Applied Panel Data Analysis – Lecture 11

Christopher F. Parmeter

AGRODEP
March 10-14th, 2015
Addis Ababa, Ethiopia

- We discussed estimation of the unobserved effects model from the standpoint of the conditional mean
- Conditional mean cannot inform us on the shape of the conditional density of y
- Can use a conditional quantile approach
- Recent research proposes a simple, two-step approach to estimate conditional quantiles of the unobserved effects model in the fixed effects framework

- Consider the standard unobserved effects model

$$y_{it} = x_{it}'\beta + c_i + \varepsilon_{it} \tag{1}$$

- Estimation of this model precedes by imposing a conditional mean assumption on ε, such as $E[\varepsilon_{it}|x_{is}] = 0$, $\forall s$ (strict exogeneity)
- However, we can think more generally about the unobserved effects model by focusing on quantiles of the conditional distribution

- Consider an unobserved effects model for the τ^{th} quantile

$$y_{it} = x'_{it}\beta(\tau) + c_i(\tau) + \varepsilon_{it}(\tau) \tag{2}$$

where the following quantile restriction holds $Q_{\varepsilon_{it}}(\tau|x_{is}) = 0$ $\forall s$

- Rosen (2009) showed that this conditional quantile restriction is not enough to identify $\beta(\tau)$

- A simple differencing or within transformation is infeasible to remove the impact of unobserved heterogeneity
- This follows given that quantiles are not linear operators
- In general, $Q_\tau(y_{it} - y_{is}|x_i) \neq Q_\tau(y_{it}|x_i) - Q_\tau(y_{is}|x_i)$

- As noted by Koenker and Hallock (2000): "Quantiles of convolutions of random variables are rather intractable objects, and preliminary differencing strategies familiar from Gaussian models have sometimes unanticipated effects."
- Without being more explicit about the relationship between c_i and x_i, it is difficult to envision an appropriate strategy for dealing with conditional quantiles
- For example, assuming c_i is a pure location shift, i.e. $c_i(\tau) = c_i \; \forall\tau$, Canay (2011) proposes a simple estimator for $\beta(\tau)$

- Canay's (2011) quantile estimator for fixed c_i across the quantiles requires two steps
- The first step is to obtain a consistent estimator for the c_i, this can be done using the standard within estimator
- The second step is to run a pooled quantile regression of $\ddot{y}_{it} = y_{it} - \hat{c}_i$ on x_{it} to obtain $\hat{\beta}(\tau)$
- The estimator works because, given that c_i is fixed across quantiles, that means it is also fixed at the mean, and so the first step helps to obtain a consistent estimator of c_i that can then be used to remove it from the unobserved effects model
- Koenker (2004) proposes a penalized estimator for $\beta(\tau)$ keeping c_i fixed across quantiles, but this involves selecting a penalty parameter which may not be ideal

- Canay (2011) proves that $\hat{\beta}(\tau)$ converges to a mean-zero Gaussian process, however, the limiting covariance function of this Gaussian process is quite complicated
- Thus, a standard wild bootstrap is recommended to conduct inference for $\hat{\beta}(\tau)$

- Wild bootstrap standard errors would be constructed as
  1. Estimate $\beta(\tau)$ and c_i using Canay's two step approach
  2. For a resample $\left\{(y_{11}^*, \hat{\varepsilon}_{11}^*), \ldots, (y_{NT_N}^*, \hat{\varepsilon}_{NT_N}^*)\right\}$ using the wild bootstrap for each individual
  3. Obtain estimates $\hat{\beta}_s^*(\tau)$ using the two step approach
  4. Repeat steps 2 and 3 B times
  5. Reject H_0 : at the α level if $t_j > z_{\alpha/2}^B$ where $t_j = \frac{\hat{\beta}_j(\tau) - \beta_j^o}{s.e.(\hat{\beta}_{j,B})}$ and

$$s.e.(\hat{\beta}_{j,B}) = \left(\frac{1}{B-1} \sum_{s=1}^{B} \left(\hat{\beta}_{j,s}^*(\tau) - \bar{\hat{\beta}}_j^*(\tau) \right)^2 \right)^{1/2}$$

where $\bar{\hat{\beta}}(\tau)_j^* = B^{-1} \sum_{s=1}^{B} \hat{\beta}_{j,s}^*(\tau)$

- Flores, Flores-Lagunes and Kapetanakis (2013) suggest a two-way quantile panel estimator where both the individual and time effects vary across quantiles
- In their application, estimation of a Kuznet's curve, they argue that these individual and time effects should vary across quantiles, thus Canay's estimator will not work
- They note "...it is likely that the factors accounted for by the state and year fixed effects will have a different impact at different quantiles of the conditional distribution of pollution. For instance, the effects of factors like fossil fuel availability and tastes in the case of state fixed effects are likely to be different at different levels of emissions ..."

- Flores, Flores-Lagunes and Kapetanakis (2013) allow both time and individual effects to vary across quantiles by brute force minimization of

$$\min_{c(\tau),d(\tau),\beta(\tau)} \sum_{i=1}^{N} \sum_{t=1}^{T} \rho_\tau \left(y_{it} - c_i(\tau) - d_t(\tau) - x_{it}'\beta(\tau) \right) \tag{3}$$

where $\rho_\tau(z) = z\left(\tau - 1\left\{z < 0\right\}\right)$ is the classic check function of Koenker and Basset (1978)

- There are a few important methodological issues with the approach of Flores, Flores-Lagunes and Kapetanakis (2013)
- First, when one estimates quantiles in the above fashion they may cross, which is theoretically inconsistent
- Second, $N \longrightarrow \infty$ is required for $\hat{d}(\tau)$ to be consistently estimated and $T \longrightarrow \infty$ is required for $\hat{c}(\tau)$ to be consistently estimated
- It is likely that substantial finite sample biases will exist for both $\hat{c}(\tau)$ and $\hat{d}(\tau)$ for moderately sized samples, bias correction methods may be required
- Flores, Flores-Lagunes and Kapetanakis (2013) do not study the asymptotic properties of their estimator but recommend inference be conducted using the bootstrap of Maasoumi and Millimet (2005) and Kato, Galvao and Montes-Rojas (2010)

- Quantile estimation for the fixed effects framework of the unobserved effects model hinges on how one vies the individual and time effects
- When the individual effects are constant across quantiles Canay (2011) proposes a simple two-step estimator
- Minimization estimators exist that allow individual effects to vary across quantiles, can be troublesome when N is large
- As Koenker (2004, pg. 76) acknowledges "At best we may be able to estimate an individual specific location-shift effect, and even this may strain credulity."