Incomplete IV models for binary outcomes

Andrew Chesher

CeMMAP and UCL

September 16th 2008

Abstract

There is a threshold crossing model for a binary outcome Y which specifies $Y = 1[p(X) < U]$ where U is a scalar continuously distributed random variable distributed independently of instrumental variables Z which are excluded from $p(\cdot)$. The model allows X to be endogenous and is incomplete in the sense that it is silent about the genesis of X. Chesher (2007) shows that the model set identifies the threshold crossing function $p(\cdot)$.

An example of this model arises when Tamer's (2003) simultaneous market entry game model is taken one equation at a time, focusing on the determination of one player's strategy, leaving the other's unspecified.

This paper focuses on the case in which X is discrete with K points of support, $\{x_k\}_{k=1}^K$, in which case the threshold crossing function is completely characterised by the K quantities $\Theta \equiv \{\theta_k\}_{k=1}^K$ where $\theta_k \equiv p(x_k)$. The identified set for Θ is studied. It is shown that at each value z of the instrumental variables Z when $P[X = x_k|Z = z] \in (0, 1)$ for all k, the identified set consists of $K!$ boxes each of dimension K embedded in the K dimensional unit hypercube. The identified set when there are many values of the instrumental variables is the intersection of the sets obtained at each value. When instruments are strong there may be substantially less than $K!$ boxes in the resulting set but generally when there is more than one box the set is not connected.

Further restriction on the threshold crossing function $p(\cdot)$ can have substantial effects on the identified sets. For example a monotonicity restriction reduces the identified set to at most 2 boxes in the unit hypercube. Consideration of the case with large K gives some insight into the identification problem that arises when X is continuous.

References

Chesher Andrew (2007): "Endogeneity and Discrete Outcomes," CeMMAP Working Paper 05/07.

Tamer, Elie (2003): "Incomplete Simultaneous Discrete Response Model with Multiple Equilibria," *Review of Economic Studies*, 70(1), 147-165.