JETEM

Journal of Economic Theory and Econometrics

The Journal of the Korean Econometric Society

(계 량 경 제 학 보)

March 2008

Volume 19. Issue 1. pp.49-59

Asymptotic Variance and Extensions of A Density-Weighted-Response Semiparametric Estimator

Myoung-Jae Lee* **· Fali Huang**† **· Young-sook Kim**‡

Abstract Building on some early works, Lewbel (2000) proposed estimators for binary and ordered discrete response models with endogenous regressors. These estimators have been extended for panel data and for truncated and censored models by later papers. The estimators are particularly innovative in that the latent linear regression functions are pulled out of the nonlinear limited dependent variable models, which are then treated as if they were the usual linear models. But understanding the estimators and their applications have been “hampered” by less-than-ideal expositions and assumptions. For this problem, this short note reviews the estimators and makes the following three points. First, the derivation and proper insight of the asymptotic variances are provided. Second, the inefficiency of the ordered discrete response version is pointed out and corrected. Third, assumptions in the panel data extension by Honoré and Lewbel (2002) are relaxed.

Keywords semiparametrics, two-stage estimation, binary response, ordered discrete response, panel data

JEL Classification C14, C33, C35

* Corresponding author. Department of Economics, Korea University, Seoul, Korea 136-701, E-mail: myoung-jae@korea.ac.kr, Phone: (82) 2-3290-2229. This research has been supported by a Korea University Grant.

† School of Economics, Social Sciences, Singapore Management U. 90 Stamford Road, Singapore 178903, E-mail: huang@smu.edu.sg

‡ Institute for National Health Insurance, National Health Insurance Corporation, 168-9, Yeomri-dong, Mapo-gu, Seoul, Korea 121-749, E-mail: yskim@nhic.or.kr

Received January 20, 2008, Accepted March 15, 2008

1. Introduction

Lewbel (2000) presented a semiparametric estimator for binary-response models with endogenous regressors, and then suggested extensions for ordered discrete response (ODR) models as well as for other models. The estimator is based on an innovative idea of pulling the regression function out of the indicator function in the binary-response models, which builds on earlier ideas in Horowitz and Härdle (1996) and Lewbel (1998). The binary response estimator has been applied in Maurin (2002), and the ODR extension in Anton et al. (2001) and Stewart (2005). Honoré and Lewbel (2002) extended the estimator to panel binary responses, and Khan and Lewbel (2006) to truncated and censored models, although the latter does not look as practical as the former and Lewbel's (2000).

In dealing with endogenous regressors in linear models, there are largely three approaches. The first is instrumental variable estimators (IVE); the second is two-stage LSE analogs replacing the endogenous regressors with their projections on exogenous variables; and the third is control function approaches adding a function to the model to remove the endogeneity. For limited dependent variable (LDV) models, methods to deal with endogenous regressors within semiparametric framework have been slow in coming. Two-stage LSE analogs were proposed by Lee (1996), and a control function approach appeared in Blundell and Powell (2004); Blundell and Powell (2003) reviewed semiparametric approaches for endogenous regressors in LDV models. In the literature, the IVE analogs were perceived to be "hopeless". Lewbel's (2000) estimators were particularly welcome in that they are genuine IVE analogs using transformed response variables.

The Lewbel (2000) binary response estimator is a two-stage estimator with a nonparametric first-stage where the first-stage nuisance parameter is a conditional density. The first-stage estimation error affects the second-stage, but the details of the derivation and insight were not provided there, which led to some difficulties in the aforementioned applications as to be shown later. Also the ODR extension turned out to be an inefficient one.

This paper does three tasks. Section 1 quickly reviews the Lewbel's (2000) estimator to show the right insight for the asymptotic variance to facilitate understanding the estimator's main idea, its extension, and its applications. Section 2 proposes an efficient ODR extension. Section 3 relaxes some assumptions in Honoré and Lewbel's (2002) extension to panel binary responses.

2. Binary-Response Estimator and Asymptotic Variance

Consider a binary response model with a 'special' regressor w, the other regressors x, and an error term u^o: $y_i = 1[\beta_w w_i + x_i'\beta^o + u_i^o > 0]$ where $1[A] = 1$ if A holds and 0 otherwise, and β_w and β^o are unknown parameters. Suppose $sign(\beta_w) > 0$—Lewbel (2000) shows how to estimate $sign(\beta_w)$. Normalizing the model by dividing through by β_w yields

$$y_i = 1[w_i + x_i'\beta + u_i > 0], \quad i = 1, ..., N, \quad \text{where } \beta \equiv \beta^o/\beta_w \text{ and } u_i^o \equiv u_i/\beta_w.$$

Suppose some components of x are endogenous and z is an instrument for x.

