Conditional inference and bias reduction for partial effects estimation of fixed-effects logit models

Francesco Bartolucci[1] · Claudia Pigini[2] · Francesco Valentini[2]

Abstract

We propose a multiple-step procedure to compute average partial effects (APEs) for fixed-effects static and dynamic logit models estimated by (pseudo) conditional maximum likelihood. As individual effects are eliminated by conditioning on suitable sufficient statistics, we propose evaluating the APEs at the maximum likelihood estimates for the unobserved heterogeneity, along with the fixed-T consistent estimator of the slope parameters, and then reducing the induced bias in the APEs by an analytical correction. The proposed estimator has bias of order $O(T^{-2})$, it performs well in finite samples and, when the dynamic logit model is considered, better than alternative plug-in strategies based on bias-corrected estimates for the slopes, especially in panels with short T. We provide a real data application based on labour supply of married women.

Keywords Average partial effects · Bias reduction · Binary panel data · Conditional Maximum Likelihood

JEL Classification C12 · C23 · C25

Francesco Bartolucci, Claudia Pigini and Francesco Valentini have contributed equally to this work.

✉ Francesco Valentini
f.valentini@univpm.it

Francesco Bartolucci
francesco.bartolucci@unipg.it

Claudia Pigini
c.pigini@univpm.it

[1] Department of Economics, University of Perugia, Via A. Pascoli 20, 06123 Perugia, PG, Italy

[2] Department of Economic and Social Sciences, Marche Polytechnic University, Piazzale R. Martelli 8, 60121 Ancona, AN, Italy

1 Introduction

Static and dynamic binary choice models are largely employed in microeconometric applications.[1] For these models, the fixed-effects approach is often advocated, as it allows for the estimation of partial effects of covariates that may be correlated with the individual specific unobserved heterogeneity in a nonparametric manner. However, unless the number of time occasions T goes to infinity, the maximum likelihood (ML) estimator of fixed-effects binary choice models is inconsistent due to the incidental parameters problem, that is, the presence of nuisance parameters whose number increases with the sample size (Lancaster 2000; Neyman and Scott 1948).

A popular method to overcome the incidental parameters problem is based on the conditional inference approach for the fixed-effects logit model (Andersen 1970; Chamberlain 1980), which admits sufficient statistics for the individual intercepts. The conditional ML (CML) method produces a fixed-T consistent estimator of the slope parameters, which makes it a particularly attractive strategy given the large availability of data sets, such as those produced by national household and workforce surveys, that are based on a rotating sampling scheme. However, a drawback of the CML approach is that plug-in estimates of the average partial effects (APEs) are not directly available, as the parameters for the individual effects are eliminated.

An alternative way to deal with the incidental parameters problem is to correct the bias of the ML estimator of the slope parameters. While the resulting estimator will not be fixed-T consistent, this approach is general and can be applied to any nonlinear model, as opposed to the CML strategy that is specific for the logit model. Several procedures have been proposed to reduce the order of the bias of the ML estimator from $O(T^{-1})$ to $O(T^{-2})$. Analytical bias corrections are provided by Fernández-Val (2009), whose derivations are based on general results for static (Hahn and Newey 2004) and dynamic (Hahn and Kuersteiner 2011) nonlinear panel data models. In the same vein, a modified log-likelihood function and score equation are proposed by Bester and Hansen (2009) and Carro (2007), respectively, to achieve the same bias reduction. An alternative bias correction method relies on the panel jackknife. A general procedure for nonlinear static panel data models is proposed by Hahn and Newey (2004), whereas a split-panel jackknife estimator is developed by Dhaene and Jochmans (2015) for dynamic models.

The advantage of relying on ML estimates is that plug-in estimators of the APEs are readily available. The estimators of the APEs share the same order of bias as the ML estimator, and the related corrections can be operated in a similar manner. However, especially in the case of dynamic binary choice models, the ML estimator is known to be severely biased (Heckman 1981) and the bias-corrected ML estimator exhibits a greater distortion compared to fixed-T consistent ones with short T (Carro

---

[1] Estimators of (dynamic) discrete choice models have been employed in seminal papers related to labour market participation (Heckman and Borjas 1980), and specifically to female labour supply and fertility choices (Hyslop 1999), self-reported health status (Contoyannis et al. 2004), poverty traps (Cappellari and Jenkins 2004), household finance (Alessie et al. 2004), and unionization of workers (Wooldridge 2005). More recently their application has been extended to the fields of firms' access to credit (Pigini et al. 2016), migrants' remitting behaviour (Bettin et al. 2018), energy consumption (Drescher and Janzen 2021), and innovation (Arroyabe and Schumann 2022).

2007; Bartolucci and Nigro 2012). Practitioners are also usually more familiar with the CML approach, as it has become prominent in the applied literature perhaps due to its long-standing presence in graduate textbooks (Cameron and Trivedi 2005; Hsiao 2005; Wooldridge 2010) and its built-in implementation in popular microeconometrics software.

In this paper, we propose a multiple-step procedure to estimate the APEs in the fixed-effects logit model that combines the conditional inference approach with a bias reduction method. The APEs are evaluated at the fixed-T consistent CML estimator of the slope parameters and at the ML estimator for the unobserved heterogeneity, obtained by maximizing the log-likelihood evaluated at the CML estimate. Plugging in the estimated fixed effects produces an additional source of bias in the APEs; we reduce the order of this bias from $O(T^{-1})$ to $O(T^{-2})$ by applying the analytical correction proposed by Fernández-Val (2009). In this respect, our approach extends the work of Stammann et al. (2016), who study the same plug-in estimator without the bias correction.

The proposed procedure cannot be directly extended to the dynamic logit model (Hsiao 2005), for which CML inference for the slope parameters is not viable in a simple form. This is overcome by Bartolucci and Nigro (2010), who propose a quadratic exponential (QE) formulation (Cox 1972) for dynamic binary panel data models, which has the advantage of admitting sufficient statistics for the individual intercepts. Furthermore, Bartolucci and Nigro (2012) propose a QE model that closely approximates the dynamic logit model, the parameters of which can easily be estimated by pseudo CML (PCML). The resulting PCML estimator is consistent in the absence of state dependence, because in this case the QE model corresponds to the dynamic logit model and, as shown by Bartolucci and Nigro (2012) by simulation, it otherwise exhibits a moderate bias. We therefore extend the proposed procedure to include PCML estimates in the APEs when a dynamic logit is specified.

As it happens with the APE estimators based on analytical and jackknife corrections, the proposed method reduces the order of the bias from $O(T^{-1})$ to $O(T^{-2})$. However, such a bias is asymptotically negligible under rectangular array asymptotics as plug-in average-effect estimators converge at the rate $n^{-1/2}$ (Dhaene and Jochmans 2015), where n is the sample size. Yet in spite of the asymptotic equivalence of bias-corrected and ML plug-in APE estimators, the simulation evidence provided by Dhaene and Jochmans (2015) suggests that operating some bias reduction entails a non-negligible improvement in small samples, especially with small values of T.

The proposed combination of the conditional inference approach with bias reduction provides a way to readily obtain APE estimates for the fixed-effects static and dynamic logit models. By means of an extensive simulation study, we show that the proposed approach has finite sample performance comparable to the ML and bias-corrected estimators with the static logit model, while it outperforms them when the dynamic logit is considered, especially when n and T are small. This is the result of plugging a fixed-T consistent estimator of the slope and state dependence parameters of the QE model into the APEs.

It is worth clarifying that while the CML and PCML estimators are fixed-T consistent, the asymptotic theory for the APE estimator here proposed is based on both $n, T \to \infty$, due to the presence of the ML estimates of the individual intercepts and,

consequently, of the bias correction, which is derived by means of a large-T asymptotic expansion.[2] This means that a few typical drawbacks of this setting theoretically apply here as well, namely time effects are ruled out, as they also are incidental parameters when $T \to \infty$, and the rest of the covariates are required to be stationary (Dhaene and Jochmans 2015; Fernández-Val 2009). However, by means of dedicated simulation exercises, we show that the proposed approach is able to handle the presence of time dummies in the model specification and that violations of the stationarity assumption does not seem to affect in practice the finite-sample performance of all the approaches considered.

The rest of the paper is organized as follows: In Sect. 2, we briefly discuss how the incidental parameters problem affects the APEs estimator; in Sect. 3, we recall the bias correction strategies for APE estimators, and then, we illustrate the proposed methodology and its extension to accommodate the dynamic logit model; in Sect. 4, we investigate by simulation the finite sample performance of the proposed estimator, compare it with the panel jackknife and analytical bias correction strategies, and illustrate the results of some robustness exercises; in Sect. 5, we provide a real data application based on labour supply of married women; finally, Sect. 6 provides the main conclusions.

2 Average partial effects and the incidental parameters problem

We consider n units, indexed with $i = 1, \ldots, n$, observed at time occasions $t = 1, \ldots, T$. Let y_{it} be the binary response variable for unit i at occasion t and $\boldsymbol{x}_{it}$ the corresponding vector of K strictly exogenous covariates. Under the static model, we assume that the y_{it} are conditionally independent, given α_i and $\boldsymbol{x}_{it}$, across i and t. Consider the logit formulation

$$p(y_{it}|\boldsymbol{x}_{it}; \alpha_i, \boldsymbol{\beta}) = \frac{\exp\left[y_{it}(\alpha_i + \boldsymbol{x}_{it}'\boldsymbol{\beta})\right]}{1 + \exp(\alpha_i + \boldsymbol{x}_{it}'\boldsymbol{\beta})}, \tag{1}$$

where α_i is the individual specific intercept and the vector $\boldsymbol{\beta}$ collects the regression parameters.

The fixed-effects estimator is obtained by ML, treating each individual effect α_i as a parameter to be estimated. The ML estimator of $\boldsymbol{\beta}$ is obtained by concentrating out the α_i as the solution to

$$\hat{\boldsymbol{\beta}} = \operatorname*{argmax}_{\boldsymbol{\beta}} \sum_{i=1}^{n} \sum_{t=1}^{T} \log p(y_{it}|\boldsymbol{x}_{it}; \hat{\alpha}_i(\boldsymbol{\beta}), \boldsymbol{\beta}),$$

$$\hat{\alpha}_i(\boldsymbol{\beta}) = \operatorname*{argmax}_{\alpha_i} \sum_{t=1}^{T} \log p(y_{it}|\boldsymbol{x}_{it}; \alpha_i, \boldsymbol{\beta}).$$

---

[2] In any case, the focus on the large n and large T perspective is necessary as APEs are often not point identified with fixed T (Chernozhukov et al. 2013).

Notice that $\hat{\alpha}_i(\boldsymbol{\beta})$ is estimated using only the data for subject i and it is therefore not consistent for α_{i0} unless $T \to \infty$. As a consequence, with T fixed and only $n \to \infty$, the ML estimator of $\hat{\boldsymbol{\beta}}$ will be plagued by the estimation noise in $\hat{\alpha}_i(\boldsymbol{\beta})$ and will not be consistent for $\boldsymbol{\beta}_0$, with $\underset{n\to\infty}{\text{plim}} \ \hat{\boldsymbol{\beta}} \equiv \boldsymbol{\beta}_T \neq \boldsymbol{\beta}_0$. This is the well-known incidental parameters problem (Lancaster 2000; Neyman and Scott 1948). In particular, Hahn and Newey (2004) show that $\boldsymbol{\beta}_T = \boldsymbol{\beta}_0 + B/T + O(T^{-2})$, which clarifies that $\boldsymbol{\beta}_T \to \boldsymbol{\beta}_0$ if $T \to \infty$ and n is fixed. Moreover, if both $n, T \to \infty$, $\hat{\boldsymbol{\beta}}$ will be asymptotically normal. However, Hahn and Newey (2004) show that the asymptotic distribution of the ML estimator will not be centred at its probability limit if n grows proportionally to T.

