Economics Letters 105 (2009) 181–182

Contents lists available at ScienceDirect

Economics Letters

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

The incidental parameter problem in a non-differentiable panel data model

Bryan S. Graham [a,1], Jinyong Hahn [b,*], James L. Powell [a,2]

[a] Department of Economics, University of California–Berkeley, Berkeley, CA 94720–3880, United States
[b] Department of Economics, UCLA, Mail Stop: 147703, Los Angeles, CA 90095, United States

ARTICLE INFO

Article history:
Received 10 November 2008
Received in revised form 21 July 2009
Accepted 23 July 2009
Available online 3 August 2009

Keywords:
Panel
Quantile
Incidental parameter problem

JEL Classification:
C33

ABSTRACT

We consider a panel quantile model with fixed effects. It is shown that the maximum likelihood estimator is numerically equivalent to the least absolute deviations estimator of the differenced model, and as a consequence, there is no incidental parameter problem.

1. Introduction

Although panel data in principle allow researchers to control for unobserved individual heterogeneity, obvious methods such as maximum likelihood estimation (MLE) may be problematic due to the well-known incidental parameter problem (cf., Neyman and Scott, 1948). Because of the individual-specific fixed effects, the total number of parameters in these models equals the number of individuals plus the dimension of the common parameter. When the number of individuals (n) is large relative to the time series dimension (T), the maximum likelihood estimator (MLE) typically results in inconsistent estimates of the common parameter of interest.

In linear panel models with homoscedastic normal errors and strictly exogenous regressors, however, it is well-known that the MLE does not suffer from the incidental parameter problem. The MLE is numerically equivalent to the within-group estimator, which is consistent even when the cross section dimension n increases to infinity with the time series dimension T fixed. Because the MLE is usually inconsistent with T fixed asymptotics, it is of interest to document other models with similar properties. In this note, we present one such model.

We consider a density on the error term such that the resultant log likelihood is equivalent to the objective function of a quantile regression. The objective function is nondifferentiable, and the asymptotic properties of the MLE are not immediately obvious. We show that the MLE for this model is consistent with T fixed asymptotics as long as the regressors are strictly exogenous.

Our result is closely related to Honoré (1992), who developed a moment condition for the censored panel model with fixed effects. In an appendix, he provides an alternative derivation of his moment condition as a result of maximizing the likelihood over the fixed effects, although this numerical equivalence is not emphasized elsewhere. We interpret his paper as a justification of MLE in censored panel models, and develop a similar analysis for quantile models.

2. Main result

For simplicity of exposition, we only consider the case with $T=2$. Suppose that we are given a linear panel model with fixed effects

$$y_{it} = \alpha_i + x_{it}'\beta + \varepsilon_{it},$$

where ε_{it} is iid with density equal to

$$f(e) = \begin{cases} \left(\frac{1}{\tau} + \frac{1}{1-\tau}\right)^{-1} exp(-\tau|e|) & \text{if } e>0 \\ \left(\frac{1}{\tau} + \frac{1}{1-\tau}\right)^{-1} exp(-(1-\tau)|e|) & \text{if } e<0 \end{cases}$$

* Corresponding author. Tel.: +1 310 825 2523; fax: +1 310 825 9528.
E-mail addresses: bgraham@econ.berkeley.edu (B.S. Graham), hahn@econ.ucla.edu (J. Hahn), powell@econ.berkeley.edu (J.L. Powell).
[1] Tel.: +1 510 642 4752.
[2] Tel.: +1 510 643 0709.

0165-1765/$ – see front matter © 2009 Elsevier B.V. All rights reserved.
doi:10.1016/j.econlet.2009.07.015

The x is assumed to be strictly exogenous, i.e., (x_{i1}, x_{i2}) is independent of $(\varepsilon_{i1}, \varepsilon_{i2})$. The MLE solves

$$\min_{a_1,\ldots,a_n,b} \sum_{i=1}^{n} \sum_{t=1}^{2} \left\{ \tau(y_{it} - a_i - x'_{it} b)^+ + (1-\tau)(y_{it} - a_i - x'_{it} b)^- \right\}. \quad (1)$$

Note that the log likelihood is numerically equivalent to the ('check') objective function of a quantile regression model.

