambda} \mathbf{f}_i + \mathbf{\epsilon}_i, \quad \mathbf{\Lambda} = \begin{pmatrix}
\lambda_x & 0 & 0 \\
\lambda_0 & \zeta_0 & 0_T \\
\lambda_1 & 0_T & \zeta_1
\end{pmatrix},
\]

where \( \mathbf{f}_i = (f_{ci}, f_{0i}, f_{1i})' \) is the vector of factors for subject \( i \) and \( \mathbf{\epsilon}_i = (\epsilon_{xi}, \epsilon_{0i}, \epsilon_{1i})' \) is the vector of idiosyncratic errors. Hence, the joint \((2T + 1)\)-variate distribution of \( \mathbf{\epsilon}_i \) is multivariate normal, \( \mathbf{\epsilon}_i \sim \mathcal{N}(\mathbf{0}, \mathbf{\Sigma}) \), with the covariance matrix given as

\[
\mathbf{\Sigma} = \begin{pmatrix}
\sigma_x^2 & \sigma_{x0} & \sigma_{x1} \\
\sigma_{x0} & \Sigma_0 & \Sigma_{01} \\
\sigma_{x1} & \Sigma_{01} & \Sigma_1
\end{pmatrix} = \begin{pmatrix}
1 + \lambda_x^2 & \lambda_x & \lambda_x \\
\lambda_x & \lambda_x^2 + \lambda_0^2 + \zeta_0 S_0 & \lambda_x \lambda_0 + \zeta_0 \zeta_1 \\
\lambda_x & \lambda_x \lambda_0 + \zeta_0 \zeta_1 & \lambda_x \lambda_1 + \zeta_1 S_1
\end{pmatrix}.
\]

We call a potential outcomes model, where the joint distribution of the errors \( \mathbf{\epsilon}_i = (\epsilon_{xi}, \epsilon_{0i}, \epsilon_{1i})' \) is defined by (15), a factor-augmented treatment effects model.

This model addresses the limitations of both the shared factor and the switching regression model. As correlation across panel outcomes is not attributed solely to the common factor, it is more flexible than the shared factor model which is recovered as that special case where \( \zeta_0 = \zeta_1 = 0 \). If the joint distribution of the specific factors \( f_{0i} \) and \( f_{1i} \) is left unspecified then the factor-augmented model is a switching regression model where conditioning on the latent factors, the idiosyncratic errors of latent utility and each potential outcome are independent.

In (15), the specific factors \( f_{0i} \) and \( f_{1i} \) are assumed to be univariate. Though an extension to multiple independent outcome specific factors \( f_{ji} \) is straightforward conceptually, to achieve identification of the factor loadings from the observed data the dimension of outcome specific factors is restricted by the number of available panel observations \( T \). We will return to this issue in Section 3.2 but note here that in contrast to Carneiro et al. (2003) we do not assume that additional measurements are available from which the latent factors can be identified.
3.2. Identification

An important issue in factor models is their identification, which according to Anderson and Rubin (1956) is a two-step procedure where the first step is identification of the variance contribution attributable to the latent factors, i.e. $\Lambda\Lambda'$. The second step is then identification of $A$ from $\Lambda\Lambda'$, i.e. solving the rotational identification problem.

The data structure in the factor-augmented treatment effects model proposed in Section 3.1 differs considerably from the standard factor model which assumes multivariate normal observations: only the binary treatment variable $x_i$ and the observed outcome sequence, which is a truncated version of one of the two potential outcome sequences are observed for each subject. However, for the factor-augmented treatment effects model with one common and two specific factors for $\lambda_x \neq 0$, rotational identification is unproblematic due to the generalized lower triangular structure of the factor loadings matrix, see Frühwirth-Schnatter and Lopes (2018).

In the probit model, identification of regression effects is feasible only up to the standard error $\sigma_x$ of the latent utility and hence only the standardized effects

\[
\tilde{a} = \frac{a}{\sigma_x} = \frac{a}{\sqrt{1 + \lambda_x^2}}
\]

in model (1) are identified.

Fixing $V(e_y) = \sigma_y^2 = 1$ in (1) instead of fixing $V(e_y) = 1$ in (7) would require restricting the factor loading $\lambda_x$ to the interval $(-1, 1)$, while in our specification $\lambda_x$ is unconstrained.

The observed data never provide information on the association between the two potential outcome vectors $y_{0i}$ and $y_{1i}$. Due to endogeneity selection the distribution of the observed outcome sequence $y_{ij}(x_i = j)$ is not the marginal distribution of $y_{ij}$, but the conditional distribution truncated by the respective range of the latent utility. It is given as

\[
p(y_{ij}|x_i = j) = \begin{cases} 
\frac{1}{1 - \phi(\tilde{v}_{x_i})} \int_{-\infty}^{0} p(y_{0i}|x_i = j) dx_{ij}, & j = 0, \\
\frac{1}{\Phi(\tilde{v}_{x_i})} \int_{0}^{\infty} p(y_{1i}|x_i = j) dx_{ij}, & j = 1,
\end{cases}
\]

where $\tilde{\mu}_{x_i} = \tilde{v} \tilde{a}$ is the mean of the standardized latent utility $x_i'/\sigma_x$. Identification of all parameters that are not identified from the probit model, i.e. the regression effects in the potential outcomes models, the factor loadings and the variances of the idiosyncratic errors has to be accomplished from these two conditional distributions. We will discuss necessary conditions for identification of the model parameters from the first and second moments of these two conditional distributions.

The expectation of the observed outcome sequence $y_{ij}(x_i = j)$ is given as:

\[
E(y_{ij}|x_i = 0) = E(y_{0i}|x_i = 0) = \mu_i + \mathbf{W}_i \gamma - \frac{\sigma_{0x} \phi(\tilde{v} \tilde{a})}{\sigma_x 1 - \Phi(\tilde{v} \tilde{a})},
\]

\[
E(y_{ij}|x_i = 1) = E(y_{1i}|x_i = 1) = \mu_i + \kappa_i + \mathbf{W}_i (\gamma + \theta) + \frac{\sigma_{1x} \phi(\tilde{v} \tilde{a})}{\sigma_x \Phi(\tilde{v} \tilde{a})},
\]

where $\phi(\cdot)$ and $\Phi(\cdot)$ are, respectively, the pdf and the cdf of the standard normal distribution. Details are provided in Appendix A. As the quantities $\gamma(\tilde{v} \tilde{a}) = -\frac{\phi(\tilde{v} \tilde{a})}{\Phi(\tilde{v} \tilde{a})}$ and $\kappa_i = \frac{\phi(\tilde{v} \tilde{a})}{\Phi(\tilde{v} \tilde{a})}$ are identified from the probit model, identification of the parameters $\mu_i, \kappa_i, \gamma_i, \theta$ and $\hat{\sigma}_{\tilde{v}}(\tilde{v} \tilde{a}) = \frac{\sigma_{0x} \phi(\tilde{v} \tilde{a})}{\sigma_x}$ for $j = 0, 1$ is feasible from these equations, if the design matrix in the corresponding regression model is of full rank.

$\hat{\sigma}_{\tilde{v}0}$ and $\hat{\sigma}_{\tilde{v}1}$ yield $2T$ equations for the $2T + 1$ factor loadings $\lambda_j = (\lambda_x, \lambda_0^j, \lambda_1^j)^T$ of the common factor, leaving at least one of these factor loadings unidentified.

