Robust Confidence Regions for Incomplete Models, Supplementary Material: Implementation

Larry G. Epstein Hiroaki Kaido Kyoungwon Seo

January 4, 2016

Abstract

We provide details on how to implement the inference method proposed in the main text.

1. Implementation

Construction of our confidence region requires computing the belief function ν_θ and the critical value c_θ. For simple examples, one may compute ν_θ analytically. In general, it can be computed using a simulation procedure. Once ν_θ is obtained, the critical value c_θ can be computed using another simulation procedure, as demonstrated by the Monte Carlo experiments in Section 5 in the main text. Below, we illustrate the simulation procedures using the entry game example studied by Bresnahan and Reiss (1990,1991), Berry (1992), and Ciliberto and Tamer (2009); the latter is CT henceforth.

Suppose there are K firms that are potential entrants into markets $i = 1, 2, \cdots$. For each i, we let $s_i = (s_{i1}, \cdots, s_{iK}) \in \{0, 1\}^K$ denote the vector of entry decisions made by the firms. For firm k in market i, CT consider the following profit function specification:

$$\pi_k(s_i, x_i, u_i; \theta) = \left(v_i'\alpha_k + z_{ik}'\beta_k + w_{ik}'\gamma_k + \sum_{j \neq k} \delta_j^k s_{ij} + \sum_{j \neq k} z_{ij}'^k{}_j s_{ij} + u_{ik} \right) s_{ik},$$

where v_i is a vector of market characteristics, $z_i = (z_{i1}, \cdots, z_{iK})$ is a matrix of firm characteristics that enter the profits of all firms in the market, while

$w_i = (w_{i1}, \cdots, w_{iK})$ is a matrix of firm characteristics such that w_{ik} enters firm k's profit but not other firms' profits. We let x_i collect v_i, z_i, and w_i and stack them as a vector. The unobservable payoff shifters $u_i = (u_{i1}, \cdots, u_{iK})$ follow a multivariate normal distribution $N(0, \Sigma)$ and vary across markets in an i.i.d. way.[1] The structural parameter θ includes Σ and the parameters associated with the profit functions: $\{\beta_k, \gamma_k, \{\delta_j^k, \phi_j^k\}_{j\neq k}\}_{k=1}^K$.

In this example, firm k's profit from not entering the market is 0. Hence, the set of pure-strategy Nash equilibria is given by

$$G(u_i|\theta, x_i) = \{s_i \in S : \pi_k(s_i, x_i, u_i; \theta) \geq 0, \forall k = 1, \cdots, K\}. \tag{1.1}$$

Suppose that a sample $\{(s_i, x_i), i = 1, \cdots, n\}$ of size n is available. Let A be a subset of $S = \{0, 1\}^K$. CT only use singleton events $A = \{s\}, s \in S$ and provide a simulation procedure to calculate $\nu_\theta(A|x)$ and its conjugate (called $\mathbf{H}_1$ and $\mathbf{H}_2$ in their paper). In general, one can use any event $A \subset S$ for inference, and we describe a simulation procedure for this general setting below.

Recall that the belief function of event A conditional on x was given by

$$\nu_\theta(A|x) = m_\theta(\{u \in U : G(u|\theta, x) \subset A\}). \tag{1.2}$$

Hence, a natural way to approximate $\nu_\theta(A|x)$ for any $A \subset S$ is to simulate u from the parametric distribution m_θ and calculate the frequency of the event $G(u|\theta, x) \subset A$. We summarize the procedure below.

Simulation procedure 1

Step 1 Fix the number of draws R. Given Σ, draw random vectors $u^r = (u_1^r, \cdots, u_K^r)$, $r = 1, \cdots, R$, from $N(0, \Sigma)$.

Step 2 For each $(s, x, u^r) \in S \times X \times U$, calculate

$$I(s, x, u^r; \theta) = \begin{cases} 1 & \pi_k(s, x, u^r; \theta) \geq 0, \ \forall k, \\ 0 & \text{otherwise.} \end{cases}$$

That is, $I(s, x, u^r) = 1$ if s is a pure strategy Nash equilibrium under (x, u^r) and θ.

[1] In the context of entry games played by airlines, CT model u_{ik} as a sum of independent normal random variables: firm-specific unobserved heterogeneity, market-specific unobserved heterogeneity, and airport-specific unobserved heterogeneity. This can also be handled by relaxing the i.i.d. assumption on m_θ^∞.

