ARTICLE IN PRESS

Economic Modelling xxx (xxxx) xxx–xxx

Contents lists available at ScienceDirect

Economic Modelling

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

Binary choice model with interactive effects☆

Sen Xue[a], Thomas Tao Yang[b,*], Qiankun Zhou[c]

[a] Jinan University, China
[b] Australian National University, Australia
[c] Louisiana State University, United States

ARTICLE INFO

JEL classification:
C01
C35
C38

Keywords:
Binary choice model
Interactive effects
Projection method
Partition regression
Special regressor

ABSTRACT

This paper considers the estimation of binary choice model with interactive effects. We propose an estimator via combining the projection method (Mundlak, 1978) and the special regressor method (Lewbel, 2000a). The proposed estimator requires mild distribution assumptions on unobservables. The asymptotic property of this estimator is established. Our estimator performs well in Monte Carlo simulations. We apply this estimator to study migration intention of rural residents in China. The results suggest that numbers of co-habiting children and elderly are negatively correlated with migration intention.

1. Introduction

Econometric analysis of models with interactive effects has been the subject of recent research attention, in both theoretical and empirical studies. The interactive effects are used to capture the interactions between unobserved individual and time-specific effects. Compared to models without interactive effects, the model with interactive effects provides a more reliable estimator, e.g., Bai, 2003, 2009a; Bai and Ng, 2002, 2008. Moreover, taking interactive effects into account also reduces the heterogeneity of the model and thus eliminate the source of bias in panel data models (Hsiao, 2014). A number of different approaches have been advanced for applying models with interactive effects: Pesaran (2006) proposes the common correlated effects (CCE) estimator which can be computed by least squares in augmented regressions with cross-sectional averages of the dependent variable and the individual-specific regressors; Bai (2009a) provides identification and estimation of panel data model with interactive effects through the principal component approach; other approaches can be found in Bai and Ng (2008) and the references therein.

Two issues arise with these attempts to apply models with interactive effects. First, these approaches assume that the model is linear and cannot be applied to the nonlinear case (for example, binomial response). Second, these approaches assume large N and large T when deriving the limiting behavior of the estimator, but T is often small in micro-level data.

In this paper, we extend the analysis to the binary choice case when N is large and T is fixed. Different from the usual methods of handling interactive effect proposed by Bai (2009a) and Pesaran (2006), our approach relies on a projection method which is widely used to model unobserved effects by observables, e.g., Hayakawa (2013); Semykina and Wooldridge, 2010. Specifically, we adopt the projection method of Mundlak (1978) to model the interactive effects, following Bai (2009b). Further, the identification of our approach relies on the special regressor approach, e.g., Lewbel (2000a) and Honoré and Lewbel (2002). The main idea is that the nonlinear model becomes a linear one after taking an expectation of some transformation of the dependent variable, assisted by a “special” regressor.[1] Parameters of interests can be identified from the linearized model via normal approaches thereafter. The estimator recently proposed by Fernandez-Val and Weidner (2016) needs nonlinear optimization to obtain the estimates. On the contrary, the computation of our estimator is rather simple. Another desirable feature of our approach is that we impose very mild distribution assumptions on unobservables.

☆ We greatly appreciate constructive comments from Cheng Hsiao, Arthur Lewbel, Roger Moon, Joakim Westerlund, Yonghui Zhang, James Jilu Zhang and participants from Economic Modelling special issue conference, which have greatly improved the original paper. This study was supported by the Fundamental Research Funds for the Central Universities in China (Grant No. 17JNQN023). All remaining errors are solely ours.
* Corresponding author.
E-mail address: youngtigers08@gmail.com (T.T. Yang).
[1] For recent works of special regressor method, refer to Dong and Lewbel (2015); Lewbel (2012) and Lewbel et al. (2012).

http://dx.doi.org/10.1016/j.econmod.2017.08.002
Received 30 July 2017; Accepted 7 August 2017