The conditional covariance matrices $V(y_{ij}|x_i = 0)$ and $V(y_{ij}|x_i = 1)$ are derived in Appendix A. They are given as

\[
V(y_{ij}|x_i = 0) = \Sigma_{0} - c_{00}(\tilde{v} \tilde{a}) \left(\tilde{v} \tilde{a} + c_{0}(\tilde{v} \tilde{a})\right) \sigma_{\tilde{v}0} \sigma_{\tilde{v}0},
\]

\[
V(y_{ij}|x_i = 1) = \Sigma_{1} - c_{11}(\tilde{v} \tilde{a}) \left(\tilde{v} \tilde{a} + c_{1}(\tilde{v} \tilde{a})\right) \sigma_{\tilde{v}1} \sigma_{\tilde{v}1},
\]

and have $\frac{T(T-1)}{2}$ free elements each, from which one element of $\lambda$, the factor loadings $\xi_0$ and $\xi_1$ and the variances $\sigma_0^2$ and $\sigma_1^2$ have to be identified. Thus, a necessary condition for identification is that

\[
T(T + 1) \geq 4T + 1,
\]

and hence that the observed panel outcomes are at least of length $T \geq 4$.

Generally, a necessary condition for identification of the parameters in a model with $r > 1$ outcome specific factors is that

\[
T(T + 1) \geq 2(r + 1)T + 1.
\]

Therefore, the outcome panels have to be at least of length $T = 6$ and $T = 8$ to identify the loadings of $r = 2$ or $r = 3$ outcome specific factors, respectively. Sufficient conditions for the identification of the elements of $\lambda_0, \lambda_1, \xi_0$ and $\xi_1$, however, require also that enough factor loadings are different from 0, see Anderson and Rubin (1956); Conti et al. (2014);
Frühwirth-Schnatter and Lopes (2018). Based on the corresponding counting rule in Frühwirth-Schnatter and Lopes (2018), the necessary condition (16) is also sufficient, if we leave all unknown factors loadings in (15) unrestricted.

3.3. Prior distributions

In the present paper, econometric inference relies on a Bayesian approach. To complete the Bayesian model specification, prior distributions are assigned to all model parameters. We write the model for the potential outcome of subject $i$ compactly as

$$y_{ijt} = \mathbf{w}_i \beta + \epsilon_{ijt},$$

where the vector $\beta = (\mu, v_1, \ldots, v_{v-1}, \kappa, T_1, \ldots, T_{v-1}, \gamma', \theta')'$ comprises all regression parameters in both outcome models and $\mathbf{w}_i$ denotes the corresponding vector of covariates at panel time $t$. Thus the factor-augmented treatment model is given as

$$x_i^t = \mathbf{v}_i \alpha + \lambda_x f_{ci} + \epsilon_{xi}, \quad \epsilon_{xi} \sim \mathcal{N}(0, 1),$$

$$y_{ji} = \mathbf{w}'_j \beta + \lambda_j f_{cj} + \zeta_j f_{ji} + \epsilon_{ji}, \quad \epsilon_{ji} \sim \mathcal{N}(0, S_j),$$

where the $i$-th row of $\mathbf{w}_i$ is equal to $\mathbf{w}_i$.

We assume that the regression parameters $\alpha, \beta$, the factor loadings $\lambda_x, \lambda_0, \lambda_1, \zeta_0, \zeta_1$ and the variances of the idiosyncratic errors $S_0, S_1$ are independent a priori.

Following Jacobi et al. (2016), we perform variable selection in the selection as well as the outcome model to avoid overspecification. We assume prior independence of all coefficients in $\alpha$ and $\beta$ and specify normal priors $\mathcal{N}(0, V_0)$ and $\mathcal{N}(0, V_2)$, respectively, for the coefficients in $\alpha$ and $\beta$ not subjected to variable selection. In our application, these are the intercept in the selection equation and $\mu$ in the outcome equation. In our application, we use $V_0 = 5$ and $V_2 = 10^4$.

For all coefficients in $\alpha$ and $\beta$ subjected to variable selection, we employ spike and slab priors. A spike and slab prior is a mixture of a component concentrated at zero, the spike, and a comparably flat component, the slab. We use a Dirac spike, i.e. a point mass at zero, $\delta_0(\cdot)$ and a Normal slab. For each coefficient $\beta_\ell$ in $\beta$ subjected to variable selection, the spike and slab prior is specified hierarchically as

$$p(\beta_\ell | \delta_\ell^B) = \delta_\ell^B p(\beta_\ell | \delta_\ell^B = 1) + (1 - \delta_\ell^B) \delta_0(\beta_\ell),$$

depending on a latent binary variable $\delta_\ell^B$ with prior inclusion probability $p(\delta_\ell^B = 1) = \pi_\beta$ and hyperprior $\pi_\beta \sim \mathcal{B}(a, b)$. A similar prior is introduced for each element $\alpha_\ell$ of $\alpha$ that is subjected to variable selection, involving a binary indicator $\delta_\ell^a$ with prior inclusion probability $p(\delta_\ell^a = 1) = \pi_\alpha$ and hyperprior $\pi_\alpha \sim \mathcal{B}(a, b)$. In our application, we use normal slabs with zero mean and variance five and a uniform hyperprior $\mathcal{B}(1, 1)$ on the inclusion probabilities $\pi_\alpha$ and $\pi_\beta$.

All indicator variables are subsumed in the binary vectors $\delta^a$ and $\delta^B$, respectively, and estimated along with the model parameters during MCMC estimation (see Section 3.4). Note that estimation of the indicators $\delta_\ell^B$ corresponding to coefficients in the parameter $\beta$ identifies relevant predictors for the outcome equation under control conditions. Most importantly, estimation of the indicators $\delta_\ell^B$ corresponding to coefficients in the parameter $\theta$ allows identification of covariates inducing heterogeneous treatment effects.

Finally, we assume that all elements in the factor loading vectors, $\lambda = (\lambda_x, \lambda_0, \lambda_1)'$, $\zeta_0$ and $\zeta_1$, a priori are independent standard normal. Also the variances $\sigma^2_0$ of the idiosyncratic errors are assumed to be independent and assigned an inverse Gamma prior $\mathcal{G}^{-1}(S_0, S_0)$. We use $S_0 = S_1 = 2.5$ in our application.

3.4. Posterior inference

Posterior inference is accomplished by Markov chain Monte Carlo (MCMC) methods. To describe the sampling scheme we subsume in the vectors of latent factors $f_i$, $f_0$ and $f_1$ for all subjects and in $\Lambda$ the vectors of factor loadings $\lambda = (\lambda_x, \lambda_0, \lambda_1)'$, $\zeta_0$ and $\zeta_1$. By $\sigma_0^2 = (\sigma_0^2, \ldots, \sigma_{0T}^2)$ and $\sigma_1^2 = (\sigma_{10}^2, \ldots, \sigma_{1T}^2)$, we denote the vectors of the idiosyncratic variances.

We extend and improve the MCMC sampler introduced in Jacobi et al. (2016) for the shared factor model which was designed as a blocked Gibbs sampler, where all unknowns within a block are sampled from the respective full conditional posterior distributions, conditional on all unknown parameters in the other blocks. The shared factor model is that special case of a factor-augmented model, where $\zeta_0 = \zeta_1 = \mathbf{0}$, and the full conditional sampler of Jacobi et al. (2016) can be easily extended to the case where $\zeta_0$ and $\zeta_1$ are unconstrained.