The incidental parameters problem affects the estimation of APEs as well; these effects are usually of interest to practitioners who want to quantify the influence of some regressor x on the response probability, other things being equal. For the logit model in (1), the partial effect of covariate x_{itk} for i at time t on the probability of $y_{it} = 1$ can be written, depending on the typology of the covariate, as

$$m_{itk}(\alpha_i, \boldsymbol{\beta}, \boldsymbol{x}_{it}) = \begin{cases} p(y_{it} = 1 | \boldsymbol{x}_{it}; \alpha_i, \boldsymbol{\beta})\left[1 - p(y_{it} = 1 | \boldsymbol{x}_{it}; \alpha_i, \boldsymbol{\beta})\right]\beta_k, & x_{itk} \ \text{continuous}, \\ \\ p(y_{it} = 1 | \boldsymbol{x}_{it,-k}, x_{itk} = 1; \alpha_i, \boldsymbol{\beta}) - \\ p(y_{it} = 1 | \boldsymbol{x}_{it,-k}, x_{itk} = 0; \alpha_i, \boldsymbol{\beta}), & x_{itk} \ \text{binary}, \end{cases}$$

where $\boldsymbol{x}_{it,-k}$ denotes the subvector of all covariates but x_{itk}. The true APE of the kth covariate can then be obtained by simply taking the expected value of $m_{itk}(\alpha_i, \boldsymbol{\beta}, \boldsymbol{x}_{it})$ with respect to $\boldsymbol{x}_{it}$ and α_{i0},

$$\mu_{k0} = \int m_{itk}(\alpha_{i0}, \boldsymbol{\beta}_0, \boldsymbol{x}_{it}) dG(\alpha_{i0}, \boldsymbol{x}_{it}),$$

where $G(\alpha_{i0}, \boldsymbol{x}_{it})$ denotes the joint distribution of α_{i0} and $\boldsymbol{x}_{it}$. An estimator of μ_{k0} can be obtained by plugging in the ML estimators $\hat{\boldsymbol{\beta}}$ and $\hat{\alpha}_i(\hat{\boldsymbol{\beta}})$, so that

$$\hat{\mu}_k = \frac{1}{nT} \sum_{i=1}^{n} \sum_{t=1}^{T} m_{itk}(\hat{\alpha}_i(\hat{\boldsymbol{\beta}}), \hat{\boldsymbol{\beta}}, \boldsymbol{x}_{it}). \tag{2}$$

It is now clear that, with small T, this estimator is plagued by two sources of bias: the first stems from the estimation error introduced by $\hat{\alpha}_i(\boldsymbol{\beta})$; the second is a result of using the asymptotically biased estimator $\hat{\boldsymbol{\beta}}$.

In order to better understand how these components will affect the bias of the APE, it is useful to introduce an expansion of $\hat{\alpha}_i(\boldsymbol{\beta})$ as $T \to \infty$:

$$\hat{\alpha}_i(\boldsymbol{\beta}) = \alpha_{i0} + \frac{\xi_i}{T} + \frac{1}{T} \sum_{t=1}^{T} \tau_{it} + o_p\left(\frac{1}{T}\right), \tag{3}$$

where $\frac{1}{\sqrt{T}} \sum_{t=1}^{T} \tau_{it} \xrightarrow{d} N(0, \sigma_i^2)$, which follows from higher-order asymptotics for time series data (Bao and Ullah 2007). Dhaene and Jochmans (2015) then show that the combined asymptotic bias is

$$\underset{n \to \infty}{\text{plim}}\ \hat{\mu}_k = \mu_{k0} + \frac{D + E}{T} + O(T^{-2}), \tag{4}$$

where, specifically,

$$D = \sum_{j=0}^{\infty} \mathrm{E}_T \left[\frac{\partial m_{itk}(\alpha_{i0}, \boldsymbol{\beta}_0)}{\partial \alpha_i} \tau_{i,t-j} \right] + \mathrm{E}_T \left[\frac{\partial m_{itk}(\alpha_{i0}, \boldsymbol{\beta}_0)}{\partial \alpha_i} \xi_i \right]$$
$$+ \frac{1}{2} \mathrm{E}_T \left[\frac{\partial^2 m_{itk}(\alpha_{i0}, \boldsymbol{\beta}_0)}{\partial \alpha_i^2} \sigma_i^2 \right] \tag{5}$$

is the bias generated from using $\hat{\alpha}_i(\boldsymbol{\beta})$ instead of α_{i0}.[3] Following Hahn and Newey (2004), this expression suggests that the bias introduced by plugging-in $\hat{\alpha}_i(\boldsymbol{\beta})$ has three components: (i) the asymptotic bias of $\hat{\alpha}_i(\boldsymbol{\beta})$; (ii) the correlation between $\hat{\alpha}_i(\boldsymbol{\beta})$ and $\hat{\boldsymbol{\beta}}$ depending on the same data; (iii) the variance of $\hat{\alpha}_i(\boldsymbol{\beta})$. This result can be clarified by noticing that D comes from an expansion of the APE around α_{i0}, where the term $\hat{\alpha}_i(\boldsymbol{\beta}) - \alpha_{i0}$ is characterized by (3). Furthermore,

$$E = \mathrm{E}_T \left[\frac{\partial m_{itk}(\alpha_{i0}, \boldsymbol{\beta}_0)}{\partial \boldsymbol{\beta}'} B \right]$$

is the bias from plugging in $\hat{\boldsymbol{\beta}}$, instead if using $\boldsymbol{\beta}_0$.

Further insights can be drawn from Expression (4). First, notice that even if a fixed–T consistent estimator of $\boldsymbol{\beta}_0$ was available, the asymptotic bias of the APE estimator would still be of order $O(T^{-1})$ because of the presence of D. Secondly, using a bias-corrected estimate of α_i along with a fixed-T consistent estimator of $\boldsymbol{\beta}$ would not remove the bias of order $O(T^{-1})$, as it would not take care of the last component in D.

The sources of bias discussed above, however, have been shown to become asymptotically negligible under rectangular array asymptotics, as plug-in estimators of average effects converge at a rate slower than $(nT)^{-1/2}$. Dhaene and Jochmans (2015) summarize this property in their Theorem 5.1, which is based on the following rationale. Consider the infeasible estimator

$$\mu_k^* \equiv \frac{1}{nT} \sum_{i=1}^{n} \sum_{t=1}^{T} m_{itk}(\alpha_{i0}, \boldsymbol{\beta}_0, \boldsymbol{x}_{it}),$$

---

[3] Expressions for τ_{is}, ξ_i, and σ_i^2 for panel binary choice models are given in Fernández-Val (2009).

and let μ_{ik} be the individual-specific average partial effect, with mean μ_{k0} and finite variance. Then, μ_k^* can be written as:

$$\mu_k^* = \frac{1}{n} \sum_{i=1}^{n} \mu_{ik} + \frac{1}{n} \sum_{i=1}^{n} \left(\frac{1}{T} \sum_{t=1}^{T} m_{itk}(\alpha_{i0}, \boldsymbol{\beta}_0, \boldsymbol{x}_{it}) - \mu_{ik} \right).$$

Notice that the first term converges to μ_{k0} at the rate $n^{-1/2}$, whereas the second term converges to zero at the rate $(nT)^{-1/2}$, thus implying that the infeasible APE estimator will converge no faster than $n^{-1/2}$.

From the above expression, it is straightforward to notice that any feasible average-effect estimator will converge at the same rate as μ_k^*, thus making the bias introduced by replacing α_{i0} and $\boldsymbol{\beta}_0$ with ML estimates, or their first order bias-corrected versions, asymptotically negligible. However, based on their simulation evidence, Dhaene and Jochmans (2015) still suggest using some bias correction of the APE estimator in finite samples, especially when T is small. The proposed method operates such bias reduction, as well as the alternative analytical and jackknife bias corrections recalled in the following section.

3 Estimation of average partial effects

In the following, we briefly review the existing strategies based on analytical and jackknife bias corrections, which represent the benchmark for the finite sample performance of the proposed estimator. We then illustrate the proposed methodology, which combines the consistent CML estimator of $\boldsymbol{\beta}_0$ and the analytical bias correction for the APE. Finally, we turn to the dynamic logit, for which the proposed procedure is based on a PCML estimator.

3.1 Existing strategies

The available bias reduction techniques for the estimation of APEs for fixed-effects binary choice models are based on either analytical or jackknife bias corrections.[4]

Analytical bias corrections for the APEs amount to plug-in a bias corrected estimate of $\boldsymbol{\beta}$, say $\hat{\boldsymbol{\beta}}^c = \hat{\boldsymbol{\beta}} - \hat{B}/T$, instead of the ML estimate in expression (2), along with $\hat{\alpha}_i(\hat{\boldsymbol{\beta}}^c)$. Doing so effectively removes the E component of the bias in (4), but the APE estimator is still plagued by the estimation noise in $\hat{\alpha}_i(\boldsymbol{\beta})$, giving rise to the D component. In order to remove it, the bias-corrected estimator of μ_k is computed as:

$$\hat{\mu}_k^c = \frac{1}{nT} \sum_{i=1}^{n} \sum_{t=1}^{T} m_{itk}(\hat{\alpha}_i(\hat{\boldsymbol{\beta}}^c), \hat{\boldsymbol{\beta}}^c, \boldsymbol{x}_{it}) - \hat{D}, \tag{6}$$

---

[4] In the following discussion, we will use the notation for the static logit model, unless required otherwise. Nonetheless, everything that follows can be generalized to the dynamic logit model.

where $\hat{D}$ is the sample counterpart of D in (5) and then evaluated at $\hat{\alpha}_i(\hat{\boldsymbol{\beta}}^c)$ and $\hat{\boldsymbol{\beta}}^c$. Expressions for panel binary choice models are given in Fernández-Val (2009),[5] whose derivations are based on general results for static (Hahn and Newey 2004) and dynamic (Hahn and Kuersteiner 2011) nonlinear panel data models. For the expressions as well as for further details, we refer the reader to Hahn and Newey (2004), Fernández-Val (2009), and Hahn and Kuersteiner (2011). A similar correction of the APEs is provided by Bester and Hansen (2009), who also perform a comparison of their proposal with alternative strategies.[6] Furthermore, for the static logit model, Stammann et al. (2016) develop a computationally more efficient implementation of the bias corrected estimator of structural parameters and APEs proposed by Hahn and Newey (2004), based on a pseudo-demeaning algorithm.

An alternative bias correction method for the APE estimator relies on the panel jackknife. A general procedure for nonlinear static panel data models is proposed by Hahn and Newey (2004). Let $\hat{\boldsymbol{\beta}}^{(t)}$ and $\hat{\alpha}_i^{(t)}(\hat{\boldsymbol{\beta}}^{(t)})$ be the ML estimators with the tth observation excluded for every subject. Then, the jackknife corrected estimator for the APE is

$$\hat{\mu}_k^c = T\hat{\mu}_k - \frac{T-1}{T} \sum_{t=1}^{T} \mu_k \left(\hat{\alpha}_i^{(t)}(\hat{\boldsymbol{\beta}}^{(t)}), \hat{\boldsymbol{\beta}}^{(t)} \right).$$

If the set of model covariates includes the lag of explanatory variables, then leaving out one of the T observations at a time becomes unsuitable. Instead, a block of consecutive observations has to be considered so as to preserve the dynamic structure of the data. The so-called split panel jackknife estimator is proposed by Dhaene and Jochmans (2015). A simple version of the estimator is the half-panel jackknife, which is based on splitting the panel into two half-panels, also non-overlapping if T is even and $T \geq 6$, with $T/2$ time periods. Denote the set of half-panels as

$$S = \{S_1, S_2\}, \quad S_1 = \{1, \ldots, T/2\}, S_2 = \{T/2 + 1, \ldots, T\};$$

then the half-panel jackknife estimator of the APE is

$$\hat{\mu}_k^{(1/2)} = 2\hat{\mu}_k - \frac{1}{2} \left(\bar{\mu}_k^{S_1} + \bar{\mu}_k^{S_2} \right),$$

where $\bar{\mu}_k^{S_1}$ and $\bar{\mu}_k^{S_2}$ are the plug-in estimators evaluated at the ML estimators of the individual effects and slope parameters obtained using the observations in subpanels S_1 and S_2, respectively. Dhaene and Jochmans (2015) also illustrate generalized versions of the half-panel jackknife to deal with odd T and overlapping subpanels, as well as an alternative jackknife estimator based on the split-panel log-likelihood correction.

It is finally worth mentioning that jackknife and analytical higher-order bias corrections for the slope parameters have also been proposed by Dhaene and Jochmans

---

[5] The term ξ_i is denoted by β_i and the term τ_{it} by ψ_{it} in Fernández-Val (2009).

[6] The approach by Bester and Hansen (2009) is not considered here, as it is very similar to that put forward by Fernández-Val (2009) and not specific to static and dynamic binary choice models.

(2015) and Dhaene and Sun (2021), respectively, and can be extended to the estimation of APEs, but their viability in dynamic models is limited. In the first case, the authors warn that the magnitude of the bias of the terms that are not eliminated increases in the order of the correction with the half-panel jackknife. In the second case, the correction is developed under the assumption of independent observations, thus ruling out its application to dynamic models completely.