Please cite this article as: Yang, T.T., Economic Modelling (2017), http://dx.doi.org/10.1016/j.econmod.2017.08.002

The Monte Carlo evidence shows that our special regressor approach outperforms the naive Maximum Likelihood Estimation (MLE) that ignores unobserved interactive effects. When N is relatively large, our approach is almost as good as the “oracle” infeasible MLE that assumes that the unobserved interactive effects can be observed. We apply our approach to analyze the determinants of migration intention of rural residents in China. The results show that people who live with more children are less inclined to migrate, suggesting that the phenomenon of the left-behind family is an important factor restraining migration in China.

The rest of the paper is organized as follows. Section 2 introduces the model, provides assumptions, identification and the estimator; and establishes the large sample property of our estimator. Section 3 reports the results of the Monte Carlo simulation. Section 4 applies the model to analyze the migration intention of rural residents in China. Section 5 concludes. Main proofs are in Appendix A, tables are in Appendix B, and technical proofs are relegated to Appendix C.

2. Model

2.1. Setup

We begin by considering the following discrete choice model with interactive effects

$$y_{it}^* = v_{it} + c_i + \delta_t + x'_{it}\beta + u_{it}, \quad t = 1, \ldots, T;\ i = 1, \ldots, N, \tag{1}$$

$$u_{it} = \lambda'_i f_t + \varepsilon_{it}, \tag{2}$$

$$y_{it} = 1(y_{it}^* > 0), \tag{3}$$

where y_{it} is the observation on the i-th cross-section unit at time t, δ_t is time effect, and x_{it} is a $k \times 1$ vector of observed individual-specific regressors on the i-th cross-section unit at time t, λ_i and f_t are $r \times 1$ and both are unobserved and vary across i and t respectively, and ε_{it} is the error term. $1(A)$ is the indicator function and takes value one if condition A is satisfied and zero otherwise. The number of factors r is fixed. Moreover, we assume v_{it} is a special regressor, which satisfies the following conditions: (i) v_{it} is a continuous random variable; (ii) v_{it} is independent of c_i, δ_t and u_{it} conditional on x_{it}; (iii) v_{it} has a relatively large support. These conditions will be elaborated more in the following sections.

Example 2.1. This model is widely used in economics, especially in microeconometrics. Take health insurance participation as an example. One could use y_{it}, which is an indicator of one and zero, to denote participation in health insurance. Here, participation is a function of observed characteristics x_{it}, such as income and residential location, and unobserved characteristics, such as risk preferences and innate health endowment, which are likely to be correlated with observed characteristics. The unobserved characteristics can be grouped into two categories, those with a time-invariant effect (c_i) and those with a time-varying effect (λ_i). f_t captures the potentially time-varying effect. Compared to the standard fixed effect model, this model considers the possibility that the effect of some unobserved characteristics is time-varying. Thus, this model can remove more unobserved heterogeneity and yield more reliable estimates than the standard fixed effect model.

In this paper, we focus on the case that the number of cross-section units N zis large and the number of time periods T is fixed. Therefore, it is desirable to treat (δ_t, f_t) as parameters instead of treating (c_i, λ_i) as parameters.

In many empirical applications, λ_i and observed characteristics x_{it} are correlated. Given this, we apply the projection method used for modeling unobservables with observables to model λ_i (for recent application of projection method, refer to Bai, 2009b; Hayakawa, ; Semykina and Wooldridge, 2010). Following Chamberlain (1982) and Bai (2009b), we assume that

$$E(\lambda_i | x_{i1}, x_{i2}, \ldots, x_{iT}) = \lambda + \sum_{s=1}^{T} \psi_s x_{is} \tag{4}$$

where λ is a $r \times 1$ vector and ψ_s is an $r \times k$ matrix ($s \geq 1$). Observing that the above projection contains too many parameters to identify, we instead consider a restricted version of projection (Mundlak, 1978) as follows

$$E(\lambda_i | x_{i1}, x_{i2}, \ldots, x_{iT}) = \lambda + \psi \bar{x}_i \tag{5}$$

with $\bar{x}_i = \frac{1}{T}\sum_{t=1}^{T} x_{it}$ and ψ is an $r \times k$ matrix. The above projection can be expressed as

$$\lambda_i = \lambda + \psi \bar{x}_i + \eta_i$$

where, by definition, we have $E(\eta_i | x_{i1}, x_{i2}, \ldots, x_{iT}) = 0$.

