Journal of Econometrics 175 (2013) 46–59

Contents lists available at SciVerse ScienceDirect

Journal of Econometrics

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

Estimation of a nonlinear panel data model with semiparametric individual effects[☆]

Wayne-Roy Gayle[a,*], Soiliou Daw Namoro[b]

[a] Department of Economics, University of Virginia, Monroe Hall, McCormick Rd., Room 208A, Charlottesville, VA 22903, United States

[b] Department of Economics, University of Pittsburgh, 230 S. Bouquet St., Pittsburgh, PA 15260, United States

ARTICLE INFO

Article history:
Received 9 April 2012
Received in revised form
6 November 2012
Accepted 14 March 2013
Available online 22 March 2013

JEL classification:
C13
C14
C23

Keywords:
Semiparametric estimation
Modified backfitting
Panel data
Nonlinear models

ABSTRACT

This paper investigates identification and estimation of a class of nonlinear panel data, single-index models. The model allows for unknown time-specific link functions, and semiparametric specification of the individual-specific effects. We develop an estimator for the parameters of interest, and propose a powerful new kernel-based modified backfitting algorithm to compute the estimator. We derive uniform rates of convergence results for the estimators of the link functions, and show the estimators of the finite-dimensional parameters are root-N consistent with a Gaussian limiting distribution. We study the small sample properties of the estimator via Monte Carlo techniques.

Published by Elsevier B.V.

1. Introduction

This paper is concerned with identification and estimation of the following semiparametric regression model:

$$y_{it} = \Phi_t\left(x_{it}\beta + \eta(z_i)\right) + \epsilon_{it}, \quad t = 1, \ldots, T, \tag{1.1}$$

where x_{it} is a K-dimensional row vector of random variables, z_i is an L-dimensional row vector of time-constant random variables, ϵ_{it} is an individual- and time-specific idiosyncratic shock that is assumed to be mean independent of the other explanatory variables, β is a K-dimensional vector of parameters, Φ_t is a strictly increasing and smooth unknown link function, and η is an unknown function. The parameters of primary interest are β, and $\Phi := \{\Phi_t, t = 1, \ldots, T\}$.

We propose a powerful new kernel-based algorithm to compute the estimator for the parameters of interest. The algorithm combines the profile likelihood approach of Severini and Wong (1992) with the backfitting algorithms of Buja et al. (1989), Mammen et al. (1999, 2001), and extends them to the present framework. The algorithm fully implements the identification restrictions of the model. We provide sufficient conditions under which the algorithm converges. Also, we derive uniform rates of convergence results for the estimators of the link functions, and show the resulting estimator of β is $\sqrt{N}$-consistent with a Gaussian limiting distribution. Furthermore, estimation of the finite-dimensional parameters is adaptive with respect to estimation of the link functions.

The model presented in Eq. (1.1) is a panel data version of the generalized partial-linear model (GPLM) with unknown link functions. Other methods can be used to estimate the parameters of the model, including the backfitting estimator developed in Opsomer (2000), among others, and the methods of series and sieve minimum distance estimation developed in Newey (1994a), Newey and Powell (2003), Ai and Chen (2003), Chen (2007) and Gayle and Viauroux (2007). However, the method developed in this paper has some key advantages over these alternatives.

Opsomer (2000) develops a backfitting procedure to estimate the parameters of additive and partial linear models. This procedure can be modified to the panel data framework. However, it is unclear how to impose shape constraints on the estimators of the infinite-dimensional parameters in an internally consistent way using the method developed in Opsomer (2000). Indeed, a

[☆] The authors are grateful to Mehmet Caner, George-Levi Gayle, Robert Miller, Holger Sieg, Jean-Francois Richard, and three anonymous referees for insightful comments and discussions. Comments by the participants of the 12th Conference on Panel Data at Copenhagen, Denmark, 2005, were greatly appreciated. All remaining errors are our own.

* Corresponding author. Tel.: +434 243 4336; fax: +434 982 2904.
E-mail address: wg4b@virginia.edu (W.-R. Gayle).

0304-4076/$ – see front matter. Published by Elsevier B.V.
http://dx.doi.org/10.1016/j.jeconom.2013.03.004

maintained assumption for identification of the parameters of interest in Eq. (1.1) is that the link functions are strictly monotonic. On the other hand, the modified backfitting algorithm developed in Mammen et al. (1999, 2001) accommodates shape restrictions on the infinite-dimensional parameters under the same empirical norm as the one constructed to define estimators of all the parameters. Another key advantage of the algorithm developed in Mammen et al. (1999, 2001) is that its convergence is well understood, and does not depend on initial values.

Alternatively, estimating the parameters of interest by implementing the methods of series or sieve estimation developed in Newey (1994a), Newey and Powell (2003), Ai and Chen (2003), Chen (2007), and Gayle and Viauroux (2007) is feasible. However, these methods rely on the choice of smoothers used to compute the estimators of the infinite-dimensional parameters. The estimator developed in this paper can be computed using a wide variety of smoothers.[1] We focus on the case where the smoothers are kernels. To the best of our knowledge, no existing studies investigate kernel-based estimation of panel data GPLM models such as Eq. (1.1) with shape constraints on the unknown link functions.

The model developed in this paper builds on previous work of Chamberlain (1980), Newey (1994a), Chen (1998), and Arellano and Carrasco (2003), to name a few, concerning the estimation of binary-choice, panel data models with individual-specific effects. The common strategy of these papers, as well as ours, is to impose restrictions on the conditional distribution of the individual-specific effects, conditioned on the observed regressors. However, the estimator developed here differs in a variety of ways.