Denote the conditional density and distribution function of $w|x, z$ as $f(w|x, z)$ and $F(w|x, z)$; analogous notations will be used for other densities and distribution functions. Define

$$\tilde{y}_i \equiv \frac{y_i - 1[w_i > 0]}{f(w_i|x_i, z_i)} \quad \text{and} \quad \tilde{y}_{Ni} \equiv \frac{y_i - 1[w_i > 0]}{f_N(w_i|x_i, z_i)} \quad \text{where}$$

$$f_N(w_i|x_i, z_i) = \frac{f_N(w_i, x_i, z_i)}{f_N(x_i, z_i)} \text{ is a 'leave-one-out' kernel estimator for } f(w_i|x_i, z_i);$$

for $f(x_i)$, we use $f_N(x_i) \equiv \{(N-1)h^{\dim(x)}\}^{-1} \sum_{j,j\neq i} K\{(x_j - x_i)/h\}$ where K is a kernel and h is a bandwidth, and $f_N(w_i, x_i, z_i)$ and $f_N(x_i, z_i)$ are defined analogously. The subscript i will be often omitted.

The key assumptions for the Lewbel's estimator are

$$(i) : E(zu) = 0 \qquad (ii) : F(u|w, x, z) = F(u|x, z)$$

$$(iii) : \text{the support of } f(w|x, z) \text{ is } [W_l, W_h] \text{ that includes the support of } -x'\beta - u \text{ where } -\infty \leq W_l < 0 < W_h \leq \infty.$$

$F(u, x|w, z) = F(u, x|z)$ is used in Lewbel (2000), but $F(u|w, x, z) = F(u|x, z)$ is adopted here following Honoré and Lewbel (2002), because the latter looks less restrictive than the former. This, however, entails the cost of increasing the curse-of-dimensionality problem as x gets newly conditioned on.

Following Honoré and Lewbel (2002, p.2056) and invoking (i)-(iii), we get

$$E(\tilde{y}|x, z) = x'\beta + E(u|x, z).$$

Multiply this equation by $E(xz')E^{-1}(zz')z$ and take $E(\cdot)$ to get $(E^{-1}(\cdot) \equiv \{E(\cdot)\}^{-1})$

$$E(xz')E^{-1}(zz')E(z\tilde{y}) = E(xz')E^{-1}(zz')E(zx')\beta$$

$$\Longrightarrow \quad \beta = M \cdot E(z\tilde{y}) \quad \text{where } M \equiv \{E(xz')E^{-1}(zz')E(zx')\}^{-1}E(xz')E^{-1}(zz'). \quad (2.1)$$

A consistent estimator for β is

$$b_N = M_N \cdot \frac{1}{N}\sum_i z_i \tilde{y}_{Ni}, \quad \text{where } M_N \text{ is a sample analog for } M.$$

Rewrite $N^{-1}\sum_i z_i\tilde{y}_{Ni}$ as $N^{-1}\sum_i z_i(\tilde{y}_{Ni} - x_i'\beta + x_i'\beta)$. Substituting this into b_N yields $\beta + M_N N^{-1}\sum_i z_i(\tilde{y}_{Ni} - x_i'\beta)$. Hence the main term for the asymptotic variance of $\sqrt{N}(b_N - \beta)$ is $N^{-1/2}\sum_i z_i(\tilde{y}_{Ni} - x_i'\beta)$, and the main question is how $f_N(w_i|x_i, z_i) - f(w_i|x_i, z_i)$ affects the asymptotic distribution.