Our main result is a simple corollary of the observation that, regardless of τ, the concentrated likelihood function is numerically equivalent to the least absolute deviations model of $y_{i2} - y_{i1}$ on $x_{i2} - x_{i1}$:

Theorem 1.

$$\min_{a_1,\ldots,a_n} \sum_{i=1}^{n} \sum_{t=1}^{2} \left\{ \tau(y_{it} - a_i - x'_{it} b)^+ + (1-\tau)(y_{it} - a_i - x'_{it} b)^- \right\} = \sum_{i=1}^{n} |(y_{i2} - y_{i1}) - (x_{i2} - x_{i1})' b|.$$

Proof. We omit the i subscript whenever doing so does not cause confusion. The function

$$g(a,b) \equiv \tau(y_1 - a - x'_1 b)^+ + (1-\tau)(y_1 - a - x'_1 b)^- + \tau(y_2 - a - x'_2 b)^+ + (1-\tau)(y_2 - a - x'_2 b)^-$$

is continuous and piecewise linear. Assume without loss of generality that $y_1 - x'_1 b \le y_2 - x'_2 b$. We then have

$$\frac{\partial g(a,b)}{\partial a} = \begin{cases} -2\tau & \text{if } a < y_1 - x_1 b \\ (1-2\tau) & \text{if } y_1 - x_1 b < a < y_2 - x_2 b \\ 2(1-\tau) & \text{if } a > y_2 - x_2 b \end{cases}$$

Now, if we minimize $g(a,b)$ over a with b fixed, the minimum is attained when

$$a = \begin{cases} y_1 - x'_1 b & \text{if } \tau \le \frac{1}{2} \\ y_2 - x'_2 b & \text{if } \tau \ge \frac{1}{2} \end{cases}$$

The minimized value is then equal to

$$\tau((y_2 - x'_2 b) - (y_1 - x'_1 b))$$

when $\tau \le \frac{1}{2}$ and

$$(1-\tau)((y_2 - x'_2 b) - (y_1 - x'_1 b))$$

when $\tau \ge \frac{1}{2}$. We can therefore conclude that the concentrated log likelihood is proportional to

$$\sum_{i=1}^{n} |(y_{i2} - y_{i1}) - (x_{i2} - x_{i1})' b|$$

□

Theorem 1 has an interesting implication that, even if we start with an arbitrary panel quantile regression problem (1), the concentrated log likelihood is always the least absolute deviations problem

$$\min_b \sum_{i=1}^{n} |(y_{i2} - y_{i1}) - (x_{i2} - x_{i1})' b| \quad (2)$$

Under our parametric specification, the differenced model

$$y_{i2} - y_{i1} = (x_{i2} - x_{i1})' \beta + (\varepsilon_{i2} - \varepsilon_{i1})$$

is such that its error $\varepsilon_{i2} - \varepsilon_{i1}$ has a median at zero. It follows that the b that solves Eq. (2) is consistent under fixed T asymptotics, hence there is no incidental parameter problem as long as the regressors are strictly exogenous.

We note that the error $\varepsilon_{i2} - \varepsilon_{i1}$ has a median at zero because the ε_{it} is iid over t. Therefore, the consistency result holds even when the density $f(e)$ is misspecified as long as the iid assumption is satisfied.

References

Honoré, Bo E., 1992. Trimmed lad and least squares estimation of truncated and censored regression models with fixed effects. Econometrica 60 (3), 533–565.

Neyman, J., Scott, E.L., 1948. Consistent estimation from partially consistent observations. Econometrica 16 (1), 1–32.