3.2 Proposed methodology

The proposed multiple-step strategy is based on removing the two sources of bias in (4) by (a) using the fixed-T consistent CML estimator of $\boldsymbol{\beta}$, $\tilde{\boldsymbol{\beta}}$, instead of the ML estimator $\hat{\boldsymbol{\beta}}$ and (b) reducing the order of bias of the APE plug-in estimator, induced by $\hat{\alpha}_i(\tilde{\boldsymbol{\beta}})$, from $O(T^{-1})$ to $O(T^{-2})$ by applying the analytical bias-correction of Fernández-Val (2009) reported in Eq. (6).

3.2.1 Multiple-step estimation

The first step consists in estimating by CML the structural parameters of the logit model in (1). Taking the individual intercept α_i as fixed, the joint probability of the response configuration $\boldsymbol{y}_i = (y_{i1}, \dots, y_{iT})'$ conditional on $X_i = (\boldsymbol{x}_{i1}, \dots, \boldsymbol{x}_{iT})$ can be written as:

$$p(\boldsymbol{y}_i | X_i, \alpha_i) = \frac{\exp\left(y_{i+}\alpha_i + \sum_{t=1}^{T} y_{it} \boldsymbol{x}_{it}' \boldsymbol{\beta}\right)}{\prod_{t=1}^{T} \left[1 + \exp\left(\alpha_i + \boldsymbol{x}_{it}' \boldsymbol{\beta}\right)\right]},$$

where the dependence of the probability on the left hand-side upon the slope parameters is suppressed to avoid abuse of notation. It is well known that the *total score* $y_{i+} = \sum_{t=1}^{T} y_{it}$ is a sufficient statistic for the individual intercepts α_i (Andersen 1970; Chamberlain 1980). The joint probability of $\boldsymbol{y}_i$ conditional on y_{i+} does not depend on α_i and can therefore be written as:

$$p(\boldsymbol{y}_i | X_i, y_{i+}) = \frac{\exp\left[\left(\sum_{t=1}^{T} y_{it} \boldsymbol{x}_{it}\right)' \boldsymbol{\beta}\right]}{\sum_{z:z_+=y_{i+}} \exp\left[\left(\sum_{t=1}^{T} z_t \boldsymbol{x}_{it}\right)' \boldsymbol{\beta}\right]}, \tag{7}$$

where the denominator is the sum over all the response configurations z such that $z_+ = y_{i+}$ and where the individual intercept α_i has been cancelled out. The log-likelihood function is

$$\ell(\boldsymbol{\beta}) = \sum_{i=1}^{n} \mathbf{I}(0 < y_{i+} < T) \log p(\boldsymbol{y}_i | X_i, y_{i+}),$$

where the indicator function $\mathbf{I}(\cdot)$ is included to take into account that observations with total score y_{i+} equal to 0 or T do not contribute to the log-likelihood. The above

function can be maximized with respect to $\boldsymbol{\beta}$ by a Newton–Raphson algorithm using standard results on the regular exponential family (Barndorff-Nielsen 1978), so as to obtain the CML estimator $\tilde{\boldsymbol{\beta}}$, which is $\sqrt{n}$-consistent and asymptotically normal with fixed–T (see, Andersen, 1970, and Chamberlain, 1980, for details). Therefore, if plugged into the APE formulation (2) instead of the ML estimator $\hat{\boldsymbol{\beta}}$, the E component of the bias in (4) is removed since $\tilde{\boldsymbol{\beta}} \xrightarrow{p} \boldsymbol{\beta}_0$ as $n \to \infty$.

In the next step, we obtain estimates of the individual intercepts α_i, which are not directly available as they have been cancelled out by conditioning on the total score. Our strategy is to obtain the ML estimates of α_i, denoted $\hat{\alpha}_i(\tilde{\boldsymbol{\beta}})$, for those subjects such that $0 < y_{i+} < T$, by maximizing the individual term $\sum_{t=1}^{T} \log p_{\tilde{\boldsymbol{\beta}}}(y_{it}|\boldsymbol{x}_{it}, \alpha_i)$, where $p_{\tilde{\boldsymbol{\beta}}}(y_{it}|\boldsymbol{x}_{it}, \alpha_i)$ is the probability of the logit model defined in (1) evaluated at the CML estimate, namely at $\boldsymbol{\beta} = \tilde{\boldsymbol{\beta}}$. As well as the ML estimator, the analytical and the jackknife bias correction, our proposal leads to an APE equal to zero for the subjects whose response configurations are made of only 0s and 1s, as the marginal effects are evaluated at the ML (non-finite) estimates of α_i. However, even if $\boldsymbol{\beta}$ is fixed at some $\sqrt{n}$-consistent estimate, the bias of the ML estimator of α_{i0} will still be of order $O(T^{-1})$ because $\hat{\alpha}_i(\tilde{\boldsymbol{\beta}}) \xrightarrow{p} \alpha_{i0}$ only as $T \to \infty$. Stammann et al. (2016) and Bartolucci and Pigini (2019) consider such plug-in estimator and confirm by simulation that this source of bias, although rather small for the static logit model, indeed shows up in finite samples. Moreover, Bartolucci and Pigini (2019) report that the bias is more severe for the dynamic logit model.[7] In Sect. 4, we show that correcting for the bias generated by the use of $\hat{\alpha}_i(\boldsymbol{\beta})$ instead of α_{i0}, denoted by D in (4), is necessary in finite samples, especially with short T.

In the final step, the APEs are obtained by simply replacing the ML estimators in (2) with $\tilde{\boldsymbol{\beta}}$ and $\hat{\alpha}(\tilde{\boldsymbol{\beta}})$ and reducing the bias from $O(T^{-1})$ to $O(T^{-2})$ by applying the bias correction proposed by Fernández-Val (2009), that is,

$$\tilde{\mu}_k = \frac{1}{nT} \sum_{i=1}^{n} \sum_{t=1}^{T} m_{itk}(\hat{\alpha}_i(\tilde{\boldsymbol{\beta}}), \tilde{\boldsymbol{\beta}}, \boldsymbol{x}_{it}) - \tilde{D},$$

where $\tilde{D}$ denotes the sample counterpart of (5) evaluated in $\tilde{\boldsymbol{\beta}}$ and $\hat{\alpha}(\tilde{\boldsymbol{\beta}})$. It is worth stressing that the proposed estimator exhibits the same asymptotic properties of any feasible average effect estimator under rectangular array asymptotics, as outlined in Sect. 2, since the CML estimator is also proved to be $\sqrt{nT}$-consistent.[8]

---

[7] It is worth recalling that using a bias corrected estimate of α_i, such as the one proposed by Kunz et al. (2019), along with a fixed-T consistent estimator of $\boldsymbol{\beta}$ will not reduce the order of the bias of the APE estimator to $O(T^{-2})$, as it would not take care of the last component in (5). Yet Bartolucci and Pigini (2019) show that the finite sample performance of the resulting APE estimator is superior to that of the panel jackknife with short-T, while the two estimators are comparable with moderately long panels.

[8] The result is a special case of Theorem 1 by Hanfelt and Wang (2014), who extend the results in Hahn and Newey (2004) to derive asymptotic properties for a general class of estimators based on what they call a "relaxed" conditional likelihood. The CML estimator emerges as a special case when, as in our case, responses are distributed according to the regular exponential family.

3.2.2 Standard errors

In order to derive an expression for the standard errors of the APEs $\tilde{\boldsymbol{\mu}} = (\tilde{\mu}_1, \ldots, \tilde{\mu}_K)'$, we need to account for the variability in $\boldsymbol{x}_{it}$ and the use of the estimated parameters $\tilde{\boldsymbol{\beta}}$ in the first step. For the latter, we rely on the generalized method of moments (GMM) approach by Hansen (1982) and also implemented by Bartolucci and Nigro (2012) for the quadratic exponential model. In particular, following Newey and McFadden (1994), we formulate the proposed multi-step procedure as the solution of the system of estimating equations

$$\boldsymbol{f}(\boldsymbol{\beta}, \boldsymbol{\mu}) = \boldsymbol{0},$$

where

$$\boldsymbol{f}(\boldsymbol{\beta}, \boldsymbol{\mu}) = \sum_{i=1}^{n} \boldsymbol{f}_i(\boldsymbol{\beta}, \boldsymbol{\mu}),$$

$$\boldsymbol{f}_i(\boldsymbol{\beta}, \boldsymbol{\mu}) = \begin{pmatrix} \nabla_{\boldsymbol{\beta}} \ell_i(\boldsymbol{\beta}) \\ \nabla_{\mu_1} g_i(\boldsymbol{\beta}, \mu_1) \\ \vdots \\ \nabla_{\mu_K} g_i(\boldsymbol{\beta}, \mu_K) \end{pmatrix}, \tag{8}$$

and

$$g_i(\boldsymbol{\beta}, \mu_k) = \frac{1}{T} \sum_{t=1}^{T} [m_{itk}(\alpha_i(\boldsymbol{\beta}), \boldsymbol{\beta}, \boldsymbol{x}_{it}) - \mu_k]^2, \quad k = 1, \ldots, K.$$

The asymptotic variance of $(\tilde{\boldsymbol{\beta}}', \tilde{\boldsymbol{\mu}}')'$ is then

$$\boldsymbol{W}(\tilde{\boldsymbol{\beta}}, \tilde{\boldsymbol{\mu}}) = \boldsymbol{H}(\tilde{\boldsymbol{\beta}}, \tilde{\boldsymbol{\mu}})^{-1} \boldsymbol{S}(\tilde{\boldsymbol{\beta}}, \tilde{\boldsymbol{\mu}}) [\boldsymbol{H}(\tilde{\boldsymbol{\beta}}, \tilde{\boldsymbol{\mu}})^{-1}]', \tag{9}$$

where

$$\boldsymbol{S}(\tilde{\boldsymbol{\beta}}, \tilde{\boldsymbol{\mu}}) = \sum_{i=1}^{n} \boldsymbol{f}_i(\tilde{\boldsymbol{\beta}}, \tilde{\boldsymbol{\mu}}) \boldsymbol{f}_i(\tilde{\boldsymbol{\beta}}, \tilde{\boldsymbol{\mu}})'.$$

Moreover, we have that

$$\boldsymbol{H}(\tilde{\boldsymbol{\beta}}, \tilde{\boldsymbol{\mu}}) = \sum_{i=1}^{n} \boldsymbol{H}_i(\tilde{\boldsymbol{\beta}}, \tilde{\boldsymbol{\mu}}),$$

where

$$\boldsymbol{H}_i(\boldsymbol{\beta}, \boldsymbol{\mu}) = \begin{pmatrix} \nabla_{\boldsymbol{\beta}\boldsymbol{\beta}} \ell_i(\boldsymbol{\beta}) & \boldsymbol{O} \\ \nabla_{\boldsymbol{\mu}\boldsymbol{\beta}} g_i(\boldsymbol{\beta}, \boldsymbol{\mu}) & \nabla_{\boldsymbol{\mu}\boldsymbol{\mu}} g_i(\boldsymbol{\beta}, \boldsymbol{\mu}) \end{pmatrix} \tag{10}$$

is the derivative of $\boldsymbol{f}_i(\boldsymbol{\beta}, \boldsymbol{\mu})$ with respect to $(\boldsymbol{\beta}, \boldsymbol{\mu})$, with $\boldsymbol{O}$ denoting a $K \times K$ matrix of zeros and $\boldsymbol{g}_i(\boldsymbol{\beta}, \boldsymbol{\mu})$ collects $g_i(\boldsymbol{\beta}, \mu_k)$, for $k = 1, \ldots, K$. Expressions for the derivatives in (8) are

$$\nabla_{\boldsymbol{\beta}} \ell_i(\boldsymbol{\beta}) = \sum_{t=1}^{T} y_{it} \boldsymbol{x}_{it} - \sum_{z:z_+=y_{i+}} \left(p(z|X_i, y_{i+}) \sum_{t=1}^{T} z_t \boldsymbol{x}_{it} \right),$$

and

$$\nabla_{\mu_k} g_i(\boldsymbol{\beta}, \mu_k) = -\frac{2}{T} \sum_{t=1}^{T} [m_{itk}(\alpha_i(\boldsymbol{\beta}), \boldsymbol{\beta}, \boldsymbol{x}_{it}) - \mu_k].$$

The second derivatives in (10) are

$$\nabla_{\boldsymbol{\beta}\boldsymbol{\beta}} \ell_i(\boldsymbol{\beta}) = \sum_{z:z_+=y_{i+}} p(z|X_i, y_{i+}) \boldsymbol{e}(z, X_i) \boldsymbol{e}(z, X_i)',$$

where

$$\boldsymbol{e}(z, X_i) = \sum_{t=1}^{T} z_t \boldsymbol{x}_{it} - \sum_{z:z_+=y_{i+}} \left(p(z|X_i, y_{i+}) \sum_{t=1}^{T} z_t \boldsymbol{x}_{it} \right),$$

and $\nabla_{\boldsymbol{\mu}\boldsymbol{\mu}} \boldsymbol{g}_i(\boldsymbol{\beta}, \boldsymbol{\mu})$ is a $K \times K$ diagonal matrix with elements equal to 2. Finally, for the computation of the block $\nabla_{\boldsymbol{\mu}\boldsymbol{\beta}} g_i(\boldsymbol{\beta}, \boldsymbol{\mu})$, we rely on a numerical differentiation. Once the matrix in (9) is computed, the standard errors for the APEs $\tilde{\boldsymbol{\mu}}$ may be obtained by taking the square root of the elements in the main diagonal of the lower right submatrix of $W(\tilde{\boldsymbol{\beta}}, \tilde{\boldsymbol{\mu}})$.

