Moment Conditions for Dynamic Panel Logit Models with Fixed Effects

with Martin Weidner (Oxford)

ESEM 2022

Econometric Motivation: Nonlinear Panel Data Models

A dependent variable in time period t, Y_{it}, is modelled as

$$Y_{it} \sim f\left(\cdot \,|\, X_i^T, Y_i^{t-1}, A_i; \theta\right) \quad \text{or} \quad Y_{it} \sim f\left(\cdot \,|\, X_i^t, Y_i^{t-1}, A_i; \theta\right)$$

- Interested in learning about θ from data for a small (finite) number of time periods, T.
- Nonparametric in the relationship between the heterogeneity term, A_i, and X_i^T.
- This is sometimes referred to as a “fixed effects” approach.

The problem: You cannot simply difference A_i away.

So what can we do?

The Focus of This Research (all with Martin Weidner)

Various Logit Models

- Simple logit with lagged dependent variables and strictly exogenous X.
- Ordered logit version (also with Chris Muris).
- Multinomial logit version.
- Simultaneous logit version (also with Hu and Kyriazidou)

Approaches

$$Y_{it} \sim f\left(\cdot \mid X_i^T, Y_i^{t-1}, A_i; \theta\right)$$

- Conditional Likelihood
  - Find (if you can) a sufficient statistic, S_i, for A_i.
  - Maximize likelihood conditional on A_i.
  - Rasch (1960), Andersen (1970),Hausman, Hall, and Griliches (1984), Chamberlain (1985), Magnac (2000), Aguirregabiria, Gu, and Luo (2020), and others.
- Conditional Maximum Score Version of Same Idea
  - Manski (1987), Abrevaya (1999) (kind of), and others.

Approaches (continued)

- Moment conditions
  - Case by case: Honoré (1992), Wooldridge (1997)*, Honoré and Hu (2004)*, Kitazawa (2013), and others
  - More systematic: Bonhomme (2012), Bonhomme and Graham (in progress)*,
- Moment Inequalities
  - Pakes and Porter (2016), Aristodemou (2018), Pakes, Porter, Shepard, and Calder-Wang (in progress) and others
  - Conditional Maximum Score can also be thought of in this way
- Other "Tricks"
  - Chen, Khan and Tang (2019) and others
- Sometimes Apply to "Textbook" Models. Sometime Reverse Engineer Models
  - Al-Sadoon, Li, and Pesaran (2017), Bartolucci and Nigro (2010)

Moment Conditions Bonhomme (2012)

If we can find a function m such that

$$E\left[m\left(Y_i^T, Y_{i0}, X_i^T, \theta\right) \middle| Y_{i0}, X_i^T, A_i\right] = 0$$

then we would have the conditional moment conditions

$$E\left[m\left(Y_i^T, Y_{i0}, X_i^T, \theta\right) \middle| Y_{i0}, X_i^T\right] = 0$$

and the unconditional moments

$$E\left[m\left(Y_i^T, Y_{i0}, X_i^T, \theta\right) g\left(Y_{i0}, X_i^T\right)\right] = 0$$

got any function g such that the moments exist.

But how do we find the moment function, m?

Go back to the model,

$$P\left(Y_{it} = 1 | Y_i^{t-1}, X_i^T, A_i\right) = \frac{\exp\left(Y_{it-1}\gamma + X_{it}'\beta + A_i\right)}{1 + \exp\left(Y_{it-1}\gamma + X_{it}'\beta + A_i\right)}$$

with $T = 3$ (total number of time-periods is 4)

We are looking for (a vector of) function(s), m, such that

$$E\left[m\left(Y_0, Y_1, Y_2, Y_3, X, \gamma_0, \beta_0\right) | X, Y_0, A\right] = 0$$

and hence

$$E\left[m\left(Y_0, Y_1, Y_2, Y_3, X, \gamma_0, \beta_0\right) | X, Y_0\right] = 0$$

for all values, (γ_0, β_0), of the true parameters.

If we knew m, then we could do GMM without worrying about the distribution of A.

This Is Actually Trivial

Since (Y_1, Y_2, Y_3) can take 8 values, we write this explicitly as

$$\sum_{(y_1,y_2,y_3)\in\{0,1\}^3} \Pr\left(Y_1 = y_1, Y_2 = y_2, Y_3 = y_3 \,\middle|\, Y_0 = y_0, X = x, A = \alpha\right) \times m\left(y_0, y_1, y_2, y_3, x, \gamma_0, \beta_0\right) = 0.$$

The model gives the vector of probabilities (conditional on (Y_0, X, A)) as a function of the parameters.

We are looking for the m's (there are 8, and they cannot all be 0).