For a random vector μ_i and some function $m(\cdot,\cdot)$, consider

$$\frac{1}{\sqrt{N}}\sum_i m(\mu_i, \alpha_i) \quad \text{where } \alpha_i \equiv f(x_i).$$

Newey (1994) showed, with $m_a \equiv \partial m/\partial a$ and $a_i \equiv f_N(x_i)$,

$$\frac{1}{\sqrt{N}}\sum_i m(\mu_i, a_i) = \frac{1}{\sqrt{N}}\sum_i [m(\mu_i, \alpha_i) \ + \ E(m_a|x_i)f(x_i) - E\{E(m_a|x)f(x)\}] + o_p(1) \quad (2.2)$$

where $m(\mu_i, \alpha_i)$ is for known α_i and the remaining part is the correction terms accounting for the first-stage estimation error. Newey presented also the correction terms for other types of infinite-dimensional nuisance parameters such as conditional means and score functions. For b_N, the nuisance parameter is a conditional density, for which no formula was provided. As shown next, however, a slight generalization of (2.2) yields the desired correction term with surprising ease.

Let now $\alpha_i = f(w_i|x_i, z_i) = f(w_i, x_i, z_i)/f(x_i, z_i)$. The correction term can be found using the product differentiation rule. First, the correction term for $f(w_i, x_i, z_i)$ (and then multiplied by $f(x_i, z_i)^{-1}$) is, using (2.2)

$$\begin{aligned} & E(\frac{m_a}{f(x,z)}|w_i, x_i, z_i)f(w_i, x_i, z_i) - E\{E(\frac{m_a}{f(x,z)}|w, x, z)f(w, x, z)\} \\ = \ & E(m_a|w_i, x_i, z_i)f(w_i|x_i, z_i) - E\{E(m_a|w, x, z)f(w|x, z)\}. \end{aligned}$$

Second, the correction term for $f(x_i, z_i)^{-1}$ (and then multiplied by $f(w_i, x_i, z_i)$) is

$$\begin{aligned} & E(\frac{-f(w,x,z)m_a}{f(x,z)^2}|x_i, z_i)f(x_i, z_i) - E\{E(\frac{-f(w,x,z)m_a}{f(x,z)^2}|x, z)f(x, z)\} \\ = \ & -E(m_a f(w|x,z)|x,z) + E\{E(m_a f(w|x,z)|x,z)\}. \end{aligned}$$

Putting the correction terms together, the terms with nothing conditioned on get cancelled to leave only

$$E(m_a|w_i, x_i, z_i)f(w_i|x_i, z_i) - E\{E(m_a|w, x, z)f(w|x, z)|x, z\}. \quad (2.3)$$

Apply this formula to $m(\mu_i, a_i) = z_i(\tilde{y}_{Ni} - x_i'\beta)$ where μ_i consists of the elements in $(w_i, x_i', y_i, z_i')'$ and $a_i = f_N(w_i|x_i, z_i)$ to get

$$\frac{1}{\sqrt{N}}\sum_i z_i(\tilde{y}_{Ni} - x_i'\beta) = o_p(1) + \frac{1}{\sqrt{N}}\sum_i [\ z_i(\tilde{y}_i - x_i'\beta)$$

$$-z_i E\{\frac{y - 1[w > 0]}{f(w|x,z)^2}|w_i, x_i, z_i\} f(w_i|x_i, z_i) + z_i E\{E(\frac{y - 1[w > 0]}{f(w|x,z)^2}|w, x, z) f(w|x,z)\ |x_i, z_i\}\ ]$$

$$= \frac{1}{\sqrt{N}}\sum_i z_i[\tilde{y}_i - x_i'\beta - \{E(\tilde{y}|w_i, x_i, z_i) - E(\tilde{y}|x_i, z_i)\}] + o_p(1) \qquad (2.4)$$

$$= \frac{1}{\sqrt{N}}\sum_i \zeta_i + o_p(1) \quad \text{where } \zeta_i \equiv z_i[\tilde{y}_i - x_i'\beta - \{E(\tilde{y}|w_i, x_i, z_i) - E(\tilde{y}|x_i, z_i)\}].$$

Note that $E(\zeta) = 0$. Hence $\sqrt{N}(b_N - \beta)$ is asymptotically normal with the variance $ME(\zeta\zeta')M$.