As the idiosyncratic errors $\epsilon_{xi}, \epsilon_{0i},$ and $\epsilon_{1i}$ are independent, the augmented likelihood including the unobserved latent utilities is given as

$$p(x, x', y | \alpha, \beta, \Lambda, \sigma_0^2, \sigma_1^2, f) = \prod_{i=1}^n p(x_i | x_i', \alpha, \lambda_x, f_{ci}) \times \prod_{i=1}^n p(y_{0i} | \beta, \sigma_0^2, \lambda_0, \zeta_0, f_{ci}, f_{0i}) \prod_{i=1}^n p(y_{1i} | \beta, \sigma_1^2, \lambda_1, \zeta_1, f_{ci}, f_{1i}).$$

Conditional on the latent factors $f$, the models for the latent utilities as well as the potential outcomes are regression models with the respective latent factors as additional regressors. This suggests to sample $(\alpha, \lambda_x)$ as well as $(\beta, \lambda_0, \lambda_1, \zeta_0, \zeta_1)$ in one block given $f$. 
Such a full conditional Gibbs sampler was implemented in a research report version of this paper (Wagner et al., 2021). However, using this sampling scheme resulted in highly autocorrelated posterior draws for more or less all parameters of interest, including the regression coefficients in the selection and the outcome equations. To increase sampling efficiency, we found it advantageous to work with a partially marginalized sampler in the spirit of Liu (1994), by sampling the regression coefficients in the selection and the outcome equations from their posterior distribution marginalizing over the latent factors. This leads to an MCMC scheme comprising the following steps:

1. **Probit regression with variable selection**
   (1a) For $i = 1, \ldots, n$, sample the latent utility $\chi_i^*$ from the conditional posterior marginalised over the latent factors, $p(\chi_i^* | \alpha, \beta, \Lambda, \sigma^2_\chi, \sigma^2_\eta, x_i, y_i)$.
   (1b) Perform variable selection in the selection equation: Sample the indicators $\delta^\alpha$ and the unrestricted coefficients in $\alpha$ from the conditional posterior marginalised over the latent factors, $p(\delta^\alpha, \alpha | \beta, \Lambda, \sigma^2_\alpha, \sigma^2_\beta, x^*, y)$.
   (1c) Sample the hyperparameter $\pi_\alpha$ from the full conditional posterior $p(\pi_\alpha | \delta^\alpha)$.

2. **Outcome regression with variable selection**
   (2a) Sample the latent utility $\chi_i^*$ from the conditional posterior marginalised over the latent factors, $p(\delta^\beta, \beta | \alpha, \Lambda, \sigma^2_\beta, \sigma^2_\alpha, \chi^*, y)$.
   (2b) Sample the hyperparameter $\pi_\beta$ from the full conditional posterior $p(\pi_\beta | \delta^\beta)$.

3. **Factor analysis steps**
   (3a) For $i = 1, \ldots, n$ sample the factors $f_{d,i}$ and $f_{x,i}$ from the full conditional posterior $p(f_{d,i}, f_{x,i} | \alpha, \beta, \sigma^2_\chi, \Lambda, \lambda_x, \xi_x^*, \chi^*_i, x_i, y_i)$.
   (3b) Sample the factor loadings $\lambda_x, \lambda_j$ and $\xi_j$, $j = 0$ from their full conditional. Conditional on the common factor $f_{d,i}$, the latent utility $\chi_i^*$ and the potential outcome sequences $y_{x,i}$ are independent and, as a consequence, $\lambda_x$, $\lambda_0, \xi_0$ (and $\lambda_1, \xi_1$) can be sampled independently from their respective full conditional posteriors $p(\lambda_x | \alpha, \beta, \chi^*)$ and $p(\lambda_j, \xi_j | \beta, f, f_j, \sigma^2_j, y^j)$ for $j = 0, 1$, where $y^j = [y_j | x_j = j]$ are the observed outcome sequences of all subjects with treatment $j$.
   (3c) Perform a boosting step and a sign-switch for the latent factors and the corresponding factor loadings.

(4) For $j = 0, 1$ sample the idiosyncratic error variances $\sigma^2_j$ from the full conditional posteriors $p(\sigma^2_j | \beta, f, f_j, \lambda_j, \xi_j, y^j)$.

Full details on this MCMC scheme are provided in Appendix B. Partial marginalisation in Step (1) and (2) when performing variable selection and sampling the regression parameters leads to considerable increase of efficiency for these parameters. In Step (3c), we apply boosting based on marginal data augmentation as described in Frühwirth-Schnatter and Lopes (2018), nevertheless draws of the factor loadings are still highly autocorrelated. As an alternative, factor loadings and latent factors could also be sampled jointly as proposed in Chib et al. (2006) requiring however longer computation times. Another alternative might be to extend the boosting techniques suggested recently in Zanella and Roberts (2021) for multilevel hierarchical models to factor-augmented models.

4. **Simulation Study**

We illustrate the flexibility of the proposed factor-augmented treatment effects model in a simulation study where we generated 100 artificial data sets from a factor-augmented (FA) model. We describe the simulation setup in detail in Section 4.1 and summarize estimation results in Section 4.2. In addition, we analyse the generated data sets in Section 4.3 with a switching regression and a shared factor model to illustrate how the restrictions imposed by these models may yield biased estimates for the average treatment effects.

4.1. **Simulation Setup**

For each data set we simulate covariates, selection of treatment and the observed outcomes for $n = 5000$ subjects observed at $T = 6$ panel periods. Selection into treatment was simulated according to a probit model with structural mean given in Equation (1). We generated two covariates $v_{t1}$ and $v_{t2}$, where $v_{t1}$ is standard normal and $v_{t2}$ is binary with $p(v_{t2} = 1) = 0.5$, and a binary instrument $z_i$ with $p(z_i = 1) = 0.5$. Hence, including an intercept, $v_i = (1, v_{t1}, v_{t2}, z_i)^T$ and the corresponding vector of regression effects was set to $\alpha = (-0.8, 0.6, 0, 1)$.

To generate the outcome sequences as defined in Equations (3) and (4) we specify a linear predictor with dummies for the panel time points $t = 2, \ldots, T$ and $w_{it} = (v_{t1}, v_{t2})'$ as regressors. For the control outcomes, we choose the intercept as $\mu = 3$, panel time effects $\nu = (v_{t1}, \ldots, v_{tT-1})' = (0.1, 0.15, 0.2, 0.25, 0.3)'$ and covariate effects $\psi = (1, 0)'$. For the treated outcomes, we set the treatment effect to $\kappa = -1$, the parameters $\tau = (\tau_1, \ldots, \tau_{T-1})'$ capturing heterogeneity of panel effects to $\tau = (0, -0.1, 0.1, 0.1)'$ and those capturing heterogeneity of covariate effects to $\beta = (0, -0.2)'$. Error variances were set to $\sigma^2_\beta = 0.25 1_T$ and $\sigma^2_\alpha = 0.09 1_T$. Finally, the loadings for the common factor are $\lambda_x = 0.8, \lambda_0 = (0.6, 0.55, 0.5, 0.45, 0.4, 0.35)'$ and $\lambda_j = -\lambda_0$ and the loadings of the specific factors are $\xi_0 = 0.41 1_T$ and $\xi_1 = 0.21 1_T$.

Marginalising over the latent factors, these settings imply the following marginal correlations between latent utility and the potential outcomes
Fig. 1. Bias of the estimated average treatment effects (upper panel), and the marginal correlations of latent utility and outcome sequences under control (lower panel, left) and under treatment (lower panel, right) in the simulated data, estimated using the FA model.

\[
\text{Cor}(x^*_i, y_{0i}) = (0.408, 0.389, 0.367, 0.343, 0.316, 0.286). \tag{17}
\]

\[
\text{Cor}(x^*_i, y_{1i}) = (-0.511, -0.499, -0.489, -0.465, -0.443, -0.415). \tag{18}
\]

which implies pronounced endogeneity.

As we will discuss in more detail in Section 4.3, the correlation structure of the data generating process can be recovered neither by the switching regression (SR) model nor by the shared factor (SF) model and an analysis using these models results in biased treatment effects estimates.

4.2. Estimation results

For each data set, MCMC as outlined in Section 3.4 was run without variable selection for 10,000 iterations after discarding 10,000 iterations as burn-in.

Figure 1 shows boxplots of the bias for the regression effects in the outcome model and the marginal correlation between outcome and latent utility under the FA model. Common effects as well as treatment effects show no or little bias, but the correlation between latent utility and the outcome sequences is slightly overestimated for \( \lambda_0 \) and underestimated for \( \lambda_1 \).

