CORRECTING THE FIXED-EFFECT ESTIMATOR FOR ENDOGENOUS SWITCHING

FERNANDO B. BOTELHO AND VLADIMIR P. PONCZEK
Correcting the fixed-effect estimator for endogenous switching

Fernando B. Botelho and Vladimir P. Ponczek

RESUMO

PALAVRAS CHAVES

CLASSIFICAÇÃO JEL

C33; C35; C51

ABSTRACT

In this paper, we propose a two-step estimator for panel data models in which a binary covariate is endogenous. In the first stage, a random-effects probit model is estimated, having the endogenous variable as the left-hand side variable. Correction terms are then constructed and included in the main regression.

KEY WORDS

Panel data, Fixed-effect estimator, Endogenous switching
The most important advantages of the fixed-effect estimator for panel data models is the ability to control for unobservable attributes which are constant over time. In most cases, it delivers unbiased estimates and is easily implemented. Nevertheless, it is still possible to have unobservable variables that are not constant over time but are correlated with another covariate. In this scenario, the fixed-effect estimator will not deliver unbiased estimators.

We propose a two-step estimator for panel data models in which a binary covariate is endogenous. In the first stage, a random-effects probit model is estimated, having the endogenous variable as the left-hand side variable. Correction terms are then constructed and included in the main regression.

The literature has focused on the problem related to the estimation of panel models with selectivity bias. [?] develop a test to check the presence of selectivity bias based on Heckman-type tests. [?] present two-step estimators for a range of parametric panel models, which encompass the problems addressed by [?]. Different from the previous cited works that rely on strong assumptions of the error and individual fixed-effect terms, [?] relaxes these assumptions. She follows [?] procedures and proposes a two-step semi-parametric estimator, which ‘differences out’ both individual fixed-effect and sample selection bias. [?] has an excellent survey about the sample selection literature, including panel data models. Our paper contributes to the literature by adapting the two-step parametric procedures that deal with selectivity bias to the endogenous switching problem. In this case, the dichotomic endogenous variable is ruled by a choice equation and it is potentially correlated to unobserved characteristics of the individual that are not fixed.
One motivation for our estimator is the measurement the wage differential
between formal and informal jobs. Our main objective is to estimate the effect
of the job status on the hourly earnings. The ordinary fixed-effect estimator
can deliver unbiased estimates if the choice between formal and informal jobs
is determined by some characteristic intrinsic to the worker and constant over
time. Otherwise,

Suppose the variables $w_{it}$ and $b_{it}$ are determined by following equations:

$w_{it} = \theta b_{it} + \beta' X_{it} + \mu_{i} + \epsilon_{it}$ \hfill (1)

$b_{it} = 1(\gamma' Z_{it} + \alpha_{i} + \nu_{it} \geq 0)$. \hfill (2)

$\mu_{i}$ and $\alpha_{i}$ are time-invariant individual effects (possibly correlated with each
other), and $\epsilon_{it}$ and $\nu_{it}$ are pure error terms, possibly correlated with each
other and with the individual effects. Assume that $X_{it}$ is a vector exogenous
variables, and $Z_{it}$ is predetermined. In this case

$$E(w_{it} | X_{i}, Z_{i}, b_{i}) = \theta \tilde{b}_{it} + \beta' \tilde{X}_{it} + E(\epsilon_{it} | X_{i}, Z_{i}, b_{i})$$

Demeaning (??) to eliminate the individual fixed effect, we obtain

$$E(\tilde{w}_{it} | X_{i}, Z_{i}, b_{i}) = \theta \tilde{b}_{it} + \beta' \tilde{X}_{it} + E(\tilde{\epsilon}_{it} | X_{i}, Z_{i}, b_{i}),$$

where $\tilde{a}_{it} = a_{it} - \sum_{t=1}^{T} a_{it}/T$, and $a_{i} = (a_{i1}, \ldots, a_{iT})$.

It is easy to see that the fixed-effect estimator is consistent if

$$E(\tilde{\epsilon}_{it} | X_{i}, Z_{i}, b_{i}) = 0$$

This would be true if $\epsilon_{it}$ and $\nu_{it}$ were uncorrelated. Otherwise, the fixed-
effect estimator will not deliver an unbiased estimate of the degree of seg-
mentation.

