Identification and Estimation in Discrete Choice Demand Models when Endogenous Variables Interact with the Error, By A. Gandhi, K. Kim, and A. Petrin.

Winter Meetings

January 2012.

Ariel Pakes

(Harvard University and the NBER.)

I like the basic idea in this paper. As far as I can tell it is;

novel, correct, and extends the literature on estimation of demand systems in what, perhaps with extensions, should be a relevant way.

The caveat about "relevant way" is important. I do not think we should publish every paper that extends the literature on estimating demand systems. We should only publish those papers that demonstrate that the extension is useful for the kind of data available on problems we care about. This seems like an extension that satisfies the constraint, but I would have been happier had they demonstrated it on a problem we care about (and I don't think they did this).

Part of my job as a discussant is to complain. So here are three complaints.

- The conceptual idea, an idea that is simple enough for a practitioner who is not particularly well versed in this literature could understand and use, is not set out in a transparent way.

- What I would view as the extension which is likely to make their procedure most elevant (see below) is not discussed.

- The econometric issues that arise in obtaining limit theorems for the estimated parameters are not detailed. As a result there are issues with implementation that practitioners are not made aware of and are likely to cause problems (which will undermine the usefulness of the estimator).

Basic Idea. Use the contraction mapping in BLP to solve for $\delta_j(\cdot, \theta_1)$, where θ_1 indexes the parameters determining the variance-covariance of the random coefficients. In BLP these $\delta_j(\cdot)$ were assumed to be additively separable in ξ_j, the unobserved product characteristic. Here they generalize and write

$$\delta_j(\cdot, \theta_{1,0}) = x_j\beta - \alpha p_j + \xi_j + \xi_j(x_j\gamma_x - p_j\gamma_p). \quad (1)$$

I.e. ξ_j not only enhances the product, it (potentially) intensifies the preferences for every characteristic of the product, and it does so in a way that is independent of individual's preference intensity for that characteristic.

This independence restricts the form of the demand function, but I am going to ignore resulting problems from that; after all they have to start somewhere.

Problem in Specification. I have a strong preference for allowing the unobservable in the interaction to be different from the one in the constant term. Take their example; the impact of advertising. It is clear that firms hope advertising interacts with price, as changing that interaction changes markups (and there is an empirical literature which verifies this). However there are other unobservables which also effect the demand for the product.

With two unobservables the BLP inversion still works, but the identification question remains and I have not explored that.

Estimation Issue. BLP estimate by assuming there exists a z such that $E[\xi|z] = 0$, and if the model is linear, that $[x, z]$ is of full column rank (so we can use instruments or 2SLS). This runs into a problem here because under this assumption

$$E[\xi \times p|z] \neq 0,$$

because ξ is a determinant of p.

Their solution. In its simplest form all we do is tegress ξ on $[x, z]$ and note that if $v = \xi - f(x, z)$ then $v + f(x, z) = \xi$ and

$$E[f(x, z)|z] = 0, \tag{2}$$

because $E(v|z) = 0$ by consturction and $E(\xi|z) = E(v|z) + E(f(x, z)|z) = 0$.

So provided $f(\cdot) \neq p$ we can subsitute $v + f(\cdot)$ (our "control function") for ξ in

$$\delta_j(\cdot, \theta_{1,0}) = x_j\beta - \alpha p_j + \xi_j + \xi_j(x_j\gamma_x - p_j\gamma_p),$$

estimate parameters there, and impose the constraints in equation (2) to get identification of the parameters of interest.

Econometric details. I would have liked a look at limit theorems, as it is the asymptotic distribution that is likely to be problematic in this context. Moreover, much of the source of the problem is the relationship between N (consumer sample size), R (number of random draws), and J (number of products), and they don't bother to tell us what we need for this.

- Unless I am missing something, given their identification proof, standard semi-parametric results, and the results in B Linton P, the only problem with consistency is insuring that the objective function is bounded away from zero at parameter values different from the truth. We (B Linton P) waived our hands at this. So do they. So why the repetition of (the rather tiresome) discussion in BLiP?

- There is no asymptotic normality proof, and it is going to;
  - Require conditions on the rates at which N and R grow relative to J that will be instructive to practitioners as this is a case where large N and R are crucial (the expansions have terms like s_j in the denominator and they are going to zero).
  - The semi-parametric component does not satisfy an orthogonality constraint so the variance-covariance matrix is likely to be complicated. Even if you end up doing a bootstrap (which will be computationally costly for this problem), you might want to know what the variance depends upon, as you are using a different estimation algorithm, and there is a choice of control functions, and you might want to chose the one most likely to increase precision.

Monte Carlo and Empirical Example. Here are my problems with these

- I think Monte Carlo designs should reflect empirical problems, and this one does not. Also I would use try one design where the assumptions are wrong (either two unobservables as above, or the unobservable interacting with individual specific preference parameters) to see how well the procedure does when they are wrong in relevant ways.

- The comparison to BLP is strange since they never compare to anything BLP presented, and the specification they do use is one which BLP tried and concluded was not able to provide parameter estimates that were precise enough to be useful. Also GM told us our markups were close to the truth, so if you get elasticities that are too different....