Supplement to "Information structure and statistical information in discrete response models"

(*Quantitative Economics*, Vol. 9, No. 2, July 2018, 995–1017)

Shakeeb Khan
Department of Economics, Boston College

Denis Nekipelov
Department of Economics, University of Virginia

This supplemental appendix is composed of two distinct parts. The first part contains statements and proofs of some of the theorems and those not included in the Appendix to the main paper. The second part provides a more complete discussion of models and results for simultaneous systems of equations (e.g., games).

Part 1. Triangular discrete response model

Proof for Theorem 3.2, part (ii)

We provide an *upper* bound on information; Recall the expression in the main text for the information in the incomplete information triangular model,

$$I_\alpha = \left\| D_1(x_1, x; \alpha_0, g)\right\|^2_{L_2(\pi_1)} + \left\| D_1(x_1, x; \alpha_0, g) - D_2(x_1, x; \alpha_0, g)\right\|^2_{L_2(\pi_2)}, \tag{SA.1}$$

where D_1 and D_2 are defined as in the main text.

Consider the expression for the information in the incomplete information triangular model expressed in (SA.1):

$$I_\alpha = \left\| D_1(x_1, x; \alpha_0, g)\right\|^2_{L_2(\pi_1)} + \left\| D_1(x_1, x; \alpha_0, g) - D_2(x_1, x; \alpha_0, g)\right\|^2_{L_2(\pi_2)}.$$

We construct the measure π^{**} (it may not be a probability measure) that is constructed as an integral over $\frac{d\pi^{**}}{d\nu} = \max\{\frac{d\pi_1}{d\nu}, \frac{d\pi_1}{d\nu}\}$, where the maximum is considered in the pointwise sense over all regular points of measures π^1 and π^2, and where $\frac{d\pi_1}{d\nu}$ is the Radon–Nykodim density with respect to the σ-finite measure ν. Then we note that $\pi^{**}(\mathbb{R}^2) < \Pi < \infty$, assuming that both measures are defined on the entire $\mathbb{R}^2$. We denote $w(t) = \Phi(t)$ and $t = (x - v)/\sigma$, and express

$$\begin{aligned} D_1(x_1, x; \alpha_0, g) &= \sigma^2 \int w(t) g\big(x_1 + \alpha_0 w(t), x - \sigma t\big)\, dt \\ &\le \sigma^2 \int \max_{w\in[0,1]} g(x_1 + \alpha_0 w, x - \sigma t)\, dt \end{aligned}$$

Shakeeb Khan: shakeeb.khan@bc.edu
Denis Nekipelov: denis@virginia.edu

Available at http://qeconomics.org. https://doi.org/10.3982/QE288

and

$$D_2(x_1, x; \alpha_0, g) - D_1(x_1, x; \alpha_0, g) = \sigma^2 \int \big(1 - w(t)\big) g\big(x_1 + \alpha_0 w(t), x - \sigma t\big)\, dt$$
$$\leq \sigma^2 \int \max_{w \in [0,1]} g(x_1 + \alpha_0 w, x - \sigma t)\, dt.$$

As a result, we find that

$$I_\alpha \leq \big\| D_1(x_1, x; \alpha_0, g) \big\|^2_{L_2(\pi^{**})} + \big\| D_1(x_1, x; \alpha_0, g) - D_2(x_1, x; \alpha_0, g) \big\|^2_{L_2(\pi^{**})}$$
$$\leq 2\sigma^2 \Big\| \max_{w \in [0,1]} g_u(x_1 + \alpha_0 w) \Big\|^2_{L_2(\pi^{**})}.$$

Note that $g_u(\cdot)$ is a probability density. Then, we find that

$$\Big\| \max_{w \in [0,1]} g_u(x_1 + \alpha_0 w) \Big\|^2_{L_2(\pi^{**})} \leq \Big(\sup_x g_u(x) \Big)^2 = \bar{g}_u^2,$$

given that $g(\cdot, \cdot)$ is twice continuously differentiable with finite moments. As a result, we provided an upper bound $I_\alpha \leq 2\sigma^2 \bar{g}_u^2$. As $\sigma \to 0$ this upper bound converges to 0, meaning that $I_\alpha \to 0$.

Part 2. Nontriangular systems: Games of complete and incomplete information

A static game of complete information

Here we consider a simultaneous discrete system of equations where we no longer impose the triangular structure. A leading example of this type of system is a two-player discrete game with complete information (e.g., Bjorn and Vuong (1985) and Tamer (2003)).