Proceed Numerically

For concrete values of y_0, x, β, γ and one value of α,

$$\sum_{y_1,y_2,y_3} \Pr\left(Y_1 = y_1,\, Y_2 = y_2,\, Y_3 = y_3 \,\middle|\, Y_0 = y_0, X = x, A = \alpha\right) \times m\left(y_0, y_1, y_2, y_3, x, \gamma_0, \beta_0\right) = 0.$$

gives one equation.

With Q values of α, there will be Q equations.

- Pick $(\alpha_1, \alpha_2, \cdots, \alpha_Q)$. with $Q > 8$. Try to solve for the m's.
- If we can find a non-zero solution (numerically) and it does not depend on the α's then there is hope. This is important. We will quickly know whether something is likely to be possible or not.
- In our case, it did look like it was possible.

But Then What?

Solving

$$\sum_{y_1,y_2,y_3} \Pr\left((Y_1,Y_2,Y_3)=(y_1,y_2,y_3) \mid Y_0=y_0,X=x,A=\alpha_1\right) m\left(y_0,y_1,y_2,y_3,x,\gamma_0,\beta_0\right)=0.$$

$$\sum_{y_1,y_2,y_3} \Pr\left((Y_1,Y_2,Y_3)=(y_1,y_2,y_3) \mid Y_0=y_0,X=x,A=\alpha_2\right) m\left(y_0,y_1,y_2,y_3,x,\gamma_0,\beta_0\right)=0.$$

$$\vdots$$

$$\sum_{y_1,y_2,y_3} \Pr\left((Y_1,Y_2,Y_3)=(y_1,y_2,y_3) \mid Y_0=y_0,X=x,A=\alpha_Q\right) m\left(y_0,y_1,y_2,y_3,x,\gamma_0,\beta_0\right)=0.$$

can be a mess, although there are tricks to make it easier.

For example, ...

We are looking for $m(y_0, y_1, y_2, y_3, x, \gamma_0, \beta_0)$ such that

$$\sum_{y_1,y_2,y_3} \Pr\left((Y_1,Y_2,Y_3)=(y_1,y_2,y_3) \mid Y_0= y_0,X=x,A=\alpha\right) m(y_0, y_1, y_2, y_3, x, \gamma_0, \beta_0) = 0.$$

for all α.

Of course

$$\lim_{\alpha\to\infty} \Pr\left((Y_1,Y_2,Y_3)=(1,1,1) \mid Y_0= y_0,X=x,A=\alpha\right) = 1,$$

$$\lim_{\alpha\to-\infty} \Pr\left((Y_1,Y_2,Y_3)=(0,0,0) \mid Y_0= y_0,X=x,A=\alpha\right) = 1.$$

So it is reasonable to guess that the solution will have

$$\begin{aligned} m(y_0, 1, 1, 1, x, \gamma_0, \beta_0) &= 0 \\ m(y_0, 0, 0, 0, x, \gamma_0, \beta_0) &= 0 \end{aligned}$$

These kind of "tricks" become more relevant for more complicated models.

Moment functions for $T = 3$, where $y = (y_1, y_2, y_3)$

Let $x_{ts} = x_t - x_s$. Define

$$m^{(a)}_{y_0}(y, x, \beta, \gamma) = \begin{cases} \exp\left(x_{12}'\beta + y_0\gamma\right) & \text{if } y = (0,1,0), \\ \exp\left(x_{13}'\beta - (1-y_0)\gamma\right) & \text{if } y = (0,1,1), \\ -1 & \text{if } (y_1, y_2) = (1,0), \\ \exp\left(x_{32}'\beta\right) - 1 & \text{if } y = (1,1,0), \\ 0 & \text{otherwise,} \end{cases}$$

$$m^{(b)}_{y_0}(y, x, \beta, \gamma) = \begin{cases} \exp\left(x_{23}'\beta\right) - 1 & \text{if } y = (0,0,1), \\ -1 & \text{if } (y_1, y_2) = (0,1), \\ \exp\left(x_{31}'\beta - y_0\gamma\right) & \text{if } y = (1,0,0), \\ \exp\left((1-y_0)\gamma + x_{21}'\beta\right) & \text{if } y = (1,0,1), \\ 0 & \text{otherwise.} \end{cases}$$

Key Result: These Moment Conditions Work For All α

THEOREM: If the outcomes $Y = (Y_1, Y_2, Y_3)$ are generated from the above panel logit AR(1) with $T = 3$, then for all $y_0 \in \{0, 1\}$, $x \in R^{K\times 3}$, $\alpha \in R$,

$$E\left[m^{(a)}_{y_0}(Y, X, \beta_0, \gamma_0) \,\middle|\, Y_0 = y_0,\, X = x,\, A = \alpha\right] = 0,$$
$$E\left[m^{(b)}_{y_0}(Y, X, \beta_0, \gamma_0) \,\middle|\, Y_0 = y_0,\, X = x,\, A = \alpha\right] = 0.$$

Are There More? Does This Generalize?