The estimator we propose in this paper treats both Φ and η as unknown functions. The models (Chamberlain, 1980) propose assume the link functions are known, and that η is known up to a set of finite-dimensional parameters. Newey (1994a) extends this framework to allow for η to be an unknown function, while maintaining the parametric specification of the link functions. We extend the model presented in Newey (1994a) to allow for unspecified time-specific link functions. In the discrete-choice framework, Chen (1998) modifies the framework of Newey (1994a) by relaxing the parametric specification of the link functions at the cost of increased restrictions on the finite-dimensional parameters. In this paper, we achieve identification and estimation without these additional restrictions on the finite-dimensional parameters. We restrict identification of the parameters of interest to the static panel data framework. Arellano and Carrasco (2003) develop a panel data discrete-choice model that allows for predetermined explanatory variables. However, the model (Arellano and Carrasco, 2003) present assume the link functions are known.

Other important developments in the semiparametric panel data literature include Manski (1987); Honoré and Lewbel (2002). Compared with Manski (1987), this paper imposes stronger restrictions on the joint distribution of the individual effects and observed regressors, but allows for the errors to be heteroskedastic over time. Honoré and Lewbel (2002) impose a different conditional independence assumption on the distribution of the individual-specific effects and the error term given the observables, which is neither weaker nor stronger than the one we impose in this paper. Also, Honoré and Lewbel (2002) imposes stronger support conditions on the observable and unobservable explanatory variables.

In the next section, we provide an example of how Eq. (1.1) is derived from the familiar binary-choice, single-index panel data model. However, our own interest goes beyond the binary-choice framework. Any model that can be presented in the form of Eq. (1.1) can be estimated using the method we develop in this paper.

We investigate the small sample performance of the proposed estimator in two environments by Monte Carlo analysis. The first exercise examines the performance of the estimator in a static, panel data discrete-choice model, and the second in a static, panel data continuous-outcome model. The simulation exercises show the estimator performs well in small samples.

We organize the rest of the paper as follows. Section 2 motivates Eq. (1.1) by describing how it is derived from various econometric models. Section 3 discusses identification, and Section 4 presents the estimator. Section 5 presents the algorithm used to compute the estimator. Section 6 derives the large sample properties of the estimator. Section 7 proposes an estimator for the asymptotic variance of the finite-dimensional parameters. Section 8 is devoted to the Monte Carlo simulations, and Section 9 concludes. All proofs and auxiliary lemmas are in the Appendix.

2. The model

In this section, we discuss how Eq. (1.1) may be derived from more primitive econometric models. Consider the following panel data, single-index model for a unit of observation, i:

$$y_{it} = F_t(x_{it}\beta + \mu_i) + r_{it}, \quad t = 1, \ldots, T, \tag{2.1}$$

where y_{it} is the dependent variable, x_{it} are observable time-varying explanatory variables, μ_i is the time-invariant unobserved effect, F_t is an unknown, strictly increasing function, and r_{it} is the idiosyncratic error where $E[r_{it}|x_{it}, \mu_i] = 0$, $t = 1, \ldots, T$. It is well known that Eq. (2.1) can be derived from the following panel data, single-index discrete-choice model:

$$y_{it} = 1\{x_{it}\beta + \mu_i - u_{it} \geq 0\}, \quad t = 1, \ldots, T, \tag{2.2}$$

where x_{it} and μ_i are as described, u_{it} is independent of x_{it} and μ_i with an unknown time-specific distribution function F_t that is absolutely continuous with respect to a Lebesgue measure.

Suppose that for each unit of observation, a vector of observable time-invariant explanatory variables, z_i, exists, and assume the individual-specific effects can be decomposed as follows: $\mu_i = \eta(z_i) + v_i$. Assume that v_i is independent of z_i and $x_i := (x_{i1}, \ldots, x_{iT})$ with distribution that is absolutely continuous with respect to a Lebesgue measure, and has a Radon–Nikodym derivative f_v. Under these assumptions, taking conditional expectations of y_{it} conditional on (x_{it}, z_i, v_i) in Eq. (2.1) gives

$$E[y_{it}|x_{it}, z_i, v_i] = F_t(x_{it}\beta + \mu_i) = F_t(x_{it}\beta + \eta(z_i) + v_i).$$

Eq. (2.1) is therefore obtained by defining $r_{it} = y_{it} - E[y_{it}|x_{it}, z_i, v_i]$. Furthermore, by the law of iterated expectations,

$$E[y_{it}|x_{it}, z_i] = \Phi_t(x_{it}\beta + \eta(z_i)),$$

where $\Phi_t(a) := \int F_t(a + v)f_v(v)dv$. By defining $\epsilon_{it} = y_{it} - E[y_{it}|x_i, z_i]$, we obtain Eq. (1.1), where Φ_t inherits the monotonicity constraint on F_t. Because we estimate Φ_t, and not F_t, any predictions made using these estimates should be interpreted as predictions made after integrating out the "pure" random effects component, v_i. Note that F_t is not needed to obtain average partial effects, because these effects can be computed from Φ_t, β, and η. This discussion shows that under certain assumptions, and by appropriately defining z_i, Eq. (1.1) is implied by a variety of models that are popular in applied work.

Returning to Eq. (1.1), define $w_{it} = (x_{it}, z_i)$. By taking conditional expectations of y_{it} conditioned on w_{it} in Eq. (1.1), we obtain

$$P_{it} := P(w_{it}) := E[y_{it}|w_{it}] = \Phi_t(x_{it}\beta + \eta(z_i)), \quad t = 1, \ldots, T. \tag{2.3}$$

The following assumption formalizes the monotonicity constraint on the link function we will maintain in this paper.

[1] See Mammen et al. (2001) for discussions on the implementation of the Nadaraya–Watson smoother and series smoothers.