Economics Letters 169 (2018) 15–19

Contents lists available at ScienceDirect

Economics Letters

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

Partial effects in binary response models using a special regressor

Ying-Ying Lee [a,*], Hsueh-Hsiang Li [b]

[a] University of California Irvine, United States
[b] Colorado State University, United States

HIGHLIGHTS

- We study the average structural function in binary response models.
- We estimate partial effects of endogenous regressors using a special regressor.
- The proposed estimator is a marginal integration of a logistic series estimator.
- We derive the asymptotic theory for our partial mean estimator and its derivative.

ARTICLE INFO

Article history:
Received 28 January 2018
Received in revised form 4 May 2018
Accepted 5 May 2018

JEL classification:
C14
C21
C25

Keywords:
Average structural function
Binary choice
Endogeneity
Logistic series
Special regressor

ABSTRACT

This paper estimates the average structural function and partial effects of endogenous regressors in binary response models using a special regressor. We present formal asymptotic theory for the proposed logistic series estimator and estimate migration probabilities within the US.

1. Introduction

We propose a logistic series estimator for the average structural function (ASF) and partial effects in binary response models using a special regressor introduced by Lewbel (2000). Lewbel et al. (2012) note that the ASF is difficult to calculate and propose an alternative average index function (AIF) that is widely applied empirically on migration (Duval and Wolff, 2016; Ning and Qi, 2017; Ruyssen and Salomone, 2018), water consumption (Bontemps and Nauges, 2016), pesticide usage (Zapata Diomedi and Nauges, 2016), and regional variation in business cycles (Basile et al., 2014). Nevertheless, the AIF has the shortcoming of not identifying average partial effects of endogenous explanatory variables (Lin and Wooldridge, 2015). Our ASF estimator overcomes this problem through a marginal integration of a logistic series estimator.[1]

Consider a semiparametric binary response model $D = \mathbf{1}\{V + X\beta + \varepsilon > 0\}$, where the special regressor $V \in \mathcal{R}$ is independent of the heteroskedastic error $\varepsilon \in \mathcal{R}$ conditional on $X \in \mathcal{R}^{d_x}$, and some elements of X may be endogenous or mismeasured such that $\mathbb{E}[\varepsilon X] \neq 0$. We focus on the structural choice probability, i.e., the ASF in Blundell and Powell (2003), for a given hypothetical value (v, x), defined by

$$q(v, x) \equiv \int \mathbf{1}\{v + x\beta + \varepsilon > 0\} dF_\varepsilon(\varepsilon),$$

where F_ε is the differentiable cumulative distribution function of ε. Taking identification and estimation of coefficients β as given

* Correspondence to: 3151 Social Science Plaza, University of California Irvine, Irvine, CA 92697, United States

E-mail addresses: yingying.lee@uci.edu (Y.-Y. Lee), Hsueh-Hsiang.Li@colostate.edu (H.-H. Li).

[1] Stata codes to easily implement the estimator are available in the online appendices.

https://doi.org/10.1016/j.econlet.2018.05.002

from prior studies,[2] we contribute to the literature by estimating the ASF and the corresponding marginal/partial effect $\partial q(v,x)/\partial x$ that characterizes how the probability of $D = 1$ changes when X changes.

Identification of the ASF using a special regressor is shown in Chen et al. (2016) who suggest a local constant kernel estimator without deriving its asymptotic theory. We discuss identification in Section 2 and further add new sufficient conditions for the AIF to be equal to the ASF that is identified up to the index value $v + x\beta$. Section 3 introduces our ASF estimator that is a marginal integration of a logistic series estimator. We build on the literature on linear series estimators (Newey, 1997; Chen, 2007; Belloni et al., 2015) and functionals of nonlinear series estimator (Hirano et al., 2003; Cattaneo, 2010). We also analyze asymptotic properties of the partial effect estimator that is a partial derivative of a linear functional of a nonlinear series estimator. To the best of our knowledge, a general asymptotic theory of this type of estimator has not been shown in the literature. Moreover our ASF and partial effect estimators allow the preliminary coefficient estimator $\hat{\beta}$ to converge at a rate slower than root-n. This advantageous feature is particularly relevant to the finding in Khan and Tamer (2010) that the convergence rate of $\hat{\beta}$ proposed in Dong and Lewbel (2015) may not be root-n if the conditional distribution of V does not have thick tails, or large support, or tail symmetry in Magnac and Maurin (2007). All the proofs are in the supplemental appendix to our paper: see Lee and Li (2018).

In Section 4, we apply our estimator on the same dataset used in Dong and Lewbel (2015) who estimate the derivative of the AIF as the partial effect. Our estimator of the derivative of the ASF accounts for the endogeneity of regressors and reports smaller magnitudes of the partial effects for all determinants of interstate migration, comparing with Dong and Lewbel (2015). We also find that the effects of some factors (i.e., disability, educational level, and number of children) are no longer significant.

2. Identification

The idea to identify the ASF at a hypothetical value (v, x) is to match observations with the same index value $a \equiv v + x\beta$. The intuition is to use variation in the special regressor V in the index to compensate for variation in the endogenous regressor X. Under conditions that the coefficient β is identified, we focus on identification of the ASF and AIF. Assumption 1 is a conditional independence assumption or exclusion restriction on the special regressor.