Sampling efficiency is measured through inefficiency factors which vary across parameters and data sets. For each parameter of interest, the median sampling inefficiency factor over all data sets serves as a measure of inefficiency, with sampling efficiency being the higher, the closer inefficiency is to 1. For the regression effects in the selection equation, inefficiencies are in the range of 3-97 and for the effects in the outcome equation in the range of 1.6-31, except for the common intercept \( \mu \) and the treatment effect \( \kappa \) for which median sampling inefficiencies were roughly 100. Inefficiencies regarding the average dynamic treatment effects ATE are higher and range from 180-236.

Regarding the covariance matrix \( \Sigma \), the idiosyncratic variances are sampled with low inefficiencies of 4-16, the elements of \( \Sigma_0 \) and \( \Sigma_1 \) have inefficiencies in the range of, respectively, 30-40 and 44-50, while sampling the correlations of the latent utility and the potential outcome sequences leads to higher inefficiencies in the range of 196-250 for outcome 0 and 165-180 for outcome 1.

Finally, factor loadings have high sampling inefficiencies with a median of 367 for \( \lambda_0 \) and, and inefficiencies in the range of 414-480 for \( \lambda_0 \), 330-362 for \( \lambda_1 \), 523-683 for \( \xi_0 \) and 614-835 for \( \xi_1 \), respectively.

4.3. Comparison to results from SF and SR model

The data generating process was chosen so that the assumptions of both the shared factor model as well the switching regression model are violated. We analysed the simulated data sets also with these two models to illustrate the bias in the
average dynamic treatment effects estimates resulting from the misspecification of the correlation structure in both models. This is obvious for the shared factor model which assumes that all correlation within a potential outcome sequence can be attributed to the correlation to the latent utility, whereas the specific factors imply no correlation with the latent utility.

For the switching regression model, violation of the model assumption stems from the high correlations between the potential outcomes and the latent utility. As discussed in Chib and Jacobi (2007), positive definiteness of the covariance matrices \( \Omega_j \), \( j = 0, 1 \) requires that \( \sum_{t=1}^{T} \rho_{jt} < 1 \) where \( \rho_{jt} \) is the correlation between the latent utility \( x^*_t \) and the potential outcome \( j \) at time \( t \), \( y_{jt} \). This inequality is violated for the data generating process in our simulation study which implies much more pronounced correlations, see Equations (17) and (18).

For all three modeling approaches, the average dynamic treatment effects in panel period \( t \), \( \text{ATE}_t \), over all subjects was estimated as

\[
\hat{\text{ATE}}_t = \hat{\kappa}_t + \frac{1}{n} \sum_{i=1}^{n} w_i \hat{\theta}_i
\]

based on the estimated posterior means of the regression effects \( \hat{\theta} \) and \( \hat{\kappa} \) in the potential outcomes models. Figure 2 shows the bias of the estimated treatment effects for \( t = 1, \ldots, 6 \) based on the three different models. Both the shared factor and the switching regression model yield more biased treatment effects estimates than the factor-augmented model for all time points, with the largest bias observed under the switching regression model. This bias, which results from the restrictions imposed by these two models, underlines the need for the increased flexibility of the proposed factor-augmented model.

5. Analysing Dynamic Earnings Effects of Maternity Leave

Mothers exhibit very different earnings patterns and trajectories due to child caring responsibilities. This is most evident in the widely documented motherhood wage gap that refers to the fact that mothers tend to earn less than women without children (Waldfogel, 1998; Lundberg and Rose, 2000; Blau and Kahn, 2000; Anderson et al., 2002; Kleven et al., 2019). Maternal (parental) leave policies have been introduced to help new mothers to balance work and family responsibility after the birth of a child. However, such policies vary widely in the key aspects of their design (duration, job protection, financial benefits). Further, breaks in employment after birth have ambiguous labour market implications, both from theoretical (loss of human capital, rents from good job matches, increase in labour market attachment) and empirical perspectives. Empirical findings on employment and earnings implications from leave are mixed and differ, among others, across countries, length of leave, as well as regarding the time frame with several studies pointing to at least short-term negative impacts but little medium and longer term effects (Rossin-Slater, 2017; Jacobi et al., 2016; Zweimüller et al., 2009).

Most studies focus on assessing the impact of policy changes on the labour market rather than the actual leave taken. A key challenge is that mothers will decide on the length of leave or break taken after child birth based also on factors unobserved by the researcher that are likely to directly impact earnings after return. In practice we observed a considerable variation in leave taken by mothers due to observed factors, including leave policies, as well as unobserved factors. Here we apply the factor-augmented treatment effects model to analyse the causal effects of a long leave after child birth on short- and medium run earnings of Austrian mothers after their return to the labor market.

Jacobi et al. (2016) have previously investigated the same question based on a shared factor (SF) model as well as a switching regression (SR) model. The factor-augmented (FA) approach to model the joint distribution of the errors \( \varepsilon_i \) introduced in this paper can capture more flexible dependencies as it allows for separate impacts of unobserved confounders as well as other unobserved factors on earnings rather than attributing all unobserved correlation across panel earnings to a single factor. This increased flexibility may be required as there are a number of factors that may drive selection into treatment and/or earnings that are not observed in the data, including preference on investment in child education, availability of child care through family or market providers, family income, job characteristics, career aspirations and, importantly, working hours. While the latter may be partially captured by the panel period dummies, this suggests the presence of complex
dependencies in unobservables between selection and outcome equations (confounding on unobservables) as well as within panel outcomes. Hence a flexible approach like the proposed factor-augmented approach would be required to identify both selection-related and -unrelated factors driving correlation across panel earnings.

5.1. Sample and Model Specification

The data for the analysis come from the Austrian Social Security Data Base (ASSD), which is an administrative data set of the universe of Austrian employees providing detailed information on employment and maternity leave spells as well as demographic information on mothers (Zweimüller et al., 2009). In order to address endogeneity of the leave decision we exploit a change in the parental leave policy in Austria in July 2000 which extended the payment of parental leave benefits from 18 to 30 months. Following Jacobi et al. (2016) we construct a sample of 31,051 mothers who gave birth to their last child between June 1998 till July 2002, with 58% giving birth under the new policy (13% for mothers with short leave and 95% for mothers with long leave). The policy change provides a valid instrument as it was only announced in August 2001 but applied to all mothers who gave birth from July 2000.

The left panel in Figure 3 illustrates the strong impact of the policy change on leave duration. The majority of mothers returned to the labour market within 18 months before this policy change whereas afterwards most mothers took a longer maternity leave of more than 18 months. Hence, as in Jacobi et al. (2016) we define the binary treatment in terms of maternity leave duration \( m_t \) as a maternity leave longer than 18 months (long leave)

\[
x_i = \begin{cases} 
0, & \text{if } m_t \leq 18, \\
1, & \text{if } m_t > 18,
\end{cases}
\]

i.e. the control group are mothers returning to work within 18 months after the birth of their child and the treatment group are those taking a leave of more than 18 months. The binary instrument \( z_i \) is defined as

\[
z_i = \begin{cases} 
0, & \text{child born before June 30, 2000}, \\
1, & \text{child born after June 30, 2000}.
\end{cases}
\]

The aim is to identify the causal impacts of long leave on labor market outcomes in terms of the log real earnings per year \( y_t = \{y_{t1}, y_{t2}, \ldots, y_{tT}\} \) that are observed for each mother after the end of the maternity leave. Our analysis focuses on mothers with a strong labour market attachment who were employed within 30 days after end of their leave and with at least 4 consecutive periods of non-zero earnings observed after their return. The analysis is based on an unbalanced sample of 4 – 6 panel periods with yearly earnings of at least 2,980 Euros after reentry to ensure that earnings are recorded under the social security scheme. These restrictions imply 169,539 earnings observations in total. More details on the sample construction are provided in Jacobi et al. (2016). The right panel of Figure 3 shows the mean log earnings for mothers with short and long maternity leave over 6 panel periods with a substantial gap in mean log earnings except for the last panel period.