In order to deal with that potential source of bias, we propose a correction
term to be added to the main regression such that the fixed-effect estimator
will be shielded from the potential selection bias. Define $u_{i} = \alpha_{i} \iota + \nu_{i}$, where $\iota$ is the unitary column vector of appropriate dimension. To construct
that estimator we need to impose the following structure to the model:

$$\begin{bmatrix}
\epsilon_{i} \\
\mu_{i} \\
\alpha_{i} \iota + \nu_{i}
\end{bmatrix} \sim N\left(0, 
\begin{bmatrix}
\sigma_{\epsilon}^{2} I & 0 & \rho_{\epsilon \nu} \sigma_{\eta} \sigma_{\nu} I \\
0 & \sigma_{\mu} I + \iota' & \rho_{\mu \alpha} \sigma_{\mu} \sigma_{\alpha} I \\
0 & \rho_{\eta \nu} \sigma_{\eta} \sigma_{\nu} I & \rho_{\mu \alpha} \sigma_{\mu} \sigma_{\alpha} I + \sigma_{\nu}^{2} I
\end{bmatrix}\right)$$

2
Under these assumptions, it is straightforward to show that

\[ E(\epsilon_i \mid u_i) = \rho_{\epsilon v} \sigma_{\epsilon} \sigma_v \left[ \sigma^2_{\alpha_i} I + \sigma^2_{v} \right]^{-1} u_i = \frac{\rho_{\epsilon v} \sigma_{\epsilon} \sigma_v}{\sigma^2_v} \left[ I - \frac{\sigma^2_{\alpha_i}}{\sigma^2_v + T \sigma^2_{\alpha_i}} \right] u_i \]

Each element of this vector has the form

\[ E(\epsilon_{it} \mid u_{it}) = \frac{\rho_{\epsilon v} \sigma_{\epsilon} \sigma_v}{\sigma^2_v} \left[ u_{it} - \frac{T \sigma^2_{\alpha_i}}{\sigma^2_v + T \sigma^2_{\alpha_i}} \sum_{s=1}^{T} u_{is} \right]. \]

Now, conditioning on the appropriate interval, we obtain

\[ E(\epsilon_{it} \mid b_{ti}, Z_i) = \frac{\rho_{\epsilon v} \sigma_{\epsilon} \sigma_v}{\sigma^2_v} \left[ E(u_{it} \mid b_{ti}, Z_i) - \frac{T \sigma^2_{\alpha_i}}{\sigma^2_v + T \sigma^2_{\alpha_i}} \sum_{s=1}^{T} E(u_{is} \mid b_{si}, Z_i) \right] \]

Finally, by demeaning the previous equation we have

\[ E(\bar{\epsilon}_i \mid b_{ti}, Z_i) = \frac{\rho_{\epsilon v} \sigma_{\epsilon} \sigma_v}{\sigma^2_v} E(\bar{u}_{it} \mid b_{ti}, Z_i) \]

\[ = \frac{\rho_{\epsilon v} \sigma_{\epsilon} \sigma_v}{\sigma^2_v} E(\bar{v}_{it} \mid b_{ti}, Z_i) \]

This difficulty of this procedure is to calculate \( E(v_{it} \mid b_{ti}, Z_i) \), since \( b_{ti} \) depend of \( \alpha_i \) and the vector \( v_i \). Therefore, one would have to integrate a bivariate normal distribution over the range defined by \( b_{ti} \). However, to avoid this numerically cumbersome procedure, we observe that

\[ E(v_{it} \mid b_{ti}, Z_i) = \int E(v_{it} \mid b_{ti}, Z_i, \alpha_i) f(\alpha_i \mid b_{ti}, Z_i) \, d\alpha_i. \]

It can be seen that

\[ E(v_{it} \mid b_{ti}, Z_i, \alpha_i) = E(v_{it} \mid b_{it}, Z_{it}, \alpha_i) = \frac{(2 b_{it} - 1) \phi(\gamma'Z_{it} + \alpha_i)}{\Phi[(2 b_{it} - 1)(\gamma'Z_{it} + \alpha_i)]}, \]

and the Bayes rule implies

\[ f(\alpha_i \mid b_{ti}, Z_i) = \frac{f(b_{ti} \mid Z_i, \alpha_i) f(\alpha_i \mid Z_i)}{\int f(b_{ti} \mid Z_i, \alpha'_i) f(\alpha'_i \mid Z_i) \, d\alpha'_i} \]
One of the advantages of this procedure is to allow us to perform two one-dimensional instead of one two-dimensional numerical integrations.