3.2.3 Dynamic logit model

The method proposed to obtain the APE for the logit model cannot be applied directly to the dynamic logit model (Hsiao 2005). In the latter case, the conditional probability of y_{it} is

$$p(y_{it}|\boldsymbol{x}_{it}, y_{i,t-1}; \alpha_i, \boldsymbol{\beta}, \gamma) = \frac{\exp\left[y_{it}(\alpha_i + \boldsymbol{x}'_{it}\boldsymbol{\beta} + y_{i,t-1}\gamma)\right]}{1 + \exp(\alpha_i + \boldsymbol{x}'_{it}\boldsymbol{\beta} + y_{i,t-1}\gamma)}, \tag{11}$$

where γ is the regression coefficient for the lagged response variable that measures the true state dependence. Plugging the CML estimator of $\boldsymbol{\beta}$ and γ in the APE formulation is not viable in this case because the total score is no longer a sufficient statistic for the incidental parameters if the lag of the dependent variable is included among the model covariates. Conditioning on sufficient statistics eliminates the incidental parameters only in the special case of $T = 3$ and no other explanatory variables (Chamberlain 1985). Honoré and Kyriazidou (2000) extend this approach to include explanatory

variables and the corresponding parameters can be estimated by CML on the basis of a weighted conditional log-likelihood. However, time effects cannot be included in the model specification, and the estimator's rate of convergence to the true parameter value is slower than $\sqrt{n}$.

More recently, Honoré and Weidner (2020) generalize the approach by Honoré and Kyriazidou (2000) to include any type of strictly exogenous covariate and by providing a $\sqrt{n}$-consistent generalized method of moments (GMM) estimator, based on moment conditions that are free of the incidental parameters. The viability of their approach in practice, though, has to be assessed since this estimator, as GMM estimators in general, suffers from a considerable small sample bias when built on a large number of moment conditions, which is shown to rapidly increase in the number of time occasions and in the number of covariates. A different perspective is taken by Bartolucci and Nigro (2010) who, instead of the dynamic logit, consider a QE formulation (Cox 1972) to model dynamic binary panel data, that has the advantage of admitting sufficient statistics for the individual intercepts.

Bartolucci and Nigro (2012) propose a QE model that approximates more closely the dynamic logit model, the parameters of which can easily be estimated by PCML. Under the approximating model, each y_{i+} is a sufficient statistic for the fixed effect α_i. By conditioning on the total score, the joint probability of $\mathbf{y}_i$ becomes:

$$
p^*(\mathbf{y}_i|X_i, y_{i0}, y_{i+}) = \frac{\exp\left(\sum_{t=1}^{T} y_{it}\mathbf{x}_{it}'\boldsymbol{\beta} - \sum_{t=1}^{T} \bar{q}_{it}y_{i,t-1}\gamma + y_{i*}\gamma\right)}{\sum\limits_{z:z_+=y_{i+}} \exp\left(\sum_{t=1}^{T} z_t\mathbf{x}_{it}'\boldsymbol{\beta} - \sum_{t=1}^{T} \bar{q}_{it}z_{i,t-1}\gamma + z_{i*}\gamma\right)},
$$

(12)

where $y_{i*} = \sum_{t=1}^{T} y_{i,t-1}y_{it}$ and $z_{i*} = y_{i0}z_1 + \sum_{t>1} z_{t-1}z_t$. Moreover, $\bar{q}_{it}$ is a function of given values of $\boldsymbol{\beta}$ and α_i, resulting from a first-order Taylor series expansion of the log-likelihood based on (11) around $\boldsymbol{\beta} = \bar{\boldsymbol{\beta}}$ and $\alpha_i = \bar{\alpha}_i, i = 1, \ldots, n$, and $\gamma = 0$ (see Bartolucci and Nigro 2012, for details). The expression for $\bar{q}_{it}$ is then

$$
\bar{q}_{it} = \frac{\exp(\bar{\alpha}_i + \mathbf{x}_{it}'\bar{\boldsymbol{\beta}})}{1 + \exp(\bar{\alpha}_i + \mathbf{x}_{it}'\bar{\boldsymbol{\beta}})}.
$$

Expressions for the partial effects and APEs are derived in the same way as for the static logit model. Let $\mathbf{w}_{it} = (\mathbf{x}_{it}', y_{i,t-1})'$ collect the $K + 1$ model covariates. Based on (11), the partial effect of covariate w_{itk} for i at time t on the probability of $y_{it} = 1$ can be written as:

$$
m_{itk}(\alpha_i, \boldsymbol{\theta}, \mathbf{w}_{it}) = \begin{cases} p(y_{it} = 1|\mathbf{w}_{it}; \alpha_i, \boldsymbol{\beta}, \gamma)\left[1 - p(y_{it} = 1|\mathbf{w}_{it}; \alpha_i, \boldsymbol{\beta}, \gamma)\right]\beta_k, & w_{itk} \text{ continuous,} \\[2ex] \begin{aligned} &p(y_{it} = 1|\mathbf{w}_{it,-k}, w_{itk} = 1; \alpha_i, \boldsymbol{\beta}, \gamma) - \\ &p(y_{it} = 1|\mathbf{w}_{it,-k}, w_{itk} = 0; \alpha_i, \boldsymbol{\beta}, \gamma), \end{aligned} & w_{itk} \text{ binary,} \end{cases}
$$

where $\mathbf{w}_{it,-k}$ denotes the vector $\mathbf{w}_{it}$ excluding w_{itk}, and $\boldsymbol{\theta} = (\boldsymbol{\beta}', \gamma)'$. This expression may also be used to compute the APE of the lagged response variable. Notice that this

function does not depend on $\bar{\boldsymbol{\beta}}$, since the probability in (11) does not depend on $\bar{q}_{it}$. The APE of the kth covariate can then be obtained by taking the expected value of $m_{itk}(\alpha_i, \boldsymbol{\theta}, \boldsymbol{w}_{it})$ with respect to $\boldsymbol{w}_{it}$ and α_{i0}, and can be written as

$$\mu_{k0} = \int m_{itk}(\alpha_{i0}, \boldsymbol{\theta}_0, \boldsymbol{w}_{it})dG(\alpha_{i0}, \boldsymbol{w}_{it}),$$

where $G(\alpha_{i0}, \boldsymbol{w}_{it})$ denotes the joint distribution of α_{i0} and $\boldsymbol{w}_{it}$.

As for the static logit model, the estimate of μ_{k0} is based on those of the α_i, which we obtain in the same manner as in the first step described in Sect. 3.2.1. In addition, here the CML estimation of $\boldsymbol{\theta}$ based on (12) relies on a preliminary step in order to obtain $\bar{q}_{it}$. In the first step, a preliminary estimate of $\bar{\boldsymbol{\beta}}$ is obtained by maximizing the conditional log-likelihood

$$\ell(\bar{\boldsymbol{\beta}}) = \sum_{i=1}^{n} \mathbf{I}(0 < y_{i+} < T)\ell_i(\bar{\boldsymbol{\beta}}),$$

where

$$\ell_i(\bar{\boldsymbol{\beta}}) = \log \frac{\exp\left[\left(\sum_{t=1}^{T} y_{it}\boldsymbol{x}_{it}\right)' \bar{\boldsymbol{\beta}}\right]}{\sum\limits_{z:z_+=y_{i+}} \exp\left[\left(\sum_{t=1}^{T} z_t\boldsymbol{x}_{it}\right)' \bar{\boldsymbol{\beta}}\right]},$$

which is the same conditional log-likelihood of the static logit model and may be maximized by a standard Newton–Raphson algorithm. We denote the resulting CML estimator by $\check{\boldsymbol{\beta}}$. The estimate $\check{\alpha}_i$ is then computed by maximizing the individual log-likelihood

$$\ell_i(\bar{\alpha}_i) = \sum_{t=1}^{T} \log \frac{\exp\left[y_{it}(\bar{\alpha}_i + \boldsymbol{x}_{it}'\check{\boldsymbol{\beta}})\right]}{1 + \exp(\bar{\alpha}_i + \boldsymbol{x}_{it}'\check{\boldsymbol{\beta}})},$$

where $\check{\boldsymbol{\beta}}$ is fixed. The probability $\bar{q}_{it}$ in (12) is estimated as $\check{q}_{it} = \exp(\check{\alpha}_i + \boldsymbol{x}_{it}'\check{\boldsymbol{\beta}})/\left[1 + \exp(\check{\alpha}_i + \boldsymbol{x}_{it}'\check{\boldsymbol{\beta}})\right]$.

In the second step, we estimate $\boldsymbol{\theta}$ by maximizing the conditional log-likelihood

$$\ell(\boldsymbol{\theta}) = \sum_{i=1}^{n} \mathbf{I}(0 < y_{i+} < T) \log p_{\check{\boldsymbol{q}}_i}^*(\boldsymbol{y}_i|X_i, y_{i0}, y_{i+}),$$

where $p_{\check{\boldsymbol{q}}_i}^*(\boldsymbol{y}_i|X_i, y_{i0}, y_{i+})$ is the joint probability in (12) evaluated at $\check{\boldsymbol{q}}_i = (\check{q}_{i1}, \ldots, \check{q}_{iT})'$. The above function can easily be maximized with respect to $\boldsymbol{\theta}$ by a Newton–Raphson algorithm, so as to obtain the PCML estimator $\tilde{\boldsymbol{\theta}}$, which is a $\sqrt{n}$-consistent estimator of $\boldsymbol{\theta}_0$ only if $\gamma_0 = 0$, representing the special case in which the QE

model corresponds to the dynamic logit model.[9] Nonetheless, Bartolucci and Nigro (2012) show that the PCML estimator has a limited bias in finite samples even in the presence of non-negligible state dependence.

The next step consists of recovering the estimates of α_i, $\tilde{\alpha}_i(\tilde{\boldsymbol{\theta}})$, by maximizing the concentrated log-likelihood evaluated at $\tilde{\boldsymbol{\theta}}$. In the final step, the APEs can then be estimated by plugging $\tilde{\alpha}_i(\tilde{\boldsymbol{\theta}})$ and $\tilde{\boldsymbol{\theta}}$ in the APE formulation and applying the same correction shown in Sect. 3.2.1, so as to obtain

$$\tilde{\mu}_k = \frac{1}{nT} \sum_{i=1}^{n} \sum_{t=1}^{T} m_{itk}(\hat{\alpha}_i(\tilde{\boldsymbol{\theta}}), \tilde{\boldsymbol{\theta}}, \boldsymbol{w}_{it}) - \tilde{D}.$$

Standard errors for $\tilde{\mu}_k$ can be obtained exactly in the same way as illustrated in Sect. 3.2.2 with the appropriate change of notation.

4 Monte Carlo simulation study

In the following, we illustrate the design and report the main results of the simulation studies aimed at assessing the finite sample performance of the estimators of the APEs for the static and dynamic logit models. We also discuss the results of some robustness exercises, with the related tables reported in Appendix.

4.1 Simulation set-up

We generate data for the logit model according to the formulation proposed by Honoré and Kyriazidou (2000), that is, for $i = 1, \ldots, n$,

$$y_{it} = \mathbf{I}(\alpha_i + y_{i,t-1}\gamma + x_{it}\beta + \varepsilon_{it} > 0), \quad t = 1, \ldots, T, \tag{13}$$

$$y_{i0} = \mathbf{I}(\alpha_i + x_{i0}\beta + \varepsilon_{i0} > 0), \tag{14}$$

where $x_{it} \sim N(0, \pi^2/3)$ and ε_{it} follows a standard logistic distribution for $t = 0, \ldots, T$.