For the analysis we include a set of baseline covariates (number of children, white/blue collar status, working experience before maternity leave and baseline earnings before the first child) in the selection Equation (1), in addition to the policy instrument. The covariates in the panel outcome Equations (10) and (11) include the set of time invariant baseline variables (but not the policy instrument), an indicator for returning to the same employer, as well as time varying covariates in terms of dummy variables for panel time effect and a quadratic calendar year effect. The proportion of mothers returning to the same employer is high for the control as well as the treatment group, with 80% for short leave and 69% for long leave.
mothers. The latter can be explained by the common practice of long leave mothers to return to work for very limited hours while still on official leave before the end of the job protection period. This option, available under the leave policy as long as earnings remain under a threshold, provides an opportunity for mothers to stay connected with their employers which appeals to mothers with a strong attachment to the labour market that we consider in this analysis.

5.2. Results

The estimation results under the proposed factor-augmented model provided below are based on 500,000 iterations after a burn-in of 10,000 with variable selection starting after 5,000 iterations of the burn-in. As in the simulation study, sampling efficiency is high for the regression effects in the selection and outcome equation and the variances of the idiosyncratic errors, but low for the factor loadings \( \lambda_j, \zeta_j, j = 0, 1 \). Different from the simulation study, the loading of the common factor in the latent utility \( \lambda_0 \) is sampled with high efficiency. More details on sampling inefficiency factors are provided in Appendix C.

Table 1 presents the coefficient estimates for the panel outcome equations in terms of point estimates, their precision (posterior standard deviation) and corresponding posterior inclusion probabilities. The first set of results refers to the common effects in both potential outcomes while the second set of results refers to the modification of these effects under treatment conditions. Firstly, the results for the modification of the intercept point to a negative treatment effect. Secondly we observe panel effects that increase over time. These are further enhanced under the treatment with additional positive panel effects again increasing over time, almost doubling the effect in most periods.

Baseline labour market characteristics (experience, blue collar status and base earnings quartiles) impact earnings after return. The number of children does not impact earnings, although having at least 2 children increases the probability of choosing a longer maternity leave (see Table C1 in the Appendix that also confirms the strong policy effect on treatment selection). We observe a small positive impact from returning to the same employer (80% under short leave and 69% under long leave) for both groups. Heterogenous treatment effects are only present for mothers with base-earnings in the upper two quartiles (Q3 and Q4). For these mothers taking a long leave lowers the positive impact from baseline earnings.

We now turn to the correlation structure noting also the strong policy effect on treatment selection (see Table C1). A key motivation for the proposed factor-augmented treatment effects model was to allow for more flexible dependencies within

---

**Table 1**

Results outcome equation: posterior means (mean), standard deviation (sd, in parentheses) and estimated posterior inclusion probabilities (prob) of regression effects

<table>
<thead>
<tr>
<th></th>
<th>Common effect ((\mu, \nu, \gamma))</th>
<th>Treatment effect ((s, r, \theta))</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>mean (sd)</td>
<td>prob</td>
</tr>
<tr>
<td>intercept</td>
<td>9.310 (0.013)</td>
<td>–</td>
</tr>
<tr>
<td>panel ( t = 2 )</td>
<td>0.066 (0.004)</td>
<td>1.000</td>
</tr>
<tr>
<td>panel ( t = 3 )</td>
<td>0.107 (0.006)</td>
<td>1.000</td>
</tr>
<tr>
<td>panel ( t = 4 )</td>
<td>0.150 (0.009)</td>
<td>1.000</td>
</tr>
<tr>
<td>panel ( t = 5 )</td>
<td>0.202 (0.011)</td>
<td>1.000</td>
</tr>
<tr>
<td>panel ( t = 6 )</td>
<td>0.253 (0.013)</td>
<td>1.000</td>
</tr>
<tr>
<td>child 2</td>
<td>–0.000 (0.001)</td>
<td>0.011</td>
</tr>
<tr>
<td>child ( \geq 3 )</td>
<td>0.000 (0.001)</td>
<td>0.015</td>
</tr>
<tr>
<td>experience</td>
<td>–0.091 (0.009)</td>
<td>1.000</td>
</tr>
<tr>
<td>blue collar</td>
<td>–0.108 (0.006)</td>
<td>1.000</td>
</tr>
<tr>
<td>interaction</td>
<td>0.001 (0.004)</td>
<td>0.043</td>
</tr>
<tr>
<td>experience/blue collar</td>
<td>0.001 (0.004)</td>
<td>0.043</td>
</tr>
<tr>
<td>base earnings Q2</td>
<td>0.066 (0.006)</td>
<td>1.000</td>
</tr>
<tr>
<td>base earnings Q3</td>
<td>0.286 (0.010)</td>
<td>1.000</td>
</tr>
<tr>
<td>base earnings Q4</td>
<td>0.606 (0.010)</td>
<td>1.000</td>
</tr>
<tr>
<td>equal employer</td>
<td>0.049 (0.005)</td>
<td>1.000</td>
</tr>
<tr>
<td>(year – 1999)</td>
<td>0.050 (0.005)</td>
<td></td>
</tr>
<tr>
<td>(year – 1999)(^2)</td>
<td>–0.005 (0.000)</td>
<td></td>
</tr>
</tbody>
</table>

**Table 2**

Estimated factor loadings and idiosyncratic variances under each treatment: posterior means and standard deviations (in parentheses)

<table>
<thead>
<tr>
<th>treatment status ( x = 0 )</th>
<th>treatment status ( x = 1 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( t )</td>
<td>( t )</td>
</tr>
<tr>
<td>( \lambda_0 )</td>
<td>( \lambda_1 )</td>
</tr>
<tr>
<td>( \zeta_0 )</td>
<td>( \zeta_1 )</td>
</tr>
<tr>
<td>( \sigma^2_0 )</td>
<td>( \sigma^2_1 )</td>
</tr>
<tr>
<td>1</td>
<td>0.297 (0.016)</td>
</tr>
<tr>
<td>2</td>
<td>0.335 (0.018)</td>
</tr>
<tr>
<td>3</td>
<td>0.377 (0.012)</td>
</tr>
<tr>
<td>4</td>
<td>0.412 (0.006)</td>
</tr>
<tr>
<td>5</td>
<td>0.434 (0.003)</td>
</tr>
<tr>
<td>6</td>
<td>0.415 (0.003)</td>
</tr>
</tbody>
</table>
a factor-approach by modelling impacts of unobserved confounders and other unobserved factors on earnings via different factors rather than attributing all unobserved correlation across panel earnings to one confounding factor. Table 2 gives the estimates for the factor loadings $\lambda_j$ that capture the impact of factors reflecting endogeneity, and $\xi_j$ that capture the impact of factors unrelated to the leave decision on earnings as well estimates for the idiosyncratic error variances.

A first finding based on the relative magnitude of the loadings in $\lambda_j$ is the prominent role of the unobserved confounding factor in the overall correlation pattern across all panel periods. The negative correlations for short leave mothers suggest that these mothers accounted for unobserved factors that would have negatively impacted their earnings had they chosen a longer leave. For long leave mothers the estimates suggest the opposite, again indicating that these mothers took unobserved factors into account in their decision. These factors may include the availability of child care through family or market providers or family income. Secondly, the role of unobserved factors unrelated to the leave decision is supported by the estimates of $\xi_j$ which may reflect job characteristics, working hours after return, career opportunities, ability and job skills. While their magnitude varies across panel periods depending on the treatment and is overall slightly less pronounced than the endogenous factor, the results strongly support the presence of both factors. Note also that the magnitude of the idiosyncratic error variances is rather small compared to the factor effects. These results confirm the motivation for the introduction of the factor-augmented model to build on the parsimonious factor approach to model correlation patterns, while addressing the potential limitation of the shared factor approach that captures correlation through a single factor.