The influence function ζ shows clearly that $\tilde{y}_i - x_i'\beta$ is the term with α_i known, and the correction term is $E(\tilde{y}|w_i, x_i, z_i) - E(\tilde{y}|x_i, z_i)$ that carries a negative sign because α_i appears in the denominator of $\tilde{y}_i$. Lewbel (2000, p.156) defines

$$q \equiv z\tilde{y} + zE(\tilde{y}|x,z) - zE(\tilde{y}|w,x,z) \qquad (\textit{notations modified}) \qquad (2.5)$$

to call this non-zero-mean term an 'influence function', and then introduces $q - zx'\beta$ which is the same as ζ. Our derivation thus agrees with this, but shows better what is going on in the asymptotic variance. The rather unnatural decomposition of ζ with q led to the following statement in Stewart (2005, p.559) for ODR with $x = z$: "the variance of b_N can be estimated as the Huber/White sandwich estimator of the coefficient estimates from an OLS regression of $\tilde{y}_i + E(\tilde{y}_i|x_i) - E(\tilde{y}_i|w_i, x_i)$ on x_i (*notations modified*)." Here, the proper insight of the ζ-decomposition is lost, and q with z removed is proposed to be regressed on x. This scheme may work, but we conjecture that the lack of the proper insight on the asymptotic variance led Stewart (2005) to use a bootstrap instead of the scheme. Maurin (2002) also applied the Lewbel's estimator to find the effect of parental income on repeating the same school grade without showing how the asymptotic variance was computed.

3. Extension for ODR

Consider an ODR model with R-many categories $(0, 1, ..., R-1)$:

$$y_i = \sum_{r=1}^{R-1} 1[w_i + x_i'\beta + u_i \geq \gamma_r] \quad \text{where } \gamma_r\text{'s are unknown thresholds.}$$

Let $x_i = (1, \tilde{x}_i')'$ and $\beta = (\beta_1, \tilde{\beta}')'$, and rewrite the ODR as

$$y_i = \sum_{r=1}^{R-1} y_{ri} \quad \text{where } y_{ri} \equiv 1[\beta_1 - \gamma_r + w_i + \tilde{x}_i'\tilde{\beta} + u_i \geq 0] = 1[y_i \geq r].$$

This is decomposing the ODR y_i into the binary responses y_{ri}, $r = 1, ..., R-1$. The y_r model has intercept $\beta_1 - \gamma_r$ and slope $\tilde{\beta}$. With $M = (M_1', \tilde{M}')'$ where M_1 is the first row of M and $\tilde{M}$ is the remaining rows, we thus get (recall (2.1))

$$M_1 \cdot E(z\tilde{y}_r) = \beta_1 - \gamma_r, \quad r = 1, ..., R-1, \tag{3.1}$$

$$\tilde{M} \cdot E(z\tilde{y}_r) = \tilde{\beta}, \quad r = 1, ..., R-1. \tag{3.2}$$

Since the same $\tilde{\beta}$ satisfies (3.2) for all r, Lewbel (2000, p.161) suggested to use the unweighted average across r: take $(R-1)^{-1}\sum_{r=1}^{R-1}(\cdot)$ on (3.2) to get

$$\tilde{M} \cdot E\left(z\frac{(R-1)^{-1}\sum_{r=1}^{R-1} y_r - 1[w > 0]}{f(w|x,z)}\right) = \tilde{\beta} \tag{3.2'}$$

as a way to extend the binary-response estimator to ODR. But this suggestion is ad hoc, because any combination $\sum_{r=1}^{R-1}\omega_r y_r$ with $\sum_{r=1}^{R-1}\omega_r = 1$ can be used, not just $(R-1)^{-1}\sum_{r=1}^{R-1} y_r$. The best (i.e., efficient) combination can be found using minimum distance estimation (MDE), which combines the $R-1$ binary response $(y_1, ..., y_{R-1})$ model estimators for $\tilde{\beta}$.

For instance, with $R = 3$ for three-category ODR and $\tilde{k} = \dim(\tilde{\beta})$, (3.1) and (3.2) can be written as

$$\begin{bmatrix} \psi_{11} \\ \psi_{1s} \\ \psi_{21} \\ \psi_{2s} \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & I_{\tilde{k}} \\ 0 & 1 & 0 \\ 0 & 0 & I_{\tilde{k}} \end{bmatrix} \begin{bmatrix} \beta_1 - \gamma_1 \\ \beta_1 - \gamma_2 \\ \tilde{\beta} \end{bmatrix} \tag{3.3}$$