Based on the above design, we generate data from the static and dynamic logit models as follows. For the static logit model, data are generated under assumption (13) with $\gamma = 0$ and $\beta = 1$, for $t = 1, \ldots, T$. Here the individual intercepts are given by $\alpha_i = \sum_{t=1}^{4} x_{it}/4$. For the dynamic logit model, data are generated for $t = 0, \ldots, T$ using also (14), γ in (13) takes values $(0.25, 0.5, 0.75)$, $\beta = 1$, and the individual heterogeneity is generated as $\alpha_i = \sum_{t=0}^{3} x_{it}/4$. We consider the same scenarios for both the static and dynamic logit model, corresponding to $n = 100, 500$ and $T = 4, 8, 12$, and the number of Monte Carlo replications is 1000.

For the static logit model, we compare the finite sample performance of the proposed APE estimator (denoted by CML-BC) with: (a) the ML plug-in estimator (ML);

---

[9] The correspondence refers to the log-odds ratio. This is clarified by Theorem 1 in Bartolucci and Nigro (2012).

(b) Hahn and Newey (2004)'s jackknife bias corrected estimator (Jackknife-BC); (c) the ML estimator with the analytical bias correction (Analytical-BC) provided by Fernández-Val (2009), also mentioned in the previous section.

For the dynamic logit model, we compare the finite sample performance of the proposed APE estimator (PCML-BC) with: (a) the ML plug-in estimator; (b) Dhaene and Jochmans (2015)'s half-panel jackknife bias-corrected estimator (Jackknife-BC); (c) the analytically bias-corrected estimator (Analytical-BC) by Fernández-Val (2009). It must be noted that the half-panel Jackknife-BC estimator cannot be computed for $T = 4$.

For each scenario, we report the mean and the median of the ratio $\tilde{\mu}/\mu^*$, the standard deviation of $\tilde{\mu}$, the rejection frequency at the 5% and 10% nominal value of a t-test for the true value of the APE, and the mean ratio between the estimator standard error and standard deviation.[10]

4.2 Main results

Table 1 reports the simulation results for the static logit model. It emerges that the proposed estimator (CML-BC) has good finite sample performance with both small n and T. On the contrary, the Jackknife-BC and the Analytical-BC exhibit a sizable bias when $T = 4$ and produce unreliable coverage intervals. Actually, the simulation results of the ML estimator, especially for $T \geq 8$, suggest that the bias correction is unnecessary in this case. Therefore, it emerges here, and also from the results reported by Fernández-Val (2009), that for the APEs of static models the sources of bias are negligible not only asymptotically, as discussed in Sect. 2, but also in finite samples.

Bias corrections are bound to be more relevant for the dynamic logit, as ML is known to produce a severely biased estimator of the state dependence parameter in auto-regressive formulations for both linear and nonlinear models (Heckman 1981; Nickell 1981). Tables 2, 3 and 4 report the results for the partial effect relative to the state dependence parameter, μ_y. The simulation results of the APEs for the covariate, denoted μ_x, are reported in Tables 8, 9 and 10 in Appendix A.1.

Results confirm that the bias of plug-in APEs based on ML estimates is greater, especially concerning the partial effect of the state dependence parameters. While all bias corrections produce a remarkable improvement over ML, the proposed estimator outperforms the Analytical-BC and Jackknife-BC. With $\gamma = 0.25$, the advantage is noticeable across all the scenarios considered. Furthermore, the proposed methodology seems to provide the most reliable confidence intervals among the examined estimators. In this regard, it is worth noticing that when $T = 4$, all the estimators provide poor coverage. As for the APE of the covariate μ_x, the PCML-BC exhibits a better performance with $T = 4$, whereas all the bias correction strategies have comparable performance with larger values of T. These results suggest that the use of a fixed-T consistent estimator, even though for an approximating model, offers an

---

[10] ML standard errors are computed for Hahn and Newey (2004)'s Jackknife-BC estimator. Bootstrapped standard errors (500 replications) are computed for Dhaene and Jochmans (2015)'s half-panel Jackknife-BC estimator.

Table 1 Simulation results for $\tilde{\mu}$, static logit model

n	T		Mean ratio	Median ratio	SD	Rejection rate		SE/SD
						5%	10%	
100	4							
		ML	0.981	0.977	0.013	0.098	0.161	0.805
		Jackknife-BC	1.231	1.220	0.022	0.407	0.499	0.806
		Analytical-BC	0.704	0.742	0.026	0.960	0.982	0.340
		CML-BC	0.911	0.909	0.012	0.108	0.193	1.132
100	8							
		ML	1.008	1.008	0.008	0.048	0.112	0.873
		Jackknife-BC	0.999	0.997	0.008	0.018	0.048	1.052
		Analytical-BC	0.983	0.982	0.008	0.055	0.104	0.898
		CML-BC	0.991	0.991	0.008	0.044	0.090	0.910
100	12							
		ML	1.006	1.005	0.006	0.054	0.119	0.853
		Jackknife-BC	0.998	0.997	0.006	0.035	0.068	0.971
		Analytical-BC	0.997	0.996	0.006	0.038	0.072	0.946
		CML-BC	1.000	0.998	0.006	0.039	0.078	0.916
500	4							
		ML	0.973	0.972	0.006	0.164	0.236	0.822
		Jackknife-BC	1.210	1.209	0.010	0.935	0.964	0.743
		Analytical-BC	0.755	0.758	0.005	1.000	1.000	0.703
		CML-BC	0.908	0.908	0.005	0.581	0.730	1.156
500	8							
		ML	1.007	1.007	0.004	0.067	0.117	0.876
		Jackknife-BC	1.001	1.000	0.004	0.029	0.061	1.033
		Analytical-BC	0.983	0.983	0.003	0.120	0.189	0.899
		CML-BC	0.991	0.990	0.004	0.067	0.122	0.917
500	12							
		ML	1.003	1.003	0.003	0.056	0.104	0.861
		Jackknife-BC	0.996	0.996	0.003	0.033	0.066	0.976
		Analytical-BC	0.995	0.994	0.003	0.039	0.086	0.952
		CML-BC	0.997	0.997	0.003	0.039	0.080	0.926

1000 replications. ML is the plug-in estimator for the static logit model; Jackknife-BC denotes Hahn and Newey (2004)'s jackknife bias correction; Analytical-BC denotes Hahn and Newey (2004)'s analytical bias correction; CML-BC denotes the proposed estimator

advantage when the bias is sizable, as removing only the $O(T^{-1})$ component may not provide enough of a reduction.

As discussed in Section 3.2.3, while the PCML estimator is consistent only when $\gamma = 0$, Bartolucci and Nigro (2012) show that its finite-sample bias is limited even in presence of non-negligible state dependence. The results in Tables 2, 3 and 4 confirm that the PCML-BC APE estimator has the same behaviour: although the advantage

Table 2 Simulation results for $\tilde{\mu}_y$, dynamic logit model, $\gamma = 0.25$

n	T		Mean ratio	Median ratio	SD	Rejection rate		SE/SD
						5%	10%	
100	4							
		ML	−3.143	−3.185	0.039	0.963	0.979	0.895
		Analytical-BC	1.884	1.768	0.050	0.152	0.227	0.771
		PCML-BC	0.954	0.888	0.052	0.018	0.047	1.234
100	8							
		ML	−1.260	−1.231	0.029	0.845	0.903	0.945
		Jackknife-BC	0.612	0.631	0.036	0.077	0.134	1.006
		Analytical-BC	0.924	0.963	0.030	0.093	0.135	0.896
		PCML-BC	1.012	1.058	0.033	0.064	0.113	0.978
100	12							
		ML	−0.594	−0.605	0.024	0.691	0.786	0.930
		Jackknife-BC	0.800	0.812	0.029	0.091	0.142	0.946
		Analytical-BC	0.900	0.886	0.025	0.077	0.147	0.900
		PCML-BC	0.931	0.915	0.027	0.071	0.131	0.921
500	4							
		ML	−3.199	−3.193	0.018	1.000	1.000	0.867
		Analytical-BC	1.557	1.549	0.019	0.254	0.335	0.870
		PCML-BC	0.902	0.896	0.024	0.021	0.049	1.208
500	8							
		ML	−1.294	−1.291	0.013	1.000	1.000	0.950
		Jackknife-BC	0.602	0.609	0.016	0.154	0.236	1.019
		Analytical-BC	0.903	0.899	0.014	0.081	0.147	0.904
		PCML-BC	0.979	0.977	0.015	0.053	0.096	0.994
500	12							
		ML	−0.528	−0.528	0.011	0.999	1.000	0.954
		Jackknife-BC	0.881	0.881	0.013	0.086	0.132	0.987
		Analytical-BC	0.970	0.972	0.011	0.076	0.130	0.925
		PCML-BC	1.004	1.006	0.012	0.073	0.113	0.950

1000 replications. ML is the plug-in estimator for the dynamic logit model; Jackknife-BC denotes Dhaene and Jochmans (2015)'s half-panel jackknife bias corrected estimator; Analytical-BC denotes Fernández-Val (2009)'s analytical bias correction; PCML-BC denotes the proposed estimator

over the alternative approaches is reduced with $\gamma = 0.5$ and $\gamma = 0.75$, the proposed approach remains the most effective strategy to correct APEs for the state dependence parameter, especially with short T.

It is finally worth recalling that further to plugging fixed-T (pseudo)consistent estimators for the slope parameters in the APEs, the good finite-sample performance of the proposed approach also benefits form the correction aimed to remove the bias originating from the ML estimate of the unobserved heterogeneity, denoted by D in (4). In order to illustrate the indirect bias effect introduced by $\hat{\alpha}_i(\boldsymbol{\beta})$, we report in

Table 3 Simulation results for $\tilde{\mu}_y$, dynamic logit model, $\gamma = 0.5$

n	T		Mean ratio	Median ratio	SD	Rejection rate		SE/SD
						5%	10%	
100	4							
		ML	-1.221	-1.264	0.041	0.975	0.986	0.840
		Analytical-BC	1.167	1.103	0.048	0.102	0.166	0.783
		PCML-BC	0.883	0.831	0.054	0.025	0.047	1.231
100	8							
		ML	-0.210	-0.218	0.030	0.861	0.910	0.919
		Jackknife-BC	0.780	0.760	0.040	0.093	0.157	0.950
		Analytical-BC	0.889	0.883	0.032	0.096	0.175	0.863
		PCML-BC	0.982	0.978	0.035	0.076	0.126	0.943
100	12							
		ML	0.177	0.172	0.024	0.728	0.825	0.960
		Jackknife-BC	0.916	0.913	0.028	0.061	0.118	1.007
		Analytical-BC	0.939	0.931	0.025	0.068	0.129	0.931
		PCML-BC	0.987	0.981	0.026	0.066	0.125	0.943
500	4							
		ML	-1.224	-1.218	0.018	1.000	1.000	0.867
		Analytical-BC	1.027	1.033	0.018	0.082	0.125	0.905
		PCML-BC	0.873	0.865	0.025	0.026	0.064	1.200
500	8							
		ML	-0.204	-0.206	0.013	1.000	1.000	0.961
		Jackknife-BC	0.800	0.795	0.017	0.140	0.230	0.997
		Analytical-BC	0.903	0.903	0.014	0.117	0.182	0.908
		PCML-BC	0.994	0.993	0.015	0.052	0.105	0.993
500	12							
		ML	0.182	0.177	0.010	1.000	1.000	0.990
		Jackknife-BC	0.923	0.921	0.012	0.071	0.132	1.027
		Analytical-BC	0.943	0.940	0.011	0.079	0.148	0.957
		PCML-BC	0.991	0.991	0.011	0.059	0.105	0.983

1000 replications. ML is the plug-in estimator for the dynamic logit model; Jackknife-BC denotes Dhaene and Jochmans (2015)'s half-panel jackknife bias-corrected estimator; Analytical-BC denotes Fernández-Val (2009)'s analytical bias correction; PCML-BC denotes the proposed estimator

Table 5 the results of a simulation exercise where we compare the performance of the proposed approach (P)CML-BC with that of an APE estimator based on (P)CML estimates ignoring the bias correction. The latter approach is employed by Stammann et al. (2016) and was considered on an earlier version of this work by Bartolucci and Pigini (2019). It clearly emerges that a bias reduction is needed for both the static and dynamic logit model, especially with $T = 4$. Besides, these results also highlight that plugging fixed-T consistent estimators already brings a sizable improvement over the ML-based APEs.