Step 3 Compute the frequency of event $G(u^r|\theta, x) \subseteq A$ across simulation draws by computing that of $A^c \subseteq G^c(u^r|\theta, x)$:

$$\nu_\theta^R(A|x) = \frac{1}{R}\sum_{r=1}^{R}\prod_{s\in A^c}(1 - I(s, x, u^r; \theta)). \tag{1.3}$$

After implementing the simulation procedure above, one can evaluate the test statistic:

$$T_n(\theta) = \max_{(x,j)\in X\times\{1,\ldots,J\}}\left\{\frac{\nu_\theta^R(A_j \mid x) - \Psi_n(s^\infty, x^\infty)(A_j \mid x)}{\sqrt{var_\theta^R(A_j \mid x)/n}}\right\}, \tag{1.4}$$

where $var_\theta^R(A_j \mid x) = \nu_\theta^R(A_j \mid x)(1 - \nu_\theta^R(A_j \mid x))$. The remaining task is to compute the critical value c_θ, which can be done by feeding Λ_θ into a commonly-used simulator for multivariate normal random vectors.

Simulation procedure 2

Step 1 Compute the covariance matrix Λ_θ, which is a $|X|\,J$-by-$|X|\,J$ block-diagonal matrix where $\Lambda_{\theta,x_1}, ..., \Lambda_{\theta,x_{|X|}}$ are the blocks.:

The (j, j')-th entry of each block $\Lambda_{\theta,x}$ is the covariance matrix, conditional on x: $(\Lambda_{\theta,x})_{jj'} = cov_\theta(A_j, A_{j'} \mid x)$, where $cov_\theta(A_j, A_{j'} \mid x)$ is calculated as

$$cov_\theta(A_i, A_j \mid x) = \nu_\theta^R(A_i \cap A_j \mid x) - \nu_\theta^R(A_i \mid x)\,\nu_\theta^R(A_j \mid x). \tag{1.5}$$

Step 2 Decompose Λ_θ as LDL' for a lower triangular matrix L and a diagonal matrix D.

Step 3 Generate $w^r \overset{i.i.d.}{\sim} N(0, I_{|X|J})$ for $r = 1, \cdots, R$. Generate $z^r = LD^{1/2}w^r$, $r = 1, \cdots, R$.

Step 4 Calculate c_θ as the $1-\alpha$ quantile of $\max_{k=1,\cdots,|X|J} z_k/\sigma_{\theta,k}$:

$$c_\theta = \min\left(c \geq 0 : \frac{1}{R}\sum_{r=1}^{R} I(\max_{k=1,\cdots,|X|J} z_k^r/\sigma_{\theta,k} \leq c) \geq 1-\alpha\right).$$

Steps 2-3 in simulation procedure 2 are based on the Geweke-Hajivassiliou-Keane (GHK) simulator. The GHK simulator is widely used in econometrics (see, for example, Hajivassiliou, McFadden, and Ruud (1996) for details). The only difference from the standard GHK-simulator is Step 2, in which we recommend to use the LDL decomposition instead of Cholesky decomposition. This is because Λ_θ may only be positive semidefinite.

Simulation procedure 2 yields a critical value c_θ. Hence, one can determine whether or not a value of the structural parameter should be included in the confidence region by checking if $T_n(\theta) \leq c_\theta$ holds. For constructing a confidence region, one needs to repeat the procedures above for different values of $\theta \in \Theta$. To save computational costs, one can draw $\{(u_1^r, \cdots, u_K^r)\}_{r=1}^R$ and $\{w^r\}_{r=1}^R$ only once and use them repeatedly across all values of θ.

A final remark is that the procedures described above extend to other settings. In other models, the researcher may use a different solution concept (e.g. pairwise stability of networks) that defines the correspondence $G(\cdot|\theta, x)$, or a different parametric specification for the latent variables in the payoff function (e.g. random coefficients following a mixed logit specification). In such cases, one need modify only Steps 1 and 2 in simulation procedure 1.

References

[1] S. Berry, Estimation of a model of entry in the airline industry, *Econometrica* 60 (1992), 889-917.

[2] T.F. Bresnahan and P.C. Reiss, Entry in monopoly markets, *Rev. Econ. Stud.* 57 (1990), 531-553.

[3] T.F. Bresnahan and P.C. Reiss, Empirical models of discrete games, *J. Econometrics* 48 (1991), 57-81.

[4] F. Ciliberto and E. Tamer, Market structure and multiple equilibria in airline markets, *Econometrica* 77 (2009), 1791-1828.

[5] V. Hajivassiliou, D. McFadden, and P. Ruud, Simulation of multivariate normal rectangle probabilities and their derivatives theoretical and computational results, *J. Econometrics* 72 (1996), 85–134.