As a result,

$$y_{it}^* = v_{it} + c_i + (\delta_t + \lambda f_t) + x'_{it}\beta + \bar{x}'_i \psi' f_t + \eta'_i f_t + \varepsilon_{it}.$$

Abusing the notation a bit, we continue using δ_t to denote $\delta_t + \lambda f_t$ for simplicity. Then, our model becomes:

$$y_{it}^* = v_{it} + c_i + \delta_t + x'_{it}\beta + \bar{x}'_i \psi' f_t + \eta'_i f_t + \varepsilon_{it}, \tag{6}$$

$$y_{it} = 1(y_{it}^* > 0) t = 1, \ldots, T;\ i = 1, \ldots, N.$$

Remark 2.1. The Mundlak's projection method of Eq. (6) is closely related to the augmented regression method proposed by Pesaran (2006), which suggests approximating f_t by observable proxies (δ_t, x_{it}).

2.2. Identification

Without further restrictions, the binary choice model can only be identified in the logistic case (Chamberlain, 2010), if the supports of the observed independent variables are bounded. However, the identification of special regressor approach is somewhat different from the usual approach. The price to pay in this paper is the existence of a “special” regressor (see below for details). With the aid of the special regressor, the expectation of a transformation of the dependent variable has a linear representation (see Eq. (8)). The identification of β can be achieved via standard methods on that transformation.

We impose technical assumptions below following Bai (2009b). We let $e_{it} \equiv f'_t \eta_i + \varepsilon_{it}$ for notional conveniences, and denote the conditional cumulative distribution of e_{it} on $x_{it}, \bar{x}_i$ as $F_{e_{it}}(e_{it} | x_{it}, \bar{x}_i)$ and denote its support as $\Omega_{e_{it}}$.

Assumption 1. Observations are *i.i.d.* over i.

Assumption 2. $E(e_{it} | x_{i1}, x_{it}, \bar{x}_i) = 0$.

For the special regressor, v_{it}, we impose the following assumptions on its support and distribution. Most of Assumptions S1 - S3 are standard in the literature, for instance, Lewbel (2000a); Honoré and Lewbel (2002); Gao et al. (2015).

Assumption S1: The conditional distribution of v_{it} given x_{it} is continuous with density $f_t(v_{it} | x_{it})$ that possibly varies across t. The support of v_{it} conditional on x_{it} is $[L_t, K_t]$ where $-\infty < L_t < 0 < K_t < \infty$, and $\inf_{v_{it} \in [L_t, K_t]} f_t(v_{it} | x_{it}) > 0$.

Assumption S2: $c_i, \eta_i, \varepsilon_{it} \perp v_{it} | x_{it}, \bar{x}_i$ and $f_t(v_{it} | x_{it}, x_{i1}, \bar{x}_i) = f_t(v_{it} | x_{it})$.

Assumption S3: The support of $s_{it} \equiv -c_i - \delta_t - x'_{it}\beta - \bar{x}'_i \psi' f_t - e_{it}$ is a subset of $[L_t, K_t]$.

We assume i.i.d. in Assumption 1. In Assumption 2, we add x_{i1} into the control variables, because we conduct the first-order difference for observations at time $t \neq 1$ with those at $t = 1$ to remove c_i. We do so in Assumption S2 for a similar reason. For Assumption S1, we permit heteroskedasticity of v_{it} at t dimension. In Assumption S2, we require $\bar{x}_i$