Table 4 Simulation results for $\tilde{\mu}_y$, dynamic logit model, $\gamma = 0.75$

n	T		Mean ratio	Median ratio	SD	Rejection rate		SE/SD
						5%	10%	
100	4							
		ML	−0.561	−0.583	0.040	0.984	0.992	0.865
		Analytical-BC	0.949	0.928	0.069	0.100	0.160	0.546
		PCML-BC	0.877	0.865	0.054	0.028	0.042	1.263
100	8							
		ML	0.130	0.133	0.029	0.906	0.938	0.937
		Jackknife-BC	0.823	0.818	0.038	0.105	0.159	0.995
		Analytical-BC	0.867	0.867	0.031	0.116	0.196	0.886
		PCML-BC	0.967	0.961	0.034	0.064	0.126	0.977
100	12							
		ML	0.434	0.431	0.024	0.753	0.827	0.964
		Jackknife-BC	0.960	0.958	0.029	0.064	0.116	0.985
		Analytical-BC	0.950	0.944	0.025	0.068	0.131	0.930
		PCML-BC	1.008	1.004	0.027	0.068	0.124	0.945
500	4							
		ML	−0.565	−0.559	0.018	1.000	1.000	0.878
		Analytical-BC	0.855	0.857	0.018	0.180	0.250	0.917
		PCML-BC	0.861	0.860	0.024	0.032	0.068	1.302
500	8							
		ML	0.139	0.141	0.013	1.000	1.000	0.954
		Jackknife-BC	0.843	0.838	0.017	0.171	0.270	1.005
		Analytical-BC	0.882	0.885	0.014	0.206	0.302	0.903
		PCML-BC	0.979	0.979	0.015	0.052	0.098	0.996
500	12							
		ML	0.423	0.427	0.011	1.000	1.000	0.964
		Jackknife-BC	0.944	0.943	0.013	0.073	0.134	1.005
		Analytical-BC	0.937	0.941	0.011	0.123	0.189	0.938
		PCML-BC	0.994	0.998	0.012	0.063	0.116	0.962

1000 replications. ML is the plug-in estimator for the dynamic logit model; Jackknife-BC denotes Dhaene and Jochmans (2015)'s half-panel jackknife bias-corrected estimator; Analytical-BC denotes Fernández-Val (2009)'s analytical bias correction; PCML-BC denotes the proposed estimator

4.3 Robustness exercises

In this section, we illustrate the design and discuss the results of a series of robustness exercises aimed to test the performance of the proposed approach under potentially problematic departures from the baseline setting. The results are collected in Appendix A.2.

The first set of exercises concerns the violation of standard hypotheses that are necessary for the asymptotic results on the bias correction of the ML estimator to

Table 5 Simulation results for the static and dynamic logit, comparison between (P)CML and (P)CML-BC

n	T		Static logit $\tilde{\mu}$		Dynamic logit, $\gamma = 0.5$ $\tilde{\mu}_x$		$\tilde{\mu}_y$	
			Mean ratio	Median ratio	Mean ratio	Median ratio	Mean ratio	Median ratio
100	4	(P)CML	0.805	0.801	0.786	0.786	0.780	0.782
		(P)CML-BC	0.908	0.909	0.887	0.885	0.881	0.876
100	8	(P)CML	0.918	0.917	0.906	0.904	0.905	0.899
		(P)CML-BC	0.991	0.990	0.986	0.985	0.982	0.976
500	4	(P)CML	0.804	0.803	0.781	0.782	0.785	0.799
		(P)CML-BC	0.911	0.910	0.889	0.889	0.885	0.894
500	8	(P)CML	0.917	0.918	0.905	0.905	0.907	0.910
		(P)CML-BC	0.990	0.991	0.985	0.985	0.985	0.985

1000 replications. (P)CML denotes the CML-based APE for the static and dynamic logit models without bias correction. (P)CML-BC denotes the proposed estimator for the static and dynamic logit models

hold. In fact, it is worth recalling the characterization of this correction that is based on a large-T asymptotic expansion, which theoretically rules out the possibility of including time dummies in the model specification as, with large-T asymptotics, time fixed-effects are incidental parameters as well.[11] We devise an experiment where data are generated from the dynamic logit model defined in (13)-(14), with the addition of a trending regressor $\eta_t = -1 + \frac{2(t-1)}{T-1}$ in $[-1, 1]$ and the model specification then includes time effects as well. The results for $T = 4, 8$ are reported in Table 11 and suggest overall robustness of the approaches considered to the inclusion of time dummies, as also documented by Fernández-Val (2009).

Large-T asymptotics also requires stationarity of covariates, which may be rather restrictive in applications. Yet both Fernández-Val (2009) and Dhaene and Jochmans (2015) report similar performance of the approaches under departures from this assumption. We confirm the same behaviour for the proposed approach by means of a simulation exercise where data are generated from a logit model with a trending regressor. The design is based on the one adopted by Hahn and Newey (2004), where

$$y_{it} = \mathbf{I}(\alpha_i + x_{it}\beta + \varepsilon_{it} > 0), \quad i = 1, \ldots, n, \ t = 1, \ldots, T,$$

with $\beta = 1$, $\alpha_i \sim N(0, 1)$, the error terms ε_{it} follow a standard logistic distribution, and

$$x_{it} = t/10 + x_{i,t-1}/2 + u_{it},$$

with $u_{it} \sim U[-0.5, 0.5]$ and $x_{i0} = u_{i0}$. According to the results reported in Table 12, there are no remarkable differences with respect to the baseline scenario.[12]

The second set of exercises explores the finite-sample performance of the proposed approach with different settings for the unobserved heterogeneity. We first consider the design based on assumptions (13)-(14) with individual intercepts generated as $\alpha_i = \sum_{t=0}^{3} x_{it}/4 + (u_i - 1)$, with $u_i \sim \chi_1^2$. The corresponding results are reported in Table 13 for the scenarios with $T = 4, 8$. As expected with fixed-effects approaches, results are unaffected by the distribution of the unobserved heterogeneity.

We then consider a shift in the distribution of the individual effects and generate them as $\alpha_i = \sum_{t=0}^{3} x_{it}/4 - 3.5$. In this way, we generate samples where the frequency of 1s amounts to about 12%, whereas in the baseline setting was around 53%, thus simulating a response variable describing rare events, which is often realistic in applications. Results are reported in Table 14 and, as expected, document a deterioration in

---

[11] Fernández-Val and Weidner (2016) provide an analytical formulation and a jackknife procedure to accommodate this case. Their simulation setup and results, however, suggest that their method is more suitable for dyadic network data, or in any case panel datasets where the dimensions of n and T are comparable, which is why their approach is not considered in our analysis, more focused on the large-n short-T framework. The conditional inference approach exploiting sufficient statistics for both individual and time effects is instead considered by Charbonneau (2017), Jochmans (2018), and Bartolucci et al. (2021).

[12] Actually, the Honoré and Kyriazidou's design also violates the stationarity assumption, in that the initial observations (y_{i0}, x_{i0}) are not drawn from their state distribution and the relationship between x and the unobserved heterogeneity changes drastically after the fourth time occasion. The only stationary setting should therefore be that of the static logit model with $T = 4$. Nevertheless, the results in Table 12 with $T = 4$ do not stress any deterioration in the performance of the approaches considered.

the performance of all the approaches considered. This is due to the limited availability of useful response configurations, which generates a larger small sample bias in the parameter estimates and increases the number of instances in which the PE must be set to zero.

5 Empirical application

We apply the proposed formulation to the problem of estimating the labour supply of married women. The same empirical application is considered by Fernández-Val (2009) and Dhaene and Jochmans (2015), after the seminal work of Hyslop (1999). The sample is drawn from the Panel Study of Income Dynamics (PSID), which consists of $n = 1,908$ married women between 19 and 59 years of age in 1980, followed for $T = 6$ time occasions, from 1980 to 1985, further to an additional observation in 1979 exploited as initial condition in dynamic models. We specify a static logit model for the probability of being employed at time t, conditional on the number of children of a certain age in the family, namely the number of kids between 0 and 2 years old, between 3 and 5, and between 6 and 17, and on the husband's income. We also specify a dynamic logit model, that is, we include lagged participation in the set of model covariates.

The estimation results for the static logit model are reported in Table 6, which shows the ML, Hahn and Newey (2004)'s panel Jackknife-BC, Hahn and Newey (2004)'s Analytical-BC, and CML estimates of the model parameters. The CML, Analytical-BC, and Jackknife-BC estimates of the parameters are all similar to each other and smaller (in absolute value) than the uncorrected ML ones. These suggest a negative effect on labour participation of having children younger than 17 in the household as well as of the level of the husband's income. The estimated APEs obtained with the proposed method suggest that having an additional child between 0 and 2 reduces the probability of working by 8.9 percentage points, and having a child between 3 and 5 years old reduces the employment probability by 6.1 percentage points. The APE estimates obtained with the Analytical-BC and Jackknife-BC estimators point toward the same results, with the exception of having children between 6 and 17 years old, which appear to be not statistically significant, according to our procedure.

Table 7 reports the results for the dynamic logit specification. Here, we report the ML, Dhaene and Jochmans (2015)'s half-panel Jackknife-BC, Fernández-Val (2009)'s Analytical-BC, and PCML estimates of the model parameters. The effect of the exogenous model covariates is now smaller and all the APE estimates suggest a negative and statistically significant effect of having children between 0 and 5 years old in the household.

The PCML estimator detects a strong state dependence in the labour force participation of married women, as the estimated coefficient for lagged participation amounts to 1.706. In terms of APE, this is translated into an increase of 15.7 percentage points in the probability of being employed at time t for a woman who was working in $t-1$, with respect to a woman who was not working in $t-1$.

Table 6 Female labour force participation: static logit model

Labour Force Participation	Model parameters β			
	ML	Jackknife BC	Analytical BC	CML
# Children 0–2	−1.331***	−1.051***	−1.090***	−1.092***
	(0.145)	(0.145)	(0.117)	(0.109)
# Children 3–5	−0.922***	−0.706***	−0.755***	−0.756***
	(0.147)	(0.147)	(0.112)	(0.103)
# Children 6–17	−0.193	−0.141	−0.157*	−0.157*
	(0.123)	(0.123)	(0.088)	(0.081)
Husband income	−0.011*	−0.005	−0.009**	−0.009***
	(0.006)	(0.006)	(0.004)	(0.004)
	Average partial effects μ			
	ML	Jackknife BC	Analytical BC	CML-BC
# Children 0–2	−0.091***	−0.092***	−0.088***	−0.089***
	(0.009)	(0.009)	(0.008)	(0.021)
# Children 3–5	−0.063***	−0.066***	−0.061***	−0.061***
	(0.009)	(0.010)	(0.008)	(0.022)
# Children 6–17	−0.013	−0.012	−0.013**	−0.013
	(0.009)	(0.009)	(0.006)	(0.019)
Husband income	−0.001	−0.001	−0.001**	−0.001
	(0.001)	(0.001)	(0.000)	(0.001)

Standard errors in round brackets. ***$p < 0.01$; **$p < 0.05$; *$p < 0.10$. ML denotes the maximum likelihood estimator; Jackknife-BC denotes Hahn and Newey (2004)'s jackknife bias correction; Analytical-BC denotes Hahn and Newey (2004)'s analytical bias correction; CML denotes the conditional maximum likelihood estimator; CML-BC denotes the proposed estimator for the APE. Source: PSID 1980-1985

6 Conclusion

We develop a multiple-step procedure to compute APEs for fixed-effects logit models that are estimated by CML. Our strategy amounts to building a plug-in APE estimator based on the fixed-T consistent CML estimator of the slope parameters and bias-corrected estimates of APEs.

The proposed estimator is asymptotically equivalent to the plug-in ML and alternative bias-corrected APE estimators, and it exhibits comparable finite sample performance when the static logit model is considered. On the contrary, the proposed approach for the dynamic logit model has a remarkable advantage in finite samples. In this respect, the multiple-step procedure here developed could be particularly useful for practitioners who often deal with short-T datasets, such as rotated surveys, and/or highly unbalanced panels.

