Journal of Econometrics 150 (2009) 71–85

Contents lists available at ScienceDirect

Journal of Econometrics

journal homepage: www.elsevier.com/locate/jeconom

Fixed effects estimation of structural parameters and marginal effects in panel probit models

Iván Fernández-Val *

Department of Economics, Boston University, United States

ARTICLE INFO

Article history:
Received 10 February 2007
Received in revised form
4 February 2009
Accepted 17 February 2009
Available online 6 March 2009

JEL classification:
C23
C25
J22

Keywords:
Panel data
Bias
Discrete choice models
Probit
Incidental parameters problem
Fixed effects
Labor force participation

ABSTRACT

Fixed effects estimators of nonlinear panel models can be severely biased due to the incidental parameters problem. In this paper, I characterize the leading term of a large-T expansion of the bias of the MLE and estimators of average marginal effects in parametric fixed effects panel binary choice models. For probit index coefficients, the former term is proportional to the true value of the coefficients being estimated. This result allows me to derive a lower bound for the bias of the MLE. I then show that the resulting fixed effects estimates of ratios of coefficients and average marginal effects exhibit no bias in the absence of heterogeneity and negligible bias for a wide variety of distributions of regressors and individual effects in the presence of heterogeneity. I subsequently propose new bias-corrected estimators of index coefficients and marginal effects with improved finite sample properties for linear and nonlinear models with predetermined regressors.

© 2009 Elsevier B.V. All rights reserved.

1. Introduction

Panel data models are widely used in empirical economics because they allow researchers to control for unobserved individual time-invariant heterogeneity. However, these models pose important technical challenges in nonlinear and/or dynamic settings. In particular, if individual heterogeneity is left completely unrestricted, then estimates of model parameters suffer from the incidental parameters problem, first noted by Neyman and Scott (1948). This problem arises because unobserved individual characteristics are replaced by sample estimates, biasing estimates of model parameters. Examples include estimators for probit models with fixed effects and (linear and nonlinear) models with lagged dependent variables and fixed effects (see, e.g., Nerlove (1967, 1971), Heckman (1981), Nickell (1981), Katz (2001), Greene (2004) and Hahn and Newey (2004)).

In this paper, I develop bias corrections for fixed effects conditional maximum likelihood estimators (MLEs) in parametric panel binary choice models. These corrections are based on expressions for the bias that intensively exploit the structure of the problem by taking expectations using the conditional parametric model. Observed quantities are therefore replaced by expected quantities in the estimation of bias. This approach is similar to the use of the conditional information matrix in the estimation of the asymptotic variances of MLEs, instead of other alternatives such as the sample average of the outer product of the scores or the sample average of the negative Hessian (see Porter (2002)). Numerical results show that this refinement improves the finite sample performance of the correction over other existing methods.

In the case of probit index coefficients, I find that the bias of the conditional MLE is the product of a matrix and the true value of these coefficients, plus a second order term. This result allows me to derive a lower bound for the first order bias, depending uniquely upon the number of time periods in the panel. This bound establishes, for example, that the bias is at least 20% for 4-period panels and 10% for 8-period panels. When there is a single regressor, the above product matrix is a positive scalar and the probit fixed effects estimates are therefore biased away from zero, providing a theoretical foundation for previous numerical evidence (see, for e.g., Greene (2004)). In the case of no individual heterogeneity, the product matrix is a positive scalar multiple of the identity matrix, suggesting that the fixed effects estimators of ratios of coefficients do not suffer from incidental parameters bias when the level of heterogeneity is moderate. These ratios are

* Corresponding address: Department of Economics, Boston University, 270 Bay State Road, 02215 Boston, MA, United States. Tel.: +1 617 353 9670; fax: +1 617 353 4449.

E-mail address: ivanf@bu.edu.

0304-4076/$ – see front matter © 2009 Elsevier B.V. All rights reserved.
doi:10.1016/j.jeconom.2009.02.007

structural parameters of interest because they can be interpreted as marginal rates of substitution in economic applications for which the index coefficients are only identified up to scale.

In nonlinear models one must often go beyond estimation of model parameters to obtain estimated marginal effects. In probit models, for example, the index coefficients cannot be interpreted as the effects of changes in the regressors on the conditional probability of the outcome. Accordingly, I also study the properties of fixed effects estimators of average marginal effects.[1] The motivation for my analysis comes from a question posed by Wooldridge: "How does treating the *individual effects* as parameters to estimate–in a 'fixed effects probit' analysis–affect estimation of the APEs *(average partial effects)*?"[2] Wooldridge (2002) conjectures that the estimators of the marginal effects have reasonable properties. Here, I characterize the analytical expression of the leading term of a large-T expansion of the bias of these average marginal effects. As Wooldridge anticipated, this bias is negligible relative to the true average marginal effect for a wide variety of distributions of regressors and individual effects and is identically zero in the absence of heterogeneity. This helps explain the small biases in the marginal effects estimates that Hahn and Newey (2004) (HN04 henceforth) find in Monte Carlo examples.