Following from the coefficient estimates on the modification of the intercept ($\kappa$) and covariates under the treatment ($\theta$) presented in Table 1 we expect heterogeneity in the treatment effect. We have negative coefficient estimates for the intercept as well as baseline variables under the long leave treatment, while positive and increasing panel coefficients suggest a reduction in the negative effect over time. The left panel in Figure 4 shows the estimated average dynamic treatment effects (averaged over all mothers in the sample) and 95% credibility intervals for the first 6 years after returning to the labour market and confirms this pattern under the factor-augmented model as well as under the shared factor and the switching regression specifications which are also presented for comparison.

In the first year after the return mothers who took a long leave on average earn more than 15% less than short leave mothers, with the gap shrinking to around 10% earnings difference in the following period. The gap finally closes after 6 periods under the FA and SR estimates. The treatment effects under the SR analysis, which also allows for different sources of correlation are closest the FA approach (except for $t = 2$). The larger flexibility due to a less rigid structure under the FA is reflected in the correlation patterns presented in the right panel of Figure 4, with considerable variation in the correlations across time periods. This could reflect issues with identification of the very general modelling separating effects from endogeneity and other factors in the correlation. The factor structure, both with one or 2 factors, yields more consistent correlation patterns across all time periods.

Fig. 4. The left panel shows estimated average dynamic treatment effects and 95%-HPD-intervals of a long maternity leave under the factor-augmented model (FA), the shared factor model (SF) and the switching regression model (SR). The right panel compares the implied correlations to the latent utility in the three models under both short (upper panel) and long leave (lower panel).
6. Conclusion

Identification and estimation of causal treatment effects in the context of endogeneity arising from self-selection into treatment is an important problem in econometrics and statistics, which is commonly investigated within a potential outcomes framework. In particular when working with observational data in areas such as in economics, approaches to control for unobserved endogeneity are needed to identify causal effects from observed behaviour (choice), such as after policy interventions, on some outcome of interest. An additional layer of complexity arises if the interest is in the estimation of treatment effects on longitudinal outcomes where correlation across outcomes can arise from two groups of unobserved factors, one related and another unrelated to the treatment selection.

This paper has introduced a factor-augmented treatment effects model for panel outcomes that extends the shared factor approach introduced in Jacobi et al. (2016) by modelling both types of unobserved factors via two different latent factors that are inferred from the data. The proposed framework builds on two benefits of the factor approach in terms the explicit modelling of both correlation sources via the more parsimonious modelling as well as the MCMC implementation via Gibbs samplers. The simulation study shows that restrictions implied by both alternative models, the SF model as well as the SR model, can lead to biased treatment effects estimates if the data generating process is more general than implied by these two models.

Our illustrative example has considered the impact of parental leave choices by new mothers on their earnings after re-entry into the labour market, revisiting the question whether longer leave choices imply permanently lower earnings. The empirical application provides evidence of both unobserved factors underpinning the correlation across earnings for mothers following re-entry in the labour market after short and long maternity leave, with the endogenous factor being somewhat more prominent and persistent across all time periods. It also affects the conclusion with the factor-augmented model providing evidence for a closing of the earnings gap of long leave mothers after 6 periods, which is not the case in the shared factor model, as well as smoother paths towards closing the gap.

The empirical results show similarities to the switching regression framework in which the endogenous factor is also separately modelled via a latent factor, while unobserved treatment-unrelated factors affecting earnings across all panels are captured via correlations between latent utility and the observed outcomes. However, due to the imposed restrictions on the correlation between treatment selection and potential outcomes, the resulting treatment effects can be severely biased as we exemplified in our simulation study. Hence, the proposed factor-augmented framework can be considered a flexible alternative to both the switching regression and the shared factor approach.

A challenging topic for future research is, among others, extending the proposed model to categorical and continuous treatments, see e.g. Alejo et al. (2018) and Lee et al. (2022).

Acknowledgment

The authors would like to thank the editor, the associate editor, two anonymous referees and the participants of the 11th European Seminar on Bayesian Econometrics for valuable comments and suggestions that helped to improve the paper considerably.

Appendix A. Moments of the observed outcomes

To derive the first two moments of the observed outcomes we start with a univariate normal random variable \(Z \sim N(\mu, \sigma^2)\), and then consider the \((T + 1)\)-variate normal random variable \((x', y')\), where the index \(i\) is dropped for simplicity.

Expectation and variance of \(Z\) truncated to \((a, b)\) are given as

\[
E(Z|a < Z < b) = \mu - \sigma \frac{\phi\left(\frac{b - \mu}{\sigma}\right) - \phi\left(\frac{a - \mu}{\sigma}\right)}{\Phi\left(\frac{b - \mu}{\sigma}\right) - \Phi\left(\frac{a - \mu}{\sigma}\right)},
\]

\[
V(Z|a < Z < b) = \sigma^2 \left[1 - \frac{1}{\sigma} \frac{(b - \mu)\phi\left(\frac{b - \mu}{\sigma}\right) - (a - \mu)\phi\left(\frac{a - \mu}{\sigma}\right)}{\Phi\left(\frac{b - \mu}{\sigma}\right) - \Phi\left(\frac{a - \mu}{\sigma}\right)} - \frac{\phi\left(\frac{b - \mu}{\sigma}\right) - \phi\left(\frac{a - \mu}{\sigma}\right)}{\Phi\left(\frac{b - \mu}{\sigma}\right) - \Phi\left(\frac{a - \mu}{\sigma}\right)}\right]^2,
\]

where \(\phi(\cdot)\) and \(\Phi(\cdot)\) are, respectively, the pdf and the cdf of the standard normal distribution.

Let \((x', y') \sim N(\mu, \Sigma)\) with moments

\[
\mu = \begin{pmatrix} \mu_x \\ \mu_y \end{pmatrix} \quad \text{and} \quad \Sigma = \begin{pmatrix} \sigma^2_x & \sigma_{xy} \\ \sigma_{yx} & \sigma^2_y \end{pmatrix},
\]

(A.1)

with the conditional distribution of \(y|x'\) given as

\[
y|x' \sim N\left(\mu_y + \frac{\sigma_{xy}}{\sigma^2_x}(x' - \mu_x), \frac{\sigma^2_y - \sigma_{xy}^2}{\sigma^2_x}\right).
\]

(A.2)
We are interested in the first two moments of the conditional distributions \( y(x^* < 0) \) and \( y(x^* > 0) \) of the outcomes \( y \) knowing the treatment status or, equivalently, the sign of \( x^* \). The conditional expectation \( E(y|x^* < 0) \) results as

\[
E(y|x^* < 0) = \mu_y + \frac{\sigma_{xy}}{\sigma_x^2} E(x^* - \mu_x|x^* < 0) = \mu_y + \frac{\sigma_{xy}}{\sigma_x^2} \frac{\phi(\mu_x/\sigma_x)}{1 - \Phi(\mu_x/\sigma_x)}.
\]

The conditional second moment of \( (y - \mu_y)(y - \mu_y)'|x^* < 0 \) can be derived as

\[
E((y - \mu_y)(y - \mu_y)'|x^* < 0) = \sum_y - \frac{\sigma_{xy}^2}{\sigma_x^2} + \frac{\sigma_{xy}^2}{\sigma_x^2} E((x^* - \mu_x)^2|x^* < 0),
\]

and, hence, the conditional covariance matrix \( V(y|x^* < 0) \) is given as

\[
V(y|x^* < 0) = \sum_y - \frac{\sigma_{xy}^2}{\sigma_x^2} + \left[ \frac{\mu_x}{\sigma_x} \frac{\phi(\mu_x/\sigma_x)}{1 - \Phi(\mu_x/\sigma_x)} - \left( \frac{\phi(\mu_x/\sigma_x)}{1 - \Phi(\mu_x/\sigma_x)} \right)^2 \right].
\]