Table 7 Female labour force participation: dynamic logit model

Labour force participation	Model parameters θ			
	ML	Jackknife BC	Analytical BC	PCML
# Children 0–2	−1.269***	−0.895***	−0.930***	−0.912***
	(0.141)	(0.246)	(0.122)	(0.095)
# Children 3–5	−0.823***	−0.503**	−0.532***	−0.503***
	(0.141)	(0.250)	(0.117)	(0.091)
# Children 6–17	−0.173	−0.019	−0.106	−0.092
	(0.117)	(0.225)	(0.092)	(0.074)
Husband income	−0.011*	−0.005	−0.009**	−0.008**
	(0.006)	(0.010)	(0.004)	(0.004)
Lagged Participation	0.569***	2.107***	1.319***	1.706***
	(0.081)	(0.119)	(0.082)	(0.103)
	Average partial effects μ			
	ML	Jackknife BC	Analytical BC	PCML-BC
# Children 0–2	−0.086***	−0.097***	−0.075***	−0.071***
	(0.009)	(0.014)	(0.008)	(0.016)
# Children 3–5	−0.056***	−0.059***	−0.043***	−0.039***
	(0.009)	(0.014)	(0.008)	(0.015)
# Children 6–17	−0.012	−0.009	−0.009	−0.007
	(0.008)	(0.013)	(0.006)	(0.012)
Husband income	−0.001	−0.001	−0.001**	−0.001
	(0.001)	(0.001)	(0.001)	(0.001)
Lagged Participation	0.041***	0.124***	0.121***	0.157***
	(0.008)	(0.010)	(0.006)	(0.021)

Standard errors in round brackets. ***$p < 0.01$; **$p < 0.05$; *$p < 0.10$. ML denotes the maximum likelihood estimator; Jackknife-BC denotes Dhaene and Jochmans (2015)'s jackknife bias correction; Analytical-BC denotes Fernández-Val (2009)'s analytical bias correction; PCML denotes the pseudo conditional maximum likelihood estimator; PCML-BC denotes the proposed estimator for the APE. Source: PSID 1979–1985

Data and computer code availability The datasets of this paper (1. code and programs, 2. data, 3. detailed readme files) are available in the GitHub repository https://github.com/fravale/replication_ape_logit.

Declarations

Appendix A

A.1 Simulation results: additional tables

See Tables 8, 9 and 10.

Table 8 Simulation results for $\tilde{\mu}_x$, dynamic logit model, $\gamma = 0.25$

n	T		Mean ratio	Median ratio	SD	Rejection rate		SE/SD
						5%	10%	
100	4							
		ML	0.922	0.914	0.013	0.206	0.287	0.849
		Analytical-BC	0.633	0.693	0.036	0.975	0.989	0.276
		PCML-BC	0.891	0.888	0.012	0.164	0.268	1.118
100	8							
		ML	0.996	0.995	0.008	0.039	0.087	0.959
		Jackknife-BC	1.057	1.057	0.011	0.077	0.146	1.051
		Analytical-BC	0.987	0.985	0.008	0.052	0.104	0.892
		PCML-BC	0.990	0.989	0.008	0.045	0.093	0.914
100	12							
		ML	0.998	0.997	0.006	0.020	0.052	1.020
		Jackknife-BC	1.009	1.008	0.007	0.041	0.092	1.003
		Analytical-BC	0.997	0.996	0.006	0.030	0.067	0.957
		PCML-BC	0.997	0.996	0.006	0.038	0.074	0.943
500	4							
		ML	0.922	0.921	0.006	0.575	0.673	0.826
		Analytical-BC	0.707	0.713	0.007	1.000	1.000	0.592
		PCML-BC	0.894	0.893	0.005	0.701	0.817	1.090
500	8							
		ML	0.991	0.991	0.004	0.059	0.113	0.936
		Jackknife-BC	1.055	1.054	0.005	0.333	0.457	0.991
		Analytical-BC	0.983	0.982	0.004	0.126	0.202	0.872
		PCML-BC	0.986	0.987	0.004	0.092	0.150	0.900
500	12							
		ML	0.996	0.998	0.003	0.027	0.069	0.985
		Jackknife-BC	1.009	1.010	0.003	0.048	0.094	0.983
		Analytical-BC	0.996	0.997	0.003	0.040	0.082	0.933
		PCML-BC	0.996	0.997	0.003	0.042	0.094	0.917

1000 replications. ML is the plug-in estimator for the dynamic logit model; Jackknife-BC denotes Dhaene and Jochmans (2015)'s half-panel jackknife bias-corrected estimator; Analytical-BC denotes Fernández-Val (2009)'s analytical bias correction; PCML-BC denotes the proposed estimator

Table 9 Simulation results for $\tilde{\mu}_x$, dynamic logit model, $\gamma = 0.5$

n	T		Mean ratio	Median ratio	SD	Rejection rate		SE/SD
						5%	10%	
100	4							
		ML	0.916	0.914	0.014	0.228	0.319	0.835
		Analytical-BC	0.640	0.696	0.034	0.967	0.983	0.286
		PCML-BC	0.885	0.882	0.012	0.153	0.252	1.125
100	8							
		ML	0.990	0.990	0.008	0.050	0.096	0.926
		Jackknife-BC	1.056	1.056	0.012	0.089	0.157	0.996
		Analytical-BC	0.982	0.983	0.008	0.067	0.118	0.861
		PCML-BC	0.986	0.987	0.008	0.054	0.103	0.891
100	12							
		ML	0.994	0.994	0.006	0.021	0.062	1.006
		Jackknife-BC	1.008	1.009	0.007	0.039	0.070	1.003
		Analytical-BC	0.994	0.994	0.006	0.027	0.077	0.948
		PCML-BC	0.993	0.993	0.006	0.028	0.083	0.930
500	4							
		ML	0.911	0.909	0.006	0.669	0.769	0.852
		Analytical-BC	0.707	0.712	0.007	1.000	1.000	0.574
		PCML-BC	0.885	0.882	0.005	0.747	0.846	1.169
500	8							
		ML	0.989	0.989	0.004	0.071	0.115	0.956
		Jackknife-BC	1.059	1.059	0.005	0.366	0.506	1.028
		Analytical-BC	0.981	0.981	0.004	0.134	0.210	0.882
		PCML-BC	0.985	0.985	0.004	0.097	0.169	0.919
500	12							
		ML	0.996	0.998	0.003	0.037	0.082	0.974
		Jackknife-BC	1.013	1.013	0.003	0.075	0.141	0.990
		Analytical-BC	0.996	0.997	0.003	0.044	0.093	0.923
		PCML-BC	0.996	0.997	0.003	0.050	0.103	0.908

1000 replications. ML is the plug-in estimator for the dynamic logit model; Jackknife-BC denotes Dhaene and Jochmans (2015)'s half-panel jackknife bias-corrected estimator; Analytical-BC denotes Fernández-Val (2009)'s analytical bias correction; PCML-BC denotes the proposed estimator

Table 10 Simulation results for $\tilde{\mu}_x$, dynamic logit model, $\gamma = 0.75$

n	T		Mean ratio	Median ratio	SD	Rejection rate		SE/SD
						5%	10%	
100	4							
		ML	0.907	0.901	0.013	0.247	0.343	0.843
		Analytical-BC	0.634	0.692	0.035	0.976	0.987	0.284
		PCML-BC	0.878	0.875	0.012	0.154	0.258	1.169
100	8							
		ML	0.989	0.989	0.008	0.048	0.097	0.929
		Jackknife-BC	1.063	1.061	0.012	0.081	0.155	1.021
		Analytical-BC	0.980	0.981	0.008	0.071	0.132	0.867
		PCML-BC	0.984	0.984	0.008	0.057	0.122	0.901
100	12							
		ML	0.994	0.994	0.006	0.030	0.061	0.999
		Jackknife-BC	1.013	1.014	0.007	0.034	0.077	1.004
		Analytical-BC	0.993	0.995	0.006	0.036	0.078	0.939
		PCML-BC	0.993	0.994	0.006	0.046	0.083	0.924
500	4							
		ML	0.904	0.904	0.006	0.710	0.810	0.824
		Analytical-BC	0.707	0.712	0.007	1.000	1.000	0.575
		PCML-BC	0.879	0.880	0.006	0.730	0.851	1.129
500	8							
		ML	0.985	0.985	0.004	0.087	0.158	0.893
		Jackknife-BC	1.063	1.062	0.005	0.383	0.515	0.975
		Analytical-BC	0.978	0.978	0.004	0.180	0.267	0.833
		PCML-BC	0.981	0.981	0.004	0.124	0.214	0.866
500	12							
		ML	0.995	0.995	0.003	0.034	0.072	0.981
		Jackknife-BC	1.015	1.015	0.003	0.074	0.147	0.985
		Analytical-BC	0.995	0.995	0.003	0.044	0.089	0.926
		PCML-BC	0.994	0.994	0.003	0.050	0.095	0.915

1000 replications. ML is the plug-in estimator for the dynamic logit model; Jackknife-BC denotes Dhaene and Jochmans (2015)'s half-panel jackknife bias-corrected estimator; Analytical-BC denotes Fernández-Val (2009)'s analytical bias correction; PCML-BC denotes the proposed estimator

A.2 Simulation results: robustness exercises

See Tables 11, 12, 13 and 14.

Table 11 Simulation results for the dynamic logit model, $\gamma = 0.5$, time dummies

n	T		Mean ratio	Median ratio	SD	Rejection rate		SE/SD
						5%	10%	
		μ_y						
100	4							
		ML	−1.385	−1.406	0.041	0.981	0.992	0.857
		Analytical-BC	1.860	1.725	0.068	0.250	0.335	0.690
		PCML-BC	0.867	0.834	0.054	0.028	0.060	1.205
100	8							
		ML	−0.276	−0.284	0.028	0.888	0.929	0.951
		Analytical-BC	0.868	0.871	0.030	0.091	0.153	0.913
		PCML-BC	0.976	0.991	0.033	0.002	0.005	1.755
500	4							
		ML	−1.372	−1.368	0.018	1.000	1.000	0.881
		Analytical-BC	1.476	1.443	0.024	0.422	0.518	0.741
		PCML-BC	0.863	0.855	0.024	0.035	0.060	1.215
500	8							
		ML	−0.264	−0.265	0.012	1.000	1.000	0.972
		Analytical-BC	0.891	0.892	0.013	0.118	0.187	0.928
		PCML-BC	0.992	0.989	0.015	0.001	0.004	1.783
		μ_x						
100	4							
		ML	0.896	0.899	0.014	0.255	0.336	0.854
		Analytical-BC	0.348	0.518	0.060	0.978	0.986	0.281
		PCML-BC	0.862	0.861	0.013	0.138	0.255	1.186
100	8							
		ML	0.995	0.994	0.009	0.049	0.094	0.922
		Analytical-BC	0.984	0.984	0.008	0.068	0.136	0.845
		PCML-BC	0.990	0.991	0.009	0.054	0.098	0.905
500	4							
		ML	0.893	0.892	0.006	0.756	0.829	0.855
		Analytical-BC	0.573	0.598	0.015	1.000	1.000	0.298
		PCML-BC	0.868	0.867	0.005	0.780	0.880	1.281
500	8							
		ML	0.990	0.990	0.004	0.061	0.122	0.907
		Analytical-BC	0.982	0.981	0.004	0.143	0.226	0.845
		PCML-BC	0.986	0.986	0.004	0.080	0.144	0.905

1000 replications. ML is the plug-in estimator for the dynamic logit model; analytical-BC denotes Fernández-Val (2009)'s analytical bias correction; CML-BC denotes the proposed estimator. With $n = 100$ and $T = 4$, we removed 3 Monte Carlo draws for Analytical-BC as convergence failed

Table 12 Simulation results for the static logit model ($\tilde{\mu}$), trending regressor

n	T		Mean ratio	Median ratio	SD	Rejection rate		SE/SD
						5%	10%	
100	4							
		ML	0.988	0.993	0.072	0.048	0.102	1.006
		Jackknife-BC	0.988	0.990	0.077	0.050	0.090	1.001
		Analytical-BC	0.926	0.932	0.067	0.071	0.132	0.938
		CML-BC	0.929	0.934	0.068	0.025	0.071	1.126
100	8							
		ML	0.998	0.999	0.028	0.046	0.101	0.985
		Jackknife-BC	0.999	1.000	0.028	0.036	0.090	1.026
		Analytical-BC	0.984	0.985	0.028	0.046	0.100	0.973
		CML-BC	0.985	0.986	0.028	0.039	0.086	1.025
100	12							
		ML	0.992	0.996	0.016	0.066	0.127	0.906
		Jackknife-BC	0.996	0.999	0.016	0.056	0.110	0.936
		Analytical-BC	0.988	0.992	0.016	0.045	0.106	0.965
		CML-BC	0.989	0.993	0.016	0.042	0.101	0.986
500	4							
		ML	1.004	0.997	0.032	0.056	0.094	1.004
		Jackknife-BC	1.010	1.004	0.035	0.056	0.113	0.994
		Analytical-BC	0.940	0.934	0.030	0.091	0.154	0.931
		CML-BC	0.943	0.936	0.031	0.030	0.087	1.120
500	8							
		ML	0.996	0.996	0.013	0.066	0.122	0.929
		Jackknife-BC	0.997	0.997	0.014	0.059	0.102	0.960
		Analytical-BC	0.982	0.982	0.013	0.080	0.139	0.909
		CML-BC	0.983	0.983	0.013	0.059	0.124	0.966
500	12							
		ML	0.999	1.000	0.007	0.075	0.126	0.910
		Jackknife-BC	1.003	1.002	0.007	0.066	0.113	0.946
		Analytical-BC	0.995	0.996	0.007	0.053	0.114	0.964
		CML-BC	0.996	0.996	0.007	0.042	0.104	0.990