Our method consists in estimating the model in two steps. First, we have to estimate $\gamma$ and $\sigma_{\alpha}$ consistently. This is done by a random effect probit in the selection equation. Notice that an important for the consistency of those estimators in the first step is

$$f(\alpha_i \mid Z_i) = f(\alpha_i).$$

In words, it means that fixed effect in the selection equation ($\alpha_i$) is orthogonal to the vector of instruments ($Z_i$). Therefore,

$$\int f(b_i \mid Z_i, \alpha_i) \, f(\alpha_i) \, d\alpha_i = \frac{\Phi[(2 \, b_{it} - 1)(\gamma'Z_{it} + \alpha_i)] \, \frac{1}{\sqrt{2\pi\sigma_{\alpha}^2}} \, \exp \left(-\frac{1}{2} \left(\frac{\alpha_i}{\sigma_{\alpha}}\right)^2\right)}{\int \Phi[(2 \, b_{it} - 1)(\gamma'Z_{it} + \alpha_i)] \, \frac{1}{\sqrt{2\pi\sigma_{\alpha}^2}} \, \exp \left(-\frac{1}{2} \left(\frac{\alpha_i}{\sigma_{\alpha}}\right)^2\right) \, d\alpha_i},$$

so the correction term is given by

$$E(v_{it} \mid b_i, Z_i) = \int \frac{(2 \, b_{it} - 1) \, \phi(\gamma'Z_{it} + \alpha_i)}{\Phi[(2 \, b_{it} - 1)(\gamma'Z_{it} + \alpha_i)]} \times \frac{\Phi[(2 \, b_{it} - 1)(\gamma'Z_{it} + \alpha_i)] \, \frac{1}{\sqrt{2\pi\sigma_{\alpha}^2}} \, \exp \left(-\frac{1}{2} \left(\frac{\alpha_i}{\sigma_{\alpha}}\right)^2\right)}{\int \Phi[(2 \, b_{it} - 1)(\gamma'Z_{it} + \alpha_i)] \, \frac{1}{\sqrt{2\pi\sigma_{\alpha}^2}} \, \exp \left(-\frac{1}{2} \left(\frac{\alpha_i}{\sigma_{\alpha}}\right)^2\right) \, d\alpha_i} \times \Phi[(2 \, b_{it} - 1)(\gamma'Z_{it} + s \, r_i)] \, \frac{1}{\sqrt{2\pi\sigma_{\alpha}^2}} \, \exp \left(-\frac{1}{2} \left(\frac{\alpha_i}{\sigma_{\alpha}}\right)^2\right) \, dr_i.$$

In order to have a standard normal distribution of the fixed effect, we have to redefine the integration variable as

$$r_i = \frac{\alpha_i}{\sqrt{2\sigma_{\alpha}^2}} \rightarrow \alpha_i = s \, r_i.$$

Finally, the correction term can be written as

$$E(v_{it} \mid b_i, Z_i) \approx \sum_{r_i \in R} \frac{(2 \, b_{it} - 1) \, \phi(\gamma'Z_{it} + s \, r_i)}{\Phi[(2 \, b_{it} - 1)(\gamma'Z_{it} + s \, r_i)]} \times \frac{\Phi[(2 \, b_{it} - 1)(\gamma'Z_{it} + s \, r_i)] \, \frac{1}{\sqrt{2\pi\sigma_{\alpha}^2}} \, \exp \left(-\frac{1}{2} \left(\frac{\alpha_i}{\sigma_{\alpha}}\right)^2\right)}{\sum_{r'_i \in R} \Phi[(2 \, b_{it} - 1)(\gamma'Z_{it} + s \, r'_i)] \, \frac{1}{\sqrt{2\pi\sigma_{\alpha}^2}} \, \exp \left(-\frac{1}{2} \left(\frac{\alpha_i}{\sigma_{\alpha}}\right)^2\right)} \exp(-r_i^2)\, dr_i.$$

The last term is a numerical approximation obtained by the Gauss-Hermite numeric integration method.
We then introduce the correction term into the main equation to eliminate the possible endogeneity of $b_{it}$. Let us define $corr_i$ as the vector with correction terms and $W_i = [b_i X_i corr_i]$. In vector notation, equation ?? augmented by the correction term becomes:

$$w_i = \phi' W_i + \mu_i + \xi_i,$$

where $\phi = (\theta, \beta, \lambda)$. Also we define the following terms\(^1\):

$$M = N^{-1} \sum_{i=1}^{N} E[W_i'W_i]$$
$$V = N^{-1} \sum_{i=1}^{N} E[W_i'Var(\xi)W_i]$$
$$D = N^{-1} \sum_{i=1}^{N} E[W_i' \partial \lambda(\gamma) \partial \gamma]$$

$\phi$ is asymptotically normally distributed with covariance matrix\(^2\):

$$\lim_{n \to \infty} M^{-1}(V + DHD')M^{-1}$$

where $H$ is the Hessian matrix generated in the maximum likelihood estimation in the first stage.

In the absence of the endogeneity problem, $\lambda$ is zero and $DH'D'$ is also zero, thus, the covariance matrix becomes trivial. A straightforward way to check whether $b_{ij}$ is endogenous is to test the statistical significance of $\lambda$.

---

\(^1\)We are assuming that $\xi_i$ has spherical variance.

\(^2\)Following [?]

5