UNIVERSITY of York

Discussion Papers in Economics

No. 22/05

Equivalence of Alternative Specifications of Nonlinear Panel Data Models

Peter J. Simmons

Department of Economics and Related Studies
University of York
Heslington
York, YO10 5DD

Equivalence of Alternative Specifications of Nonlinear Panel Data Models*

Peter J. Simmons†
University of York (UK)

September 8, 2022

Panel data models are widely used with mixes of continuous and ldv models. They may involve simultaneity and nonlinearity between dependent variables and bed written in different ways to reflect this.

This note explores the relation between different specifications of such models under the assumption that any one of them is identified. We deal with three basic models; the endogenous variables are denoted y; the exogenous variables by x_i and the disturbances ε_i which are each zero mean. For simplicity we deal with the bivariate case. The three models are

$$\begin{aligned} Y &= F_1(x) + \varepsilon_1 \\ y_i &= f_{1i}(x, Y_i) + \varepsilon_{2i}, i = 1..n \quad (1) \\ y_i &= f_{3i}(x, Y)_i + \varepsilon_{3i}, i = 1..n \quad (2) \end{aligned}$$

A special case has Y consisting of 0/1 ldv variables.

The main finding is that in nondegenerate cases unless there are incompatible prior restrictions on the distributions of the disturbances, these are all equivalent. So long as the models are each identified it means the parameters and the SE's, etc of the f, F can be found from estimating just the one with simplest estimation program. It also means that tests between the spefications cannot accept one over another in any meangingful way.

To see this:

(1) vs (2): If the Jacobian of Y in (2) is nonsingular we can solve (2) for the y_i yielding (1). Conversely from (1) we just include the relevant y_i on the RHS. If the Jacobian is singular then there is effectively just a single relation in (2).

(1) vs (3): define the four outcome ldv Y by $Y = [(1;1);(1;0);(0;1);(0;0)]$ and rewrite (1) as (3). The distribution of $\varepsilon_i$is the relevant marginal distribution of ε. Conversely from (3) rewrite it as (2) using the marginal distributions and then as (1).

*The author is responsible for all, errors, omissions. I got the idea from listening to Vassilis Hajivassiliou (LSE) talk.

†Department of Economics and Related Studies, University of York, York Y01 5DD. Email: peter.simmons@york.ac.uk

(1) vs (4). From (4) rewrite it as two equations as in (2). Hence all four models are equivalent.

Implications

Typically it will be simpler to estimate (1) and then rewrite the results in the forms of (2),(3) to give equivlent alternative interpretations of the model. If diagnostic and other hypothesis results give a result in one specification then it will be true in each specification. It has very wide applicability to significant questions eg does one persons lifetime earnings profile interact with that of others? at which dates/ages?

It also applies to the case of mixes of continuous and ldv endogenous variables. For example in alternative populations, how does the time profile and changes in marital status of in one population interact with that in others?

Reference

VA Hajivassiliou, PA Ruud, Classical estimation methods for LDV models using simulation, Handbook of econometrics 4, 2383-2441, 627, 1994.