We will distinguish the behavioral models from the statistical one, where the latter corresponds to which variables are observed by the econometrician and the former corresponds to which are observed by the agents.

Economic model A simple binary game of complete information is characterized by the players' deterministic payoffs, strategic interaction coefficients, and random payoff components U and V. There are two players $i = 1, 2$ and the action space of each player consists of two points $A_i = \{0, 1\}$ with the actions denoted $Y_i \in A_i$. The payoff to player 1 from choosing action $Y_1 = 1$ can be characterized as a function of observed covariates and player 2's action,

$$Y_1^* = Z_1' \gamma_0 + \alpha_{10} Y_2 - U,$$

where Z_1 denotes a vector of covariates; the payoff of player 2 from choosing action $Y_2 = 1$ is characterized as

$$Y_2^* = Z_2' \delta_0 + \alpha_{20} Y_1 - V,$$

where Z_2 denotes a vector of covariates. The variables $(\gamma_0, \delta_0, \alpha_{10}, \alpha_{20})$ denote coefficients and, analogous to before, the econometrician is primarily interested in the parameters α_{10} and α_{20}, often referred to as the interaction parameters in the empirical industrial organization literature. Because of this, for convenience of both notation and analysis we assume the parameters γ_0 and δ_0 are known, and we change notation to $X_1 = Z_1'\gamma_0$ and $X_2 = X_2'\delta_0$. We normalize the payoff from action $Y_i = 0$ to zero, and we assume that realizations of covariates X_1 and X_2 are commonly observed by the players along with realizations of the variables U and V. Thus, in the game of complete information, each player observes all the variables in both payoff functions.

Under this information structure, the pure strategy of each player is the mapping from the observable variables into actions: $(U, V, X_1, X_2) \mapsto 0, 1$.

A pair of pure strategies constitutes a Nash equilibrium if the strategies reflect the best responses to the rival's equilibrium actions. This is the equilibrium concept we assume players use in our behavioral model.

Statistical model In the statistical model, the econometrician observes a random sample of equilibrium outcomes as well as the covariates. The realizations of the random variables U and V are not observed by the econometrician and characterize the unobserved heterogeneity of the players' payoffs.

The observed equilibrium actions are described by random variables (from the viewpoint of the econometrician) characterized by a pair of binary equations in the statistical model

$$\begin{aligned} Y_1 &= \mathbf{1}\{X_1 + \alpha_{10} Y_2 - U > 0\}, \\ Y_2 &= \mathbf{1}\{X_2 + \alpha_{20} Y_1 - V > 0\}, \end{aligned} \tag{SA.2}$$

where the unobserved (to the econometrician) variables U and V are correlated with each other with an unknown distribution. From a random sample of observations of the vector (Y_1, Y_2, X_1, X_2), the econometrician is interested in conducting statistical inference on the strategic interaction parameters α_{10} and α_{20}.

As noted in Tamer (2003), the system of simultaneous discrete response equations (SA.2) has a fundamental problem of *indeterminacy*, which relates to the lack of a unique Nash equilibrium in the game. This by itself must first be resolved to attain point identification of the interaction parameters. To do so, we impose the following additional assumption, which effectively is an equilibrium selection mechanism when multiple equilibria arise.

Assumption SA.1. *Let x_1, x_2, u, and v denote realizations of the random variables X_1, X_2, U, and V. Denote the sets*

$$\begin{aligned} S_1 &= [\alpha_{10} + x_1, x_1] \times [\alpha_{20} + x_2, x_2], \\ S_2 &= [x_1, \alpha_{10} + x_1] \times [x_2, \alpha_{20} + x_2], \\ S_3 &= [\alpha_{10} + x_1, x_1] \times [x_2, x_2 + \alpha_{20}], \end{aligned}$$

and

$$S_4 = [x_1, x_1 + \alpha_{10}] \times [\alpha_{20} + x_2, x_2].$$

Note that $S_1 = \emptyset$ *if and only if (iff)* $\alpha_{10} > 0$ *and* $\alpha_{20} > 0$, *and* $S_2 = \emptyset$ *iff* $\alpha_{10} < 0$ *and* $\alpha_{20} < 0$.