where the left-hand side is the 'reduced form (RF)' parameters: $\psi_1 \equiv (\psi_{11}, \psi_{1s}')'$ for y_{1i} and $\psi_2 \equiv (\psi_{21}, \psi_{2s}')'$ for y_{2i} such that ψ_{11} and ψ_{21} are the intercepts and ψ_{1s} and ψ_{2s} are the slopes. Defining the middle matrix as C, the efficient MDE for the right-hand side 'structural form (SF)' parameter $\Gamma \equiv (\beta_1 - \gamma_1, \beta_1 - \gamma_2, \tilde{\beta}')'$ is obtained by minimizing

$$\{(\hat{\psi}_1', \hat{\psi}_2')' - C\Gamma\}'\Omega_N^{-1}\{(\hat{\psi}_1', \hat{\psi}_2')' - C\Gamma\}$$

with respect to Γ where $\hat{\psi}_1$ and $\hat{\psi}_2$ are the RF estimators and Ω_N is their asymptotic variance estimator. The resulting efficient MDE is (see. e.g., Lee (2002) and the references therein)

$$\hat{\Gamma} \equiv (C'\Omega_N^{-1}C)^{-1}C'\Omega_N^{-1}(\hat{\psi}_1', \hat{\psi}_2')' \quad \text{with}$$

$$\sqrt{N}(\hat{\Gamma} - \Gamma) \rightsquigarrow N\{0,\ (C\Omega^{-1}C)^{-1}\}, \quad \text{where } \Omega_N \to^p \Omega.$$

To obtain Ω_N, we need influence functions for $(\hat{\psi}_1', \hat{\psi}_2')'$. Define $\tilde{y}_{ji}$ as the transformed response for y_{ji}, $j = 1, 2$, and

$$\zeta_{ji} \equiv z_i[\tilde{y}_{ji} - (\beta_1 - \gamma_j) - \tilde{x}_i'\tilde{\beta} - \{E(\tilde{y}_j|w_i, x_i, z_i) - E(\tilde{y}_j|x_i, z_i)\}], \quad j = 1, 2.$$

Let $\hat{\zeta}_{ji}$ be a sample analog for ζ_{ji}. Then an estimator of the influence function for $\hat{\psi}_j$ is

$$\hat{\xi}_{ji} = M_N\hat{\zeta}_{ji},\ j = 1, 2 \implies \hat{\psi}_j = \frac{1}{\sqrt{N}}\sum_i \hat{\xi}_{ji} + o_p(1) \text{ and } \Omega_N = N^{-1}\sum_i (\xi_{1i}', \xi_{2i}')'(\xi_{1i}', \xi_{2i}').$$

Instead of the MDE with (3.3), suppose we use (3.1) and (3.2') as suggested by Lewbel (2000), which are

$$\psi_{11} = \beta_1 - \gamma_1, \quad \psi_{21} = \beta_1 - \gamma_2, \quad \frac{\psi_{1s} + \psi_{2s}}{2} = \tilde{\beta}.$$

The Lewbel's unweighted average amounts to replacing Ω_N in the MDE with I_N, which yields an inefficient MDE. This can be seen in that the inefficient MDE is

$$(\hat{\psi}_{11}, \hat{\psi}_{21}, (\hat{\psi}_{1s}' + \hat{\psi}_{1s}')/2)' = (C'C)^{-1}C'(\hat{\psi}_1', \hat{\psi}_2')'.$$

Since the number of observations falling in each ODR category can vary much across $r = 0, ..., R-1$, the efficiency of the binary response estimators will vary much accordingly. The unweighted averaging can easily result in a rather inefficient ODR estimator. Unfortunately, Anton et al. (2001) applied the inefficient MDE to a grouped unemployment duration data.

4. Extension for Panel Binary Response

Honoré and Lewbel (2002) extended the binary response estimator to a panel data model

$$y_{it} = 1[w_{it} + x_{it}'\beta + v_{it} > 0], \quad v_{it} = \delta_i + u_{it}, \quad t = 0, 1$$

where δ_i is a time-constant error possibly related to x_{it}. Let z_i be a time-constant instrument vector. In this section, we briefly review Honoré and Lewbel (2002) and relax some assumptions there.

Assume

$(i)'$: $E(u_{it}z_i) = 0,\ t = 0, 1$ $\quad (ii)' : F(v_{it}|w_{it}, x_{it}, z_i) = F(v_{it}|x_{it}, z_i)$

$(iii)'$: the support of $f(w_{it}|x_{it}, z_i)$ is $[W_{lt}, W_{ht}]$ that includes the support of $-x_{it}'\beta - v_{it}$ where $-\infty \leq W_{lt} < 0 < W_{ht} \leq \infty$.