Similarly, expectation and covariance of \( y(x^* > 0) \) can be derived as

\[
E(y|x^* > 0) = \mu_y + \frac{\sigma_{xy}}{\sigma_x} \frac{\phi(\mu_x/\sigma_x)}{\Phi(\mu_x/\sigma_x)}.
\]

\[
V(y|x^* > 0) = \sum_y - \frac{\sigma_{xy}^2}{\sigma_x^2} + \left[ \frac{\mu_x}{\sigma_x} \frac{\phi(\mu_x/\sigma_x)}{\Phi(\mu_x/\sigma_x)} + \left( \frac{\phi(\mu_x/\sigma_x)}{\Phi(\mu_x/\sigma_x)} \right)^2 \right].
\]

Appendix B. Details on MCMC sampling

To increase sampling efficiency, we sample the regression effects in the selection equation and in the outcome equations marginalising over the latent factors. To this aim, we utilize the joint distribution of the latent utility \( x^*_i \) and the potential outcome vector \( y_{ji} \), which is equal to the \((T + 1)\)-dimensional normal distribution \( (x^*_i, y_{ji})' \sim \mathcal{N}(m_{ij}, M_j) \) with moments derived from Equation (A.1):

\[
m_i = \begin{pmatrix} \nu_{i} & \alpha \end{pmatrix} \quad \text{and} \quad M_j = \begin{pmatrix} \frac{\sigma_{x}^2}{\lambda_{i} \lambda_{j}} & \lambda_{i} \lambda_{j} \Sigma_j \end{pmatrix}.
\]

where \( \Sigma_j = \lambda_{j} \lambda_{j}^* + \xi_j \lambda_{j} + S_j \) and \( S_j = \text{Diag}(\sigma_{j1}^2, \ldots, \sigma_{jT}^2) \) is a diagonal matrix.

To draw the latent factors and factor loadings sampling from the full conditional distributions is more convenient as the resulting posterior distributions are multivariate normals and hence no Metropolis-Hastings-Step is needed.

We give more details on the sampling steps in the following Sections B1 to B5.

B1. Probit regression with variable selection

The conditional distribution \( p(x^*_i | \alpha, \beta, \lambda_{i}, \lambda_{j}, \xi, \sigma_{ij}^2, y_{ji}) \) resulting from (B.1) is the univariate normal distribution \( \mathcal{N}(m_{ij}, s_{ij}^2) \) with parameters

\[
m_{ij} = \nu_{i} + \lambda_{i} \lambda_{j} \Sigma_j^{-1} (y_{ji} - \tilde{W}_i \beta),
\]

\[
s_{ij}^2 = \sigma_{x}^2 - \lambda_{i}^2 \lambda_{j} \Sigma_j^{-1} \lambda_{j}.
\]

Thus, in Step (1a) the latent utilities \( x^*_i \) are sampled from this normal distribution truncated to the interval \(( -\infty, 0 ) \) for \( x_i = 0 \) and to \(( 0, \infty ) \) for \( x_i = 1 \).

Conditional on the potential outcomes sequences \( y_{ji} \), the shifted latent utilities

\[
\tilde{x}_i = x^*_i - \lambda_{i} \lambda_{j} \Sigma_j^{-1} (y_{ji} - \tilde{W}_i \beta)
\]

follow the normal distribution \( \mathcal{N}(\nu_{i}, s_{ij}^2) \) and hence sampling Steps (1b) and (1c) are standard sampling steps for variable selection and parameter estimation in a heteroscedastic regression model, see Appendix A.3 in Jacobi et al. (2016).
B2. Outcome regression with variable selection

The conditional distribution \( p(y_{ji} | \alpha, \beta, \lambda_x, \lambda_j, \xi_j, \sigma_j^2, x_i^*) \) resulting from (B.1) follows a \( T \)-variate normal distribution,

\[
y_{ji} | x_i^*, \cdot \sim \mathcal{N}
\left( \mathbf{W}_i \beta + \frac{\lambda_x \lambda_j}{\sigma_x^2} (x_i^* - \mathbf{v}_i \alpha), \Sigma_j - \frac{\lambda_x^2 \lambda_j^2}{\sigma_x^2} \right).
\]

(B.2)

see also (A.2).

Conditional on the latent utility \( x_i^* \), the shifted outcomes

\[
y_i = y_{ji} - \frac{\lambda_x \lambda_j}{\sigma_x} (x_i^* - \mathbf{v}_i \alpha)
\]

follow a multivariate normal distribution with mean \( \mathbf{W}_i \beta \) and the conditional variance given in (B.2). Hence, sampling steps (2a) and (2b) are again those for variable selection and parameter estimation in a regression model with correlated heteroscedastic errors, detailed in Appendix A.3 in Jacobi et al. (2016).

B3. Factor analysis steps

Given the regression coefficients \( \alpha \) and \( \beta \), the errors

\[
\mathbf{e}_i = \begin{pmatrix} e_{i1} \\ e_{ji} \end{pmatrix} = \begin{pmatrix} x_i^* - \mathbf{v}_i \alpha \\ y_{ji} - \mathbf{W}_i \beta \end{pmatrix},
\]

of subject \( i \) are determined and the factors and the factor loadings are updated conditional on \( \mathbf{e}_i \). In this factor model, the error term \( e_i \) follows the \((T + 1)\)-dimensional normal distribution \( \mathbf{e}_i \sim \mathcal{N}(\mathbf{A}_j \mathbf{f}_i, \mathbf{S}_j) \) where \( \mathbf{A}_j = \begin{pmatrix} \lambda_x & 0 \\ \lambda_j & \xi_j \end{pmatrix} \) and \( \mathbf{S}_j \) is a diagonal matrix with entries 1 and \( \sigma_j^2 \). The full conditional posterior for \( \mathbf{f}_i \) is the normal distribution \( \mathcal{N}(\mathbf{f}_{n,i}, \mathbf{F}_{n,i}) \) with posterior moments given as

\[
\mathbf{F}_{n,i} = (\mathbf{A}_j' \mathbf{S}_j^{-1} \mathbf{A}_j + \mathbf{I}_2)^{-1},
\]

\[
\mathbf{f}_{n,i} = \mathbf{F}_{n,i} \mathbf{A}_j' \mathbf{S}_j^{-1} \mathbf{e}_i.
\]

Conditional on the common factor \( f_{ci} \), the latent utility \( x_i^* \) and the potential outcomes sequence \( y_{ji,n} \) are independent. Thus the loadings \( \lambda_x, (\lambda_0, \xi_0) \) and \( (\lambda_1, \xi_1) \) can be sampled independently from the full conditionals of the regression models

\[
\mathbf{e}_x = \mathbf{f}_i \lambda_x + \mathbf{e}_x, \quad \mathbf{e}_x \sim \mathcal{N}(\mathbf{0}, \mathbf{I}_n),
\]

and for each \( j = 0, 1 \) from all subjects \( i \) where \( x_i = j \):

\[
\mathbf{e}_{ij} = \begin{pmatrix} f_{ci} \mathbf{1}_r \\ f_{ji} \mathbf{1}_r \end{pmatrix} \begin{pmatrix} \lambda_j \\ \xi_j \end{pmatrix} + \mathbf{e}_{ji}, \quad \mathbf{e}_{ji} \sim \mathcal{N}(\mathbf{0}, \mathbf{S}_j).
\]

B4. Sampling the error variances

The full conditional posterior of \( \sigma_j^2 \) is the inverse Gamma distribution \( \mathcal{G}^{-1}(s_{n,j}, r_{n,j}) \) with parameters

\[
s_{n,j} = s_{0,j} + n_{jt}/2, \quad r_{n,j} = r_{0,j} + S_{\varepsilon,j} / 2,
\]

where \( n_{jt} \) is the number of subjects for which \( y_{j,it} \) is observed in panel period \( t \) and

\[
S_{\varepsilon,j} = \sum_{i:x_i=j} (e_{j,it} - f_{ci} \lambda_{jt} - f_{ji} \xi_{jt})^2.
\]

where \( e_{j,it} = y_{j,it} - \mathbf{W}_i \beta \).