What about $T > 3$? Yes. The expressions are in the paper with Weidner.

What About More Lags? Yes. We also have results about that.

Expressions are in the paper.

Relation to Existing Results

- Kitazawa (2013) also derives valid moment conditions for this model that eliminate A_i. For the AR(1) model with $T = 3$, they coincide with ours.

- For $x_2 = x_3$ the moment conditions are transformations of the first order conditions for maximizing the conditional log-likelihood in Honoré and Kyriazidou (2000) (and Cox (1958) and Chamberlain (1985)).

- Kruiniger (2020) and Dobronyi, Gu, and Kim (2021) proves that for the model with one lag, we have derived all the available moment equality conditions.

- A recent related paper by Dobronyi, Gu, and Kim (2021) also consider moment inequalities to characterize the sharp identified set when the model is not point-identified.

If I Had More Time, I Would Talk About

Natural things to talk about:

- Identification
- Turn the conditional moments into a GMM estimation strategy
- Monte Carlo (It works OK, but there is stuff to be done computationally)
- Empirical Illustration (Employment Status; the results make sense)

Potential weakness:

- This is awfully specific to the logit case.
- So perhaps you should not care after all?

Is It Worth It? (Preliminary)

$$P\left(Y_{it}=1|Y_i^{t-1},X_i^T,A_i\right)=F\left(Y_{it-1}\gamma+X_{it}'\beta+A_i\right)$$

F perhaps not logistic.

Very simple (preliminary) design

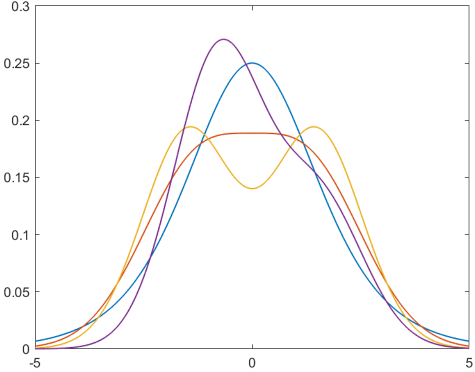
- $T=4$; $k=3$; 100 draws of X's, then probability limit.
- $\gamma=1$, $\beta=\left(\pi/\sqrt{3},0,0\right)'$.
- $A_i\sim N(0,2)$ (Essentially)
- X_{it}'s i.i.d. over time conditional on A_i:
  - $X_1=(N(0,1)+A)/\sqrt{2}-1/4$;
  - $X_2=(N(0,1)+1\{A>0\}-1\{A<0\})/\sqrt{2}$;
  - $X_3=2\cdot(N(0,1)+1\{X_1>0\})-1$.

γ and β are the probability limits of our estimator (one of them)

Bounds calculated using Pakel and Weidner (2022).

Choices of F (Mixtures of Normals: Mean 0, Variance $\pi^2/3$)

Corresponding densities



Marginal Effects ($T = 4$)

$$E\left[P\left(Y_4 = 1 \middle| Y_3^* = 1\right) - P\left(Y_4 = 1 \middle| Y_3^* = 0\right)\right]$$

	Logistic	Flat	Bimodal	Asymmetric
True	0.145	0.150	0.132	0.085
Assume Logit	(0.142, 0.149)	(0.147, 0.155)	(0.131, 0.134)	(0.087, 0.087)
LPM + IV	0.091	0.092	0.089	0.068

$$E\left[\frac{\partial P\left(Y_4 = 1\right)}{\partial x_{41}}\right]$$

	Logistic	Flat	Bimodal	Asymmetric
True	0.257	0.276	0.153	0.152
Assume Logit	(0.247, 0.269)	(0.255, 0.296)	(0.159, 0.160)	(0.154, 0.155)
LPM + IV	0.299	0.319	0.158	0.171

Extensions: Ordered Logit Model

Honoré, Muris and Weidner (under revision) :

$$Y_{it} = \begin{cases} 1 & \text{if } \; Y_{it}^* \in (-\infty, \lambda_1], \\ 2 & \text{if } \; Y_{it}^* \in (\lambda_1, \lambda_2], \\ \vdots & \\ Q & \text{if } \; Y_{it}^* \in (\lambda_{Q-1}, \infty), \end{cases}$$

where

$$Y_{it}^* = X_{it}' \beta + \sum_{q=1}^{Q} \gamma_q \, 1\{Y_{i,t-1} = q\} + A_i + \varepsilon_{it},$$

and ε_{it} is i.i.d. logistically distributed.