Then we make the following assumptions:

(i) *If* $S_1 \neq \emptyset$ *or* $S_2 \neq \emptyset$, *then*

$$\Pr\bigl(Y_1 = Y_2 = 1 | (u, v) \in S_k\bigr) \equiv \frac{1}{2} \quad \text{for } k = 1, 2.$$

(ii) *If* $S_3 \neq \emptyset$ *or* $S_4 \neq \emptyset$, *then*

$$\Pr\bigl(Y_1 = (1 - Y_2) = 1 | (u, v) \in S_k\bigr) \equiv \frac{1}{2} \quad \text{for } k = 3, 4.$$

Assumption SA.1 requires that when the system of binary responses has multiple solutions, then the realization of a particular solution is determined by a symmetric coin flip. Furthermore, in regions where the system may have no solutions, we impose solutions via randomization. This assumption addresses both the *incoherency* and *incompleteness* that may arise in these models.[1]

Assumption SA.1 is a strong condition that we deliberately impose to demonstrate how difficult it is to identify the interaction parameters in this model. Specifically, while the assumption eliminates the difficulties that arise from incompleteness and incoherency, we will show that it does not suffice to conduct standard inference on the interaction parameters. We will again demonstrate this by evaluating the Fisher information for the interaction parameters after imposing our equilibrium selection rule.

To do so, we formalize our conditions on the joint distribution of observables X_1 and X_2 and unobservables U and V with the following assumption, which is analogous to Assumption 1 in the main paper in the triangular model.

Assumption SA.2. *Make the following assumptions*:

(i) *Observables* X_1 *and* X_2 *have a continuous distribution with full support on* $\mathbb{R}^2$ *(which is not contained in any proper one-dimensional linear subspace). The parameters of interest,* α_{10} *and* α_{20}, *lie in the interior of a convex compact set* $\mathcal{A}_1 \times \mathcal{A}_2$.

(ii) *The unobservables* (U, V) *are independent of* (X_1, X_2) *and have a continuously differentiable density with full support on* $\mathbb{R}^2$, *with an integrable envelope over* v *and* u *and joint c.d.f.* $G(\cdot, \cdot)$. *The partial derivatives* $\frac{\partial G(u,v)}{\partial u}$ *and* $\frac{\partial G(u,v)}{\partial v}$ *exist and are strictly positive on* $\mathbb{R}^2$.

(iii) *For each* $t_1, t_2 \in \mathbb{R}$, *there exist functions* $q_1(\cdot)$ *and* $q_2(\cdot)$ *with* $E[q_1(X_1, X_2)^2] < \infty$ *and* $E[q_2(X_1, X_2)^2] < \infty$ *that dominate* $\frac{\partial G(x_1+t_1, x_2+t_2)}{\partial u}$ *and* $\frac{\partial G(x_1+t_1, x_2+t_2)}{\partial v}$, *respectively.*

[1] Following the terminology introduced in Tamer (2003), incoherency refers to the nonexistence of an equilibrium and incompleteness refers to multiplicity of equilibria.

Before considering the Fisher information for the interaction parameters, we first establish identification.[2]

THEOREM SA.1. *Suppose that Assumptions SA.1 and SA.2 are satisfied. Then the interaction parameters* α_{10} *and* α_{20} *in model* (SA.2) *are identified.*

Having established the identifiability of the parameters of interest, we now study the information associated with the strategic interaction parameters. The following result establishes that the information associated with the interaction parameters in the static game of complete information is zero. The insight is that in the light of the identification result in Theorem SA.1, this result is not related to the incoherency or incompleteness of the static game.

THEOREM SA.2. *Suppose that Assumptions SA.1 and SA.2 are satisfied. Then the Fisher information associated with parameters* α_{10} *and* α_{20} *in model* (SA.2) *is zero.*

PROOF. To derive the information of the model, we follow the approach in Chamberlain (1986) by demonstrating that for each complete information static game model generated by a distribution satisfying the conditions of Theorem SA.2, we can construct a parametric submodel passing through that model for which the information for parameters α_1 and α_2 is equal to zero.