B5. Sampling the inclusion probabilities

In Step (1c), sample \( \pi_\alpha | k_\alpha \) from \( B(1 + k_\alpha, 1 + N_\alpha - k_\alpha) \) and in Step (2b), sample \( \pi_\beta | k_\beta \) from \( B(1 + k_\beta, 1 + N_\beta - k_\beta) \) where \( k_\alpha = \sum \delta_\alpha^\alpha \) is the number of selected regressors for the latent utility and \( k_\beta = \sum \delta_\beta^\beta \) accordingly is the number of selected regressors for the potential outcome equations. \( N_\alpha \) and \( N_\beta \) are the number of elements in, respectively, \( \delta^\alpha \) and \( \delta^\beta \).
Table C1
Results of selection equation: posterior means (mean), standard deviation (sd) and estimated posterior inclusion probabilities (prob) of the regression effects.

<table>
<thead>
<tr>
<th></th>
<th>mean</th>
<th>sd</th>
<th>prob</th>
</tr>
</thead>
<tbody>
<tr>
<td>intercept</td>
<td>-1.622</td>
<td>0.034</td>
<td>1.000</td>
</tr>
<tr>
<td>z</td>
<td>2.934</td>
<td>0.024</td>
<td>1.000</td>
</tr>
<tr>
<td>child 2</td>
<td>0.055</td>
<td>0.037</td>
<td>0.753</td>
</tr>
<tr>
<td>child ≥ 3</td>
<td>-0.012</td>
<td>0.033</td>
<td>0.152</td>
</tr>
<tr>
<td>experience</td>
<td>0.095</td>
<td>0.028</td>
<td>0.984</td>
</tr>
<tr>
<td>blue collar</td>
<td>-0.062</td>
<td>0.045</td>
<td>0.706</td>
</tr>
<tr>
<td>interaction experience/blue collar</td>
<td>-0.017</td>
<td>0.041</td>
<td>0.177</td>
</tr>
<tr>
<td>base earnings Q2</td>
<td>0.002</td>
<td>0.011</td>
<td>0.054</td>
</tr>
<tr>
<td>base earnings Q3</td>
<td>-0.001</td>
<td>0.008</td>
<td>0.035</td>
</tr>
<tr>
<td>base earnings Q4</td>
<td>-0.158</td>
<td>0.028</td>
<td>1.000</td>
</tr>
</tbody>
</table>

Table C2
Results of the outcome equations: posterior means, standard deviations (in parentheses); inclusion of regression effects based on posterior inclusion probabilities > 0.5 is indicated by “*”.

<table>
<thead>
<tr>
<th></th>
<th>SR Model</th>
<th>SF Model</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Common effect ($\mu, v(y)$)</td>
<td>Treatment effect ($x, \tau, \theta$)</td>
</tr>
<tr>
<td>intercept</td>
<td>9.331 (0.011)</td>
<td>-0.112 (0.009)</td>
</tr>
<tr>
<td>panel t=2</td>
<td>0.072* (0.004)</td>
<td>0.061* (0.005)</td>
</tr>
<tr>
<td>panel t=3</td>
<td>0.117* (0.006)</td>
<td>0.094* (0.005)</td>
</tr>
<tr>
<td>panel t=4</td>
<td>0.162* (0.007)</td>
<td>0.107* (0.005)</td>
</tr>
<tr>
<td>panel t=5</td>
<td>0.214* (0.009)</td>
<td>0.113* (0.006)</td>
</tr>
<tr>
<td>panel t=6</td>
<td>0.261* (0.011)</td>
<td>0.132* (0.007)</td>
</tr>
<tr>
<td>child 2</td>
<td>-0.000 (0.001)</td>
<td>-0.000 (0.001)</td>
</tr>
<tr>
<td>child ≥3</td>
<td>0.000 (0.001)</td>
<td>0.000 (0.002)</td>
</tr>
<tr>
<td>experience</td>
<td>-0.091* (0.009)</td>
<td>0.010 (0.013)</td>
</tr>
<tr>
<td>blue collar</td>
<td>-0.102* (0.006)</td>
<td>0.000 (0.001)</td>
</tr>
<tr>
<td>int. exp./blue collar</td>
<td>0.000 (0.003)</td>
<td>0.010 (0.015)</td>
</tr>
<tr>
<td>base earnings Q2</td>
<td>0.069* (0.006)</td>
<td>0.000 (0.003)</td>
</tr>
<tr>
<td>base earnings Q3</td>
<td>0.292* (0.009)</td>
<td>-0.050* (0.012)</td>
</tr>
<tr>
<td>base earnings Q4</td>
<td>0.610* (0.010)</td>
<td>-0.118* (0.013)</td>
</tr>
<tr>
<td>equal employer</td>
<td>0.051* (0.005)</td>
<td>0.000 (0.002)</td>
</tr>
<tr>
<td>(year – 1999)</td>
<td>0.034 (0.003)</td>
<td></td>
</tr>
<tr>
<td>(year – 1999)$^2$</td>
<td>-0.004 (0.0002)</td>
<td></td>
</tr>
</tbody>
</table>

Appendix C. Further details for the analysis of the mother data

The data set contains information on 31,051 mothers with earnings after return to labor market observed over 4-6 consecutive panel periods. Covariates are specified as in the analysis of Jacobi et al. (2016). Covariates included in the selection equation are an indicator for the policy change (z = 1 indicates longer payment of parental leave benefits), indicator variables for the child (child 2 = 1 if the mother had already a child and child ≥ 3 = 1 if the mother had 2 or more older children before the child birth for which the maternity leave is taken), the working experience (experience = 1 if the working experience is above the median working experience in the sample), type of contract (blue collar or white collar), the interaction between these two variables and finally indicators to control for earnings before the first child in terms of quartiles. The outcome model additionally includes indicator variables for panel periods 2-6 and for return to the same employer (equal employer), and a quadratic calendar year effect.

MCMC estimation as outlined in Section 3.4 was run for 500,000 iteration after a burn-in of 10,000 with variable selection starting after 5,000 iterations of the burn-in. Employing the partially marginalized sampling scheme results in low sampling inefficiency factors for the regression effects in selection (from 5 to 36) and outcome equation (from 1 to 38) as well as the idiosyncratic variances (from 2 to 29). Also the loading of the common factor in the latent utility, $\lambda_x$, is sampled with high efficiency (inefficiency factor 14.1) but sampling inefficiencies for the loadings in the outcome equations are typically high (in the range of, respectively; 25 – 1430 for $\lambda_{0}$, 1532 – 1844 for $\lambda_{1}$, 1078 – 1471 for $\zeta_{0}$ and 1766 – 1842 for $\zeta_{1}$). We obtained essentially the same posterior means for all model parameters from the first 50,000 iterations (after burnin) of this chain as well as with 50,000 iterations (after burnin) of another chain with different starting values.

Table C1 reports results for the posterior means of the regression effects $\alpha$ in the selection equation.

The posterior mean of the factor loading of the common factor in the selection equation is $\hat{\lambda}_x = 0.277$ (sd = 0.014).

Table C2 provides estimation results for the parameter of the outcomes model from the analyses with the SF and the SR model.
References


Frühwirth-Schnatter, S., Lopes, H., 2018. Sparse Bayesian Factor Analysis when the Number of Factors is Unknown. arXiv preprint 1804.04231