Theory and empirical illustration.

Extensions: Multinomial Logit Model

Honoré and Weidner (in progress)

There are $Q \in \{2, 3, 4, \dots\}$ possible for a variable $Y_{it} \in \{1, 2, \dots, Q\}$ and that Y_{it} is generated by the model

$$Y_{it} = \underset{q \in \{1,2,\dots,Q\}}{\operatorname{argmax}} \; U_{qit},$$

for $t \in \{1, 2, \dots, T\}$, with latent variable given by

$$U_{qt} = \sum_{r=1}^{Q} \gamma_{qr} \, 1\{Y_{t-1} = r\} + X_t' \beta_q + A_q + \varepsilon_{qt}.$$

So far only theory.

Extensions: Bivariate Model of Schmidt and Strauss (1975)

Honoré, Hu, Kyriazidou and Weidner (2022) consider the econometric model

$$
\begin{aligned}
& P\left(Y_{1,it} = c_1, Y_{2,it} = c_2 \middle| \left\{ Y_{1,is}, Y_{2,is} \right\}_{s<t}, \left\{ Y_{1,is} \right\}_{s=1}^{T}, \left\{ X_{2,is} \right\}_{s=1}^{T}, A_{1,i}, A_{2,i} \right) \\
= & \frac{\exp\left(c_1\left(Z_{1,it} + A_{1,i}\right) + c_2\left(Z_{2,it} + A_{2,i}\right) + c_1 c_2 \rho\right)}{1 + \exp\left(Z_{1,it} + A_{1,i}\right) + \exp\left(Z_{2,it} + A_{2,i}\right) + \exp\left(Z_{1,it} + A_{1,i} + Z_{2,it} + A_{2,i} + \rho\right)}
\end{aligned}
$$

for $t = 1, 2, 3$ and $c_1, c_2 \in \{0, 1\}$ where

$$
\begin{aligned}
Z_{1,it} &= X'_{1,it}\beta_1 + Y_{1,it-1}\gamma_{11} + Y_{2,it-1}\gamma_{12} \\
Z_{2,it} &= Y'_{2,it}\beta_2 + Y_{1,it-1}\gamma_{21} + Y_{2,it-1}\gamma_{22}
\end{aligned}
$$

(Not as obscure as it looks)

Extensions: Moment Conditions for Other Models?

- Probit? No!
- Mixture of Logits? Yes! For larger T. For example, a static mixture of 56 logits will generate 7 moment conditions when $T = 9$. Compare to Chamberlain (2010), Johnson (2004) and Davezies, D'Haultfoeuille, and Mugnier (2020)
- $Y_{it} \sim \text{binomial}\left(M, \frac{\exp\left(Y_{it-1}\gamma + X_{it}'\beta + A_i\right)}{1+\exp\left(Y_{it-1}\gamma + X_{it}'\beta + A_i\right)}\right)$? Yes, looks like it.
- Multinomial model of Pakes, Porter, Shepard, and Calder-Wang (in progress)?

$$\mathrm{Y}_{it} = \underset{q\in\{1,2,\dots,Q\}}{\operatorname{argmax}} \ \mathrm{U}_{qit},$$

$$U_{qit} = \left(-P_{q,i,t} - 1\{Y_{i,t-1} \neq d\}\kappa\right) B_i + A_{q,i} + \varepsilon_{qit}$$

Does not look like it.

Extensions: Moment Conditions for Entry Games?

- Static entry games? Bresnahan and Reiss (1991), Tamer (2003).

$$\begin{aligned} Y_{1t} &= 1\left\{X_{1t}'\beta - \gamma Y_{2t} + A_1 + \varepsilon_{1t} > 0\right\}, \\ Y_{2t} &= 1\left\{X_{2t}'\beta - \gamma Y_{1t} + A_2 + \varepsilon_{2t} > 0\right\} \end{aligned}$$

Yes, looks like it.

- Dynamic entry games (put in lags above)?

$$\begin{aligned} Y_{1t} &= 1\left\{X_{1t}'\beta - \gamma Y_{2t} + \delta_{self} Y_{1,t-1} - \delta_{other} Y_{2,t-1} + A_1 + \varepsilon_{1t} > 0\right\}, \\ Y_{2t} &= 1\left\{X_{2t}'\beta - \gamma Y_{1t} + \delta_{self} Y_{2,t-1} - \delta_{other} Y_{1,t-1} + A_2 + \varepsilon_{2t} > 0\right\} \end{aligned}$$

Not sure.