Suppose that Γ contains all distributions of errors that satisfy the conditions of Theorem SA.2 along with distributions of indices x_1 and x_2. First we construct the likelihood function of the model and introduce the notation

$$
\begin{aligned}
P^{11}(t_1, t) &= \Pr(U \leq t_1, V \leq t) = G(t_1, t), \\
P^{01}(t_1, t) &= \Pr(U > t_1, V \leq t), \\
P^{10}(t_1, t) &= \Pr(U \leq t_1, V > t),
\end{aligned}
$$

and

$$
P^{00}(t_1, t) = \Pr(U > t_1, V > t).
$$

Without loss of generality, we focus on the case where the signs of coefficients α_1 and α_2 coincide. We construct the probability mass that corresponds to the region with multiple equilibria as

$$
\Delta(t_1, t_2; \alpha_1, \alpha_2) = \Pr(t_1 < U \leq t_1 + \alpha_1, t_2 < V \leq t_2 + \alpha_2).
$$

[2]A proof of identification follows immediately from arguments used in Tamer (2003), so we do not include it here.

We write the density of the data as

$$r(y_1, y_2, x_1, x_2; \alpha, P) = \left(P^{11}(x_1+\alpha_1, x_2+\alpha_2) - \frac{1}{2}\Delta(x_1, x_2; \alpha_1, \alpha_2)\right)^{y_1 y_2} \\ \times P^{01}(x_1+\alpha_1, x_2)^{(1-y_1)y_2} P^{10}(x_1, x_2+\alpha_2)^{y_1(1-y_2)} \\ \times \left(P^{00}(x_1, x) - \frac{1}{2}\Delta(x_1, x_2; \alpha_1, \alpha_2)\right)^{(1-y_1)(1-y_2)}$$

with respect to the measure μ defined on $\Omega = \{0,1\}^2 \times \mathbb{R}^2$ such that for any Borel set A in $\mathbb{R}^2$,

$$\mu(\{1,1\} \times A) = \mu(\{1,0\} \times A) = \mu(\{0,1\} \times A) = \mu(\{0,0\} \times A) = \nu(A),$$

where $P((X_1, X_2) \in A) = \int_A d\nu$.

Let $h_1 : \mathbb{R}^2 \mapsto \mathbb{R}$ and $h_2 : \mathbb{R}^2 \mapsto \mathbb{R}$ be continuously differentiable functions supported on the compact set with continuous derivatives in the interior of that compact set such that $\frac{\partial h_i(u,v)}{\partial u} \geq B$ and $\frac{\partial h_i(u,v)}{\partial v} \geq B$ for some constant B on that compact set, and $i = 1, 2$. Define $\tilde{\Lambda}$ as the collection of paths through the original model that we design as

$$\lambda^{11}(t_1, t_2; \delta_1, \delta_2) = P^{11}\big(t_1 + \delta_1(h_1(t_1, t_2) + 1), t_2 + \delta_2(h_2(t_1, t_2) + 1)\big),$$
$$\lambda^{01}(t_1, t_2; \delta_1, \delta_2) = P^{01}\big(t_1 + \delta_1(h_1(t_1, t_2 + \alpha_2) + 1), t_2\big),$$
$$\lambda^{10}(t_1, t_2; \delta_1, \delta_2) = P^{11}\big(t_1, t_2 + \delta_2(h_2(t_1 + \alpha_1, t_2) + 1)\big),$$

and

$$\lambda^{00}(t_1, t_2; \delta_1, \delta_2) = P^{11}(t_1, t),$$
$$\gamma(t_1, t_2; \alpha_1, \alpha_2, \delta_1, \delta_2) = \Pr\big(t_1 < U \leq t_1 + \alpha_1 + \delta_1(h_1(t_1 + \alpha_1, t_2 + \alpha_2) + 1), \\ t_2 < V \leq t_2 + \alpha_2 + \delta_2(h_2(t_1 + \alpha_1, t_2 + \alpha_2) + 1)\big),$$

where we note that these paths maintain the properties of the joint probability distribution (bounded between 0 and 1, summing up to 1) and, in a sufficiently small neighborhood about the origin containing δ, they also maintain the monotonicity of the c.d.f. (as the partial derivatives of $h_1(\cdot,\cdot)$ and $h_2(\cdot,\cdot)$ are bounded from below).