Assumption 1 (*Special Regressor*)**.** $\varepsilon \perp V|(X,Z)$, where the instrumental variable Z satisfies $\mathbb{E}[\varepsilon Z] = 0$.

We suppress Z in the following discussion to learn intuition. For a hypothetical value (v, x), suppose there exists some (v', x') in the support of (V, X) such that $v' + x'\beta = v + x\beta = a$ reaching the same index value a. Then we can identify the conditional ASF given $X = x'$ by the conditional choice probability given $V = v' = a - x'\beta$ and $X = x'$:

$$\begin{aligned}
&\int \mathbf{1}\{v + x\beta + \varepsilon > 0\} dF_{\varepsilon|X}(\varepsilon|x') \\
&= \int \mathbf{1}\{v + x\beta + \varepsilon > 0\} dF_{\varepsilon|VX}(\varepsilon|v', x') \\
&= \int \mathbf{1}\{v' + x'\beta + \varepsilon > 0\} dF_{\varepsilon|VX}(\varepsilon|v', x') \\
&= \mathbb{E}[D|V = a - x'\beta, X = x'], \qquad (1)
\end{aligned}$$

where the first equality is implied by Assumption 1 and assuming v' is in the support of V conditional on $X = x'$. The second equality uses the index constraint $v' + x'\beta = v + x\beta = a$. To learn some insight, let us consider a subpopulation whose $X = x'$ and a hypothetical index value a. If $v' = a - x'\beta$ is a feasible choice of V for the subpopulation with $X = x'$, then the ASF of such subpopulation is the same at *any* hypothetical value (v, x) satisfying $v + x\beta = a$, regardless of their actual values of V.[3] Therefore the ASF is identified up to the index value a.

To identify the unconditional ASF $q(v, x)$, we need to find v' such that $x'\beta + v' = x\beta + v = a$ for *all* $x' \in Supp(X)$. That is, we only identify $q(v, x)$ at the hypothetical index $v + x\beta = a$ satisfying the large support Assumption 2.

Assumption 2 (*Large Support*)**.** For a given a and conditional on Z, the support of V covers the support of $a - X\beta$.

Proposition 1 (*Identification*)**.** *Consider a hypothetical value (v, x) and define $a \equiv v + x\beta$. Let Assumptions 1 and 2 hold. Denote $p^*(v, x, z) \equiv \mathbb{E}[D|V = v, X = x, Z = z]$. Assuming β is identified,[4] then the ASF is identified by*

$$\begin{aligned}
q(v,x) &= \int_{Supp(X,Z)} \mathbb{E}[D|V = a - x'\beta, X = x', Z = z]\, dF_{XZ}(x', z) \\
&= \mathbb{E}\left[p^*(a - X\beta, X, Z)\right] \equiv Q(a).
\end{aligned}$$

Assuming the conditional cumulative distribution function $F_{\varepsilon|VXZ}(\varepsilon|v, x, z)$ is differentiable in v, the partial effect of V is identified by

$$\begin{aligned}
&\frac{\partial q(v,x)}{\partial v} \\
&= \int_{Supp(X,Z)} \frac{\partial}{\partial v} \mathbb{E}[D|V = v, X = x', Z = z]\Big|_{v = a - x'\beta} dF_{XZ}(x', z) \\
&\equiv Q_v(a)
\end{aligned}$$

and the partial effect of X is identified by $\partial q(v, x)/\partial x^\top = \beta Q_v(a) \equiv Q_x(a)$. The AIF is

$$\mathbb{E}[D|V + X\beta = a] = \mathbb{E}\left[p^*(a - X\beta, X, Z)\frac{f_{V|XZ}(a - X\beta|X, Z)}{\mathbb{E}\left[f_{V|XZ}(a - X\beta|X, Z)\right]}\right].$$

Proposition 1 suggests sufficient conditions for the AIF to be equal to the ASF. The first condition is that $f_{V|XZ}(a - X\beta|X, Z)$ is constant over the support of (X, Z), which implies that the index $V + X\beta$ is a sufficient statistic for the ASF for any given a. The second sufficient condition is known in the literature that (X, Z) is exogenous so that the conditional ASF given (X, Z) in (1) does not depend on (X, Z). The identification of the ASF in Proposition 1 has been shown in Chen et al. (2016). We propose logistic series estimators for $Q(a)$, $Q_v(a)$, and $Q_x(a)$ and derive their limiting distributions in the next section.

3. Estimation and inference

We estimate the ASF for a given value (v, x) and the partial effects by the following procedure. Let the observations $\{(D_i, S_i)\}_{i=1}^n$, where $S_i \equiv (V_i, X_i, Z_i)$ and the first coordinate in X_i and Z_i is one.

[2] For example, see Lewbel (2000) and Dong and Lewbel (2015) for mean-independent errors and Chen et al. (2016) for median-independent errors and conditionally symmetric errors.

[3] This is in the same spirit with Vytlacil and Yildiz (2007) for the dummy endogenous variable in weakly separable models.

[4] Under Assumption 1 and other regularity conditions, Lewbel (2000) further assumes a large support assumption that the support of V given Z covers the support of $-(X\beta + \varepsilon)$ and contains 0. Magnac and Maurin (2007) characterizes the conditions under which these assumptions are justified. With conditionally symmetric errors, Chen et al. (2016) find that a special regressor improves the identifying power and propose a weighted least absolute deviation estimator and a kernel-weighted least squares estimator.