Most of the theoretical results I derive in this paper are concerned with static probit models with exogenous regressors, but I also explore related questions in other linear and nonlinear models with predetermined regressors and fixed effects. In particular, I find numerical evidence suggesting that probit and logit fixed effects estimates of index coefficients and average marginal effects are biased downward for lagged dependent variables. This finding for marginal effects in dynamic nonlinear models resembles the analogous result for fixed effects estimators of model parameters in dynamic linear models. I subsequently develop new bias correction methods for estimates of index coefficients and marginal effects in probit and logit dynamic models that exhibit better finite sample properties than the existing alternatives. Simple linear probability models, in the spirit of Angrist (2001), also perform well in estimating average marginal effects for exogenous regressors but need to be corrected when the regressors are just predetermined.

The properties of probit and logit fixed effects estimators of model parameters and marginal effects are illustrated by an analysis of female labor force participation, using 10 waves from the Panel Survey of Income Dynamics (PSID). This analysis is motivated by similar studies in labor economics in which panel binary choice processes have been widely used to model female labor force participation decisions (see, e.g., Hyslop (1999), Chay and Hyslop (2000) and Carro (2007)). I find that fixed effects estimators, while biased for index coefficients, give very similar estimates to their bias-corrected counterparts of marginal effects in static models. On the other hand, uncorrected fixed effects estimators of both index coefficients and marginal effects are biased in dynamic models that account for true state dependence. In this case, the bias corrections I present in this paper are effective in reducing the incidental parameters problem.

The approach followed in this paper is related to the recent large-n large-T literature for panel data estimators including, e.g., Phillips and Moon (1999), Lancaster (2002), Hahn and Kuersteiner (2002), Woutersen (2002), Arellano (2003), Alvarez and Arellano (2003), Hahn and Kuersteiner (2004), HN04, and Carro (2007); see also Arellano and Hahn (2007) for a recent survey of this literature and additional references. These studies aim to provide what I refer to as large-T consistent estimates because they rely on an asymptotic approximation to the behavior of the estimator that lets both the number of individuals n and the time dimension T grow with the sample size.[3] The idea behind these methods is to expand the incidental parameters bias of the estimator on the order of magnitude T, and to subtract an estimate of the leading term of the bias from the estimator.[4] As a result, the adjusted estimator has a bias of order T^{-2}, whereas the order of bias of the initial estimator is T^{-1}. All the above papers focus mainly on estimation of model parameters. The contribution of this paper is to provide a refinement of the bias corrections for parametric binary choice models with improved finite-sample properties and to develop theoretical results for the bias of index coefficients and marginal effects in probit models that help explain previous numerical findings.

The paper is organized as follows. Section 2 describes the problem of fixed effects estimation in panel binary choice models and develops bias corrections for these models. Section 3 looks at fixed effects estimation of marginal effects and establishes the small bias property for the probit model. Monte Carlo results and an empirical illustration are given in Sections 4 and 5, respectively. Section 6 concludes. The proofs of the main results are given in the Appendix.

2. Fixed effects estimation of panel binary choice models

2.1. The model

Given a binary response Y and a $p \times 1$ regressor vector X, the response for individual i at time t is assumed to be generated by the following single index process:

$$Y_{it} = \mathbf{1}\left\{X_{it}'\theta_0 + \alpha_i - \epsilon_{it} \geq 0\right\}, \quad \text{for } i = 1, \ldots, n \text{ and } t = 1, \ldots, T,$$

where $\mathbf{1}\{C\}$ is an indicator function that takes on value one if condition C is satisfied and zero otherwise, θ_0 denotes a $p \times 1$ vector of parameters (index coefficients), α_i is a scalar unobserved individual effect, and ϵ_{it} is a time/individual-specific random shock. This is an error-components model where the unobserved error term is decomposed into a permanent individual-specific component α_i and a transitory shock ϵ_{it}. Examples of economic decisions that can be modeled within this framework include labor force participation, union membership, migration, purchase of durable goods, marital status, and fertility (see Amemiya (1981), for a survey).

2.2. Fixed effects estimators

In economic applications, regressors and individual heterogeneity are usually correlated because regressors typically include choice variables and individual heterogeneity usually represents variation in tastes or technology. To avoid imposing any structure

[1] Marginal effects are defined either as the change in the conditional outcome probability as a response to a one-unit increase in a regressor or as a local approximation to this quantity based on the slope of the conditional outcome probability. For example, in a probit model with a single regressor and constant term, the marginal effect can be defined either as $\Phi(\alpha + (x+1)\theta) - \Phi(\alpha + x\theta)$ or $\theta\phi(\alpha + x\theta)$, where $\Phi(\cdot)$ and $\phi(\cdot)$ denote the CDF and PDF of the standard normal distribution, respectively.

[2] C.f., Wooldridge (2002), p. 489 (italics mine).

[3] Fixed-T-consistent estimators have also been derived for panel logit models (see Cox (1958), Rasch (1960), Andersen (1973) and Chamberlain (1980) for the static case; and Cox (1958); Chamberlain (1985); Honoré and Kyriazidou (2000) for the dynamic case) and other semiparametric index models (see Manski (1987) for the static case; and Honoré and Kyriazidou (2000) for the dynamic case). These methods, however, do not provide estimates for individual effects, thus precluding estimation of other quantities of interest such as marginal effects.

[4] To avoid complicated terminology, in the future I will generally refer to the leading term of the large-T expansion of the bias simply as the bias.

Download English Version:

https://daneshyari.com/en/article/5097003

Download Persian Version:

https://daneshyari.com/article/5097003

Daneshyari.com