These assumptions do not require x_{it} to be 'strictly exogenous'; pre-determinedness is sufficient allowing feedbacks from the past and current responses to the future regressors.

Define now

$$\tilde{y}_{it} \equiv \frac{y_{it} - 1[w_{it} > 0]}{f(w_{it}|x_{it}, z_i)} \quad \text{and} \quad \tilde{y}_{Nit} \equiv \frac{y_{it} - 1[w_{it} > 0]}{f_N(w_{it}|x_{it}, z_i)}.$$

Then it holds that

$$E(\tilde{y}_{it}|x_{it}, z_i) = x'_{it}\beta + E(\delta_i + u_{it}|x_{it}, z_i).$$

Multiply this by z_i and take $E(\cdot)$ to get

$$E\{E(z_i\tilde{y}_{it}|x_{it}, z_i)\} = E(z_i x'_{it})\beta + E\{E(z_i\delta_i + z_i u_{it}|x_{it}, z_i)\}$$

$$\Longrightarrow \quad E(z_i\tilde{y}_{it}) = E(z_i x'_{it})\beta + E(z_i\delta_i), \quad \text{for } E(z_i u_{it}) = 0 \text{ in } (i)'. \tag{4.1}$$

Define $\Delta\tilde{y}_i \equiv \tilde{y}_{i1} - \tilde{y}_{i0}$ and $\Delta x_i \equiv x_{i1} - x_{i0}$. First-difference (4.1) to get

$$E(z_i\Delta\tilde{y}_i) = E(z_i\Delta x'_i)\beta \Longrightarrow \beta \equiv [E(\Delta x'_i z_i)E^{-1}(z_i z'_i)E(z_i\Delta x'_i)]^{-1}E(\Delta x'_i z_i)E^{-1}(z_i z'_i)E(z_i\Delta\tilde{y}'_i).$$

The asymptotic variance of the sample analog b_N for β takes the usual form $QE(\nu\nu')Q$ where Q is the matrix before $E(z_i\Delta\tilde{y}'_i)$ in β. To determine ν, proceed as in (2.4) to obtain

$$\begin{aligned}
&\frac{1}{\sqrt{N}}\sum_i(z_i\Delta\tilde{y}_{Ni} - z_i\Delta x'_i\beta) = o_p(1) + \frac{1}{\sqrt{N}}\sum_i[\ z_i(\Delta\tilde{y}_i - \Delta x'_i\beta) \\
&-z_iE\{\frac{y_{i1} - 1[w_{i1} > 0]}{f(w_{i1}|x_{i1}, z_i)}|w_{i1}, x_{i1,}z_i\} + z_iE(E\{\frac{y_{i1} - 1[w_{i1} > 0]}{f(w_{i1}|x_{i1}, z_i)}|w_{i1}, x_{i1,}z_i\}\ |x_{i1}, z_i) \\
&+z_iE\{\frac{y_{i0} - 1[w_{i0} > 0]}{f(w_{i0}|x_{i0}, z_i)}|w_{i0}, x_{i0,}z_i\} - z_iE(E\{\frac{y_{i0} - 1[w_{i0} > 0]}{f(w_{i0}|x_{i0}, z_i)}|w_{i0}, x_{i0,}z_i\}\ |x_{i0}, z_i)\ ] \\
= \ &\frac{1}{\sqrt{N}}\sum_i z_i[\Delta\tilde{y}_i - \Delta x'_i\beta - \{E(\tilde{y}_{i1}|w_{i1}, x_{i1}, z_i) - E(\tilde{y}_{i1}|x_{i1}, z_i)\} \\
&+\{E(\tilde{y}_{i0}|w_{i0}, x_{i0}, z_i) - E(\tilde{y}_{i0}|x_{i0}, z_i)\}] + o_p(1).
\end{aligned}$$

The term $\Delta\tilde{y}_i - \Delta x'_i\beta$ is for the known nuisance parameters, and the rest is the correction terms, whereas Honoré and Lewbel (2002, p.2057) list the unnatural decomposition analogous to q in (2.5).