1000 replications. ML is the plug-in estimator for the static logit model; Jackknife-BC denotes Hahn and Newey (2004)'s jackknife bias correction; Analytical-BC denotes Hahn and Newey (2004)'s analytical bias correction; CML-BC denotes the proposed estimator

Table 13 Simulation results for the dynamic logit model, $\gamma = 0.5$, non-normal α_i

n	T		Mean ratio	Median ratio	SD	Rejection rate		SE/SD
						5%	10%	
		μ_y						
100	4							
		ML	−1.242	−1.216	0.039	0.962	0.986	0.850
		Analytical-BC	1.172	1.144	0.044	0.105	0.162	0.816
		PCML-BC	0.905	0.914	0.054	0.017	0.045	1.276
100	8							
		ML	−0.237	−0.248	0.027	0.866	0.919	0.976
		Jackknife-BC	0.731	0.726	0.035	0.080	0.146	1.034
		Analytical-BC	0.869	0.852	0.029	0.092	0.158	0.916
		PCML-BC	0.974	0.960	0.032	0.041	0.087	1.087
500	4							
		ML	−1.229	−1.231	0.017	1.000	1.000	0.887
		Analytical-BC	1.023	1.026	0.017	0.079	0.137	0.903
		PCML-BC	0.867	0.864	0.023	0.013	0.032	1.335
500	8							
		ML	−0.242	−0.248	0.013	1.000	1.000	0.930
		Jackknife-BC	0.752	0.744	0.016	0.191	0.284	0.988
		Analytical-BC	0.873	0.864	0.014	0.148	0.222	0.869
		PCML-BC	0.976	0.966	0.015	0.049	0.096	1.035
		μ_x						
100	4							
		ML	0.910	0.909	0.013	0.232	0.317	0.861
		Analytical-BC	0.630	0.689	0.032	0.959	0.976	0.300
		PCML-BC	0.883	0.882	0.012	0.118	0.208	1.231
100	8							
		ML	0.987	0.986	0.009	0.043	0.085	0.903
		Jackknife-BC	1.060	1.059	0.011	0.059	0.107	1.036
		Analytical-BC	0.977	0.976	0.008	0.073	0.110	0.839
		PCML-BC	0.982	0.982	0.009	0.048	0.090	0.900
500	4							
		ML	0.905	0.905	0.006	0.685	0.769	0.858
		Analytical-BC	0.697	0.703	0.007	1.000	1.000	0.576
		PCML-BC	0.880	0.881	0.005	0.662	0.809	1.244
500	8							
		ML	0.987	0.987	0.004	0.059	0.107	0.915
		Jackknife-BC	1.064	1.064	0.005	0.335	0.482	1.023
		Analytical-BC	0.978	0.977	0.004	0.129	0.226	0.854
		PCML-BC	0.982	0.981	0.004	0.075	0.132	0.925

1000 replications. ML is the plug-in estimator for the dynamic logit model; Jackknife-BC denotes Dhaene and Jochmans (2015)'s jackknife bias correction; Analytical-BC denotes Fernández-Val (2009)'s analytical bias correction; PCML-BC denotes the proposed estimator

Table 14 Simulation results for the dynamic logit model, $\gamma = 0.5$, rare events

n	T		Mean ratio	Median ratio	SD	Rejection rate		SE/SD
						5%	10%	
		μ_y						
100	4							
		ML	−1.447	−1.452	0.028	0.914	0.945	0.837
		Analytical-BC	2.072	1.228	0.466	0.171	0.245	0.085
		PCML-BC	0.727	0.623	0.044	0.001	0.003	2.392
100	8							
		ML	−0.565	−0.591	0.021	0.768	0.840	0.940
		Jackknife-BC	0.267	0.196	0.029	0.190	0.268	1.006
		Analytical-BC	0.617	0.593	0.025	0.187	0.271	0.786
		PCML-BC	0.954	0.932	0.030	0.002	0.008	1.727
500	4							
		ML	−1.421	−1.430	0.013	1.000	1.000	0.860
		Analytical-BC	1.026	0.991	0.018	0.180	0.252	0.621
		PCML-BC	0.709	0.686	0.019	0.000	0.000	2.581
500	8							
		ML	−0.594	−0.593	0.010	1.000	1.000	0.903
		Jackknife-BC	0.273	0.275	0.013	0.510	0.634	0.968
		Analytical-BC	0.600	0.603	0.012	0.401	0.492	0.757
		PCML-BC	0.910	0.908	0.014	0.002	0.015	1.658
		μ_x						
100	4							
		ML	0.822	0.810	0.010	0.296	0.394	0.924
		Analytical-BC	0.332	0.503	0.034	0.905	0.925	0.371
		PCML-BC	0.795	0.793	0.010	0.001	0.004	2.434
100	8							
		ML	0.949	0.949	0.008	0.063	0.125	0.890
		Jackknife-BC	1.092	1.088	0.011	0.032	0.077	1.104
		Analytical-BC	0.917	0.915	0.008	0.125	0.202	0.850
		PCML-BC	0.936	0.933	0.008	0.003	0.013	1.450
500	4							
		ML	0.816	0.815	0.005	0.836	0.897	0.879
		Analytical-BC	0.513	0.551	0.011	1.000	1.000	0.293
		PCML-BC	0.800	0.800	0.004	0.046	0.174	2.355
500	8							
		ML	0.948	0.947	0.004	0.178	0.288	0.895
		Jackknife-BC	1.093	1.092	0.005	0.208	0.326	0.995
		Analytical-BC	0.920	0.919	0.003	0.449	0.578	0.859
		PCML-BC	0.936	0.936	0.003	0.022	0.069	1.440

1000 replications. ML is the plug-in estimator for the dynamic logit model; Jackknife-BC denotes Dhaene and Jochmans (2015)'s jackknife bias correction; Analytical-BC denotes Fernández-Val (2009)'s analytical bias correction; PCML-BC denotes the proposed estimator. With $n = 100$ and $T = 4$, we removed 11 Monte Carlo draws for Analytical-BC as convergence failed

References

Alessie R, Hochguertel S, van Soest A (2004) Ownership of stocks and mutual funds: a panel data analysis. Rev Econ Stat 86:783–796

Andersen EB (1970) Asymptotic properties of conditional maximum-likelihood estimators. J Roy Stat Soc B 32:283–301

Arroyabe MF, Schumann M (2022) On the estimation of true state dependence in the persistence of innovation. Oxf Bull Econ Stat 84:850–893

Bao Y, Ullah A (2007) The second-order bias and mean squared error of estimators in time-series models. J Econom 140:650–669

Barndorff-Nielsen O (1978) Information and exponential families in statistical theory. Wiley, Chichester

Bartolucci F, Nigro V (2010) A dynamic model for binary panel data with unobserved heterogeneity admitting a $\sqrt{n}$-consistent conditional estimator. Econometrica 78:719–733

Bartolucci F, Nigro V (2012) Pseudo conditional maximum likelihood estimation of the dynamic logit model for binary panel data. J Econom 170:102–116

Bartolucci F, Pigini C (2019) Partial effects estimation for fixed-effects logit panel data models (MPRA Paper No. 92243). University Library of Munich, Germany

Bartolucci F, Pigini C, Valentini F (2021) MCMC Conditional Maximum Likelihood for the two-way fixed-effects logit (MPRA Paper No. 110034). University Library of Munich, Germany

Bester CA, Hansen C (2009) A penalty function approach to bias reduction in nonlinear panel models with fixed effects. J Bus Econ Stat 27:131–148

Bettin G, Lucchetti R, Pigini C (2018) A dynamic double hurdle model for remittances: evidence from Germany. Econ Model 73:365–377

Cameron AC, Trivedi PK (2005) Microeconometrics: methods and applications. Cambridge University Press, New York

Cappellari L, Jenkins SP (2004) Modelling low income transitions. J Appl Econom 19:593–610

Carro JM (2007) Estimating dynamic panel data discrete choice models with fixed effects. J Econom 140:503–528

Chamberlain G (1980) Analysis of covariance with qualitative data. Rev Econ Stud 47:225–238

Chamberlain G (1985) Heterogeneity, omitted variable bias, and duration dependence. In: Heckman JJ, Singer B (eds) Longitudinal analysis of labor market data. Cambridge University Press, Cambridge

Charbonneau KB (2017) Multiple fixed effects in binary response panel data models. Econom J 20:S1–S13

Chernozhukov V, Ferández-Val I, Hahn J, Newey W (2013) Average and quantile effects in nonseparable panel models. Econometrica 81:535–580

Contoyannis P, Jones AM, Rice N (2004) Simulation-based inference in dynamic panel probit models: an application to health. Empir Econ 29:49–77

Cox D (1972) The analysis of multivariate binary data. Appl Stat 21:113–120

Dhaene G, Jochmans K (2015) Split-panel jackknife estimation of fixedeffect models. Rev Econ Stud 82:991–1030

Dhaene G, Sun Y (2021) Second-order corrected likelihood for nonlinear panel models with fixed effects. J Econom 220:227–252

Drescher K, Janzen B (2021) Determinants, persistence, and dynamics of energy poverty: an empirical assessment using German household survey data. Energy Econ 102:105433

Fernández-Val I (2009) Fixed effects estimation of structural parameters and marginal effects in panel probit models. J Econom 150:71–85

Fernández-Val I, Weidner M (2016) Individual and time effects in nonlinear panel models with large N, T. J Econom 192:291–312

Hahn J, Kuersteiner G (2011) Bias reduction for dynamic nonlinear panel models with fixed effects. Econom Theor 27:1152–1191

Hahn J, Newey W (2004) Jackknife and analytical bias reduction for nonlinear panel models. Econometrica 72:1295–1319

Hanfelt JJ, Wang L (2014) Simple relaxed conditional likelihood. Biometrika 101:726–732

Hansen LP (1982) Large sample properties of generalized method of moments estimators. Econometrica 50:1029–1054

Heckman JJ (1981) The incidental parameters problem and the problem of initial conditions in estimating a discrete time-discreted data stochastic process. In: Manski C, McFadden D (eds) Structural analysis of discrete data with econometric applications. MIT Press, Cambridge

Heckman JJ, Borjas GJ (1980) Does unemployment cause future unemployment? Definitions, questions and answers from a continuous time model of heterogeneity and state dependence. Economica 47:247–283

Honoré BE, Kyriazidou E (2000) Panel data discrete choice models with lagged dependent variables. Econometrica 68:839–874

Honoré BE, Weidner M (2020) Moment conditions for dynamic panel logit models with fixed effects. arXiv preprint arXiv:2005.05942

Hsiao C (2005) Analysis of panel data, 2nd edn. Cambridge University Press, New York

Hyslop DR (1999) State dependence, serial correlation and heterogeneity in intertemporal labor force participation of married women. Econometrica 67:1255–1294

Jochmans K (2018) Semiparametric analysis of network formation. J Bus Econ Stat 36:705–713

Kunz J, Śtaub KE, Winkelmann R (2019) Predicting fixed effects in panel probit models. Monash Business School 19

Lancaster T (2000) The incidental parameter problem since 1948. J Econom 95:391–413

Newey WK, McFadden D (1994) Large sample estimation and hypothesis testing. Handb Econ 4:2111–2245

Neyman J, Scott EL (1948) Consistent estimates based on partially consistent observations. Econometrica 16:1–32

Nickell S (1981) Biases in dynamic models with fixed effects. Econometrica 49:1417–1426

Pigini C, Presbitero AF, Zazzaro A (2016) State dependence in access to credit. J Financ Stab 27:17–34

Stammann A, Heiß F, McFadden D (2016) Estimating fixed effects logit models with large panel data (Tech. Rep. No. G01-V3). ZBW - Deutsche Zentralbibliothek für Wirtschaftswissenschaften, Leibniz-Informationszentrum Wirtschaft

Wooldridge JM (2005) Simple solutions to the initial conditions problem in dynamic, nonlinear panel data models with unobserved heterogeneity. J Appl Econom 20:39–54

Wooldridge JM (2010) Econometric analysis of cross section and panel data. MIT Press, Cambridge

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.