Honoré and Lewbel (2002, p.2055) state "z_i would typically consist of predetermined regressors up to period 0" (*modified slightly*). This may not give enough instruments in practice, discouraging the application of the estimator. To relax this apparent restriction on z as well as the restriction $(i)'$ in the following, recall (4.1) to see how $E(z_i\delta_i + z_i u_{it}) = E(z_i\delta_i) + E(z_i u_{it})$ was removed: $E(z_i\delta_i)$ was removed by differencing, and $E(z_i u_{it})$ by $(i)'$.

First, suppose a time-variant z_{it} is used to yield $E(z_{it}v_{it})$. To remove $E(z_{it}v_{it})$ by differencing, it is sufficient to assume the moment stationarity

$$E(z_{i1}v_{i1}) = E(z_{i0}v_{i0}). \tag{4.2}$$

This allows z_{it} to be related to δ_i and u_{it}, and $(i)'$ is not necessary.

Second, let $z_i = (z'_{i0}, z'_{i1})$, which is a way to turn time-variant instruments to time-constant ones. Then we get $E(z_iv_{it}) = E(z_i\delta_i) + E(z_iu_{it})$. $E(z_i\delta_i)$ is removed by differencing, and to remove $E(z_iu_{it})$ by differencing as well, it is sufficient to assume

$$E(z_iu_{i1}) = E(z_iu_{i0}) \iff E(z_{i0}u_{i1}) = E(z_{i0}u_{i0}) \text{ and } E(z_{i1}u_{i1}) = E(z_{i1}u_{i0}). \tag{4.3}$$

This is different from (4.2) in that no δ_i is involved but 'cross-moments' with different time indices appear, whereas (4.2) has to do only with contemporaneous moments.

Although relaxing $(i)'$ with (4.2) or (4.3) may look easy, considering the way Lewbel's (2000) influence function decomposition and ODR extension have been taken too literally in the literature, (4.2) and (4.3) may expand substantially the applicability of Honoré and Lewbel (2002).

5. Conclusions

In this paper, we provided a better insight for the asymptotic variances of Lewbel's (2000) and Honoré and Lewbel's (20002) estimators. Then we showed the efficient extension of Lewbel's (2000) estimator for ordered discrete responses. We also relaxed some assumptions regarding the instrument time-constancy and orthogonality to error terms in Honoré and Lewbel (20002). These contributions are hoped to encourage further applications of the innovative estimators.

References

Anton, A.A., A.F. Sainz, and J. Rodriguez-Poo, 2001, Semiparametric estimation of a duration model, Oxford Bulletin of Economics and Statistics 63, 517-533.

Blundell, R.W. and J.L. Powell, 2003, Endogeneity in nonparametric and semiparametric regression models, Advances in Economics and Econometrics: Theory and Applications, Eighth World Congress, Vol.II, edited by M. Dewatripont, L.P. Hansen, and S.J. Turnovsky, Cambridge University Press.

Blundell, R.W. and J.L. Powell, 2004, Endogeneity in semiparametric binary response models, Review of Economic Studies 71, 655-679.

Honoré, B. and A. Lewbel, 2002, Semiparametric binary choice panel data models without strictly exogenous regressors, Econometrica 70, 2053-2063.

Horowitz, J.L. and W. Härdle, 1996, Direct semiparametric estimation of a single index model with discrete covariates, Journal of the American Statistical Association 91, 1632-1640.

Khan, S. and A. Lewbel, 2006, Weighted and two stage least squares estimation of semiparametric truncated regression models, Econometric Theory, forthcoming.

Lee, M.J., 1996, Nonparametric two stage estimation of simultaneous equations with limited endogenous regressors, Econometric Theory 12, 305-330.

Lee, M.J., 2002, Panel data econometrics: methods-of-moments and limited dependent variables, Academic Press.

Lewbel, A., 1998, Semiparametric latent variable model estimation with endogenous or mismeasured regressors, Econometrica 66, 105-121.

Lewbel, A., 2000, Semiparametric qualitative response model estimation with unknown heteroskedasticity or instrumental variables, Journal of Econometrics 97, 145-177.

Maurin, E., 2002, The impact of parental income on early schooling transitions: a re-examination using data over three generations, Journal of Public Economics 85, 301-332.

Newey, W., 1994, The asymptotic variance of semiparametric estimators, Econometrica 62, 1349-1382.

Stewart, M.B., 2005, A comparison of semiparametric estimators for the ordered response model, Computational Statistics and Data Analysis 49, 555-573.