Denote the likelihood function corresponding to the perturbed model $l_\lambda(y_1, y_2, x_1, x_2; \alpha, \delta)$. Provided the assumed dominance condition holds, it will be mean-square differentiable at $(\alpha_0, 0)$. In other words, we can find vector functions $\psi_\alpha(x_1, x_2)$ and $\psi_\delta(x_1, x_2)$ such that

$$l_\lambda^{1/2}(\cdot; \alpha, \delta) = \psi_\alpha(x_1, x_2)'(\alpha - \alpha_0) + \psi_\delta(x_1, x_2)'\delta + R_{\alpha,\delta}$$

with

$$E\big[R^2_{\alpha,\delta}\big] / \big(|\alpha - \alpha_0| + |\delta|\big)^2 \to 0 \quad \text{as } \alpha \to \alpha_0, \delta \to 0.$$

We can explicitly derive the mean-square derivatives. For convenience, we introduce the notation

$$P^{++}(x_1,x_2;\alpha)=P^{11}(x_1+\alpha_1,x_2+\alpha_2)-\frac{1}{2}\Delta(x_1,x_2,\alpha_1,\alpha_2),$$
$$P^{-+}(x_1,x_2;\alpha)=P^{01}(x_1+\alpha_1,x_2),$$
$$P^{+-}(x_1,x_2;\alpha)=P^{10}(x_1,x_2+\alpha_2),$$

and

$$P^{--}(x_1,x_2;\alpha)=P^{00}(x_1,x_2)-\frac{1}{2}\Delta(x_1,x_2,\alpha_1,\alpha_2).$$

In particular, the components of the derivative with respect to the finite-dimensional parameter can be expressed as

$$\begin{aligned}\psi_{\alpha_1}(x_1,x_2)=&\frac{1}{4}\{y_1y_2P^{++}(x_1,x_2;\alpha)^{-1/2}-(1-y_1)(1-y_2)P^{--}(x_1,x_2;\alpha)^{-1/2}\}\\&\times\frac{\partial G(x_1+\alpha_1,x_2+\alpha_2)}{\partial u}-\frac{1}{2}(1-y_1)y_2P^{-+}(x_1,x_2;\alpha)^{-1/2}\frac{\partial G(x_1+\alpha_1,x_2)}{\partial u}\end{aligned}$$

and

$$\begin{aligned}\psi_{\alpha_2}(x_1,x_2)=&-\frac{1}{4}\{y_1y_2P^{++}(x_1,x_2;\alpha)^{-1/2}-(1-y_1)(1-y_2)P^{--}(x_1,x_2;\alpha)^{-1/2}\}\\&\times\frac{\partial G(x_1+\alpha_1,x_2+\alpha_2)}{\partial v}-\frac{1}{2}y_1(1-y_2)P^{+-}(x_1,x_2;\alpha)^{-1/2}\frac{\partial G(x_1,x_2+\alpha_2)}{\partial v}.\end{aligned}$$

The derivative with respect to λ can be expressed as

$$\begin{aligned}\psi_{\delta,1}(x_1,x_2)=&\frac{1}{4}\{y_1y_2P^{++}(x_1,x_2;\alpha)^{-1/2}-(1-y_1)(1-y_2)P^{--}(x_1,x_2;\alpha)^{-1/2}\}\\&\times\frac{\partial G(x_1+\alpha_1,x_2+\alpha_2)}{\partial u}\big(h_1(x_1+\alpha_1,x_2+\alpha_2)+1\big)\\&-\frac{1}{2}(1-y_1)y_2P^{-+}(x_1,x_2;\alpha)^{-1/2}\\&\times\frac{\partial G(x_1+\alpha_1,x_2)}{\partial u}\big(h_1(x_1+\alpha_1,x_2+\alpha_2)+1\big)\end{aligned}$$

and

$$\begin{aligned}\psi_{\delta,2}(x_1,x_2)=&-\frac{1}{4}\{y_1y_2P^{++}(x_1,x_2;\alpha)^{-1/2}-(1-y_1)(1-y_2)P^{--}(x_1,x_2;\alpha)^{-1/2}\}\\&\times\frac{\partial G(x_1+\alpha_1,x_2+\alpha_2)}{\partial v}\big(h_2(x_1+\alpha_1,x_2+\alpha_2)+1\big)\\&-\frac{1}{2}y_1(1-y_2)P^{+-}(x_1,x_2;\alpha)^{-1/2}\\&\times\frac{\partial G(x_1,x_2+\alpha_2)}{\partial v}\big(h_2(x_1+\alpha_1,x_2+\alpha_2)+1\big).\end{aligned}$$

We note that the corresponding score has mean zero.

We use the fact that the Fisher information can be bounded as

$$
\begin{aligned}
I_{\lambda,\alpha_1} &\leq 4\int (\psi_{\alpha_1}-\psi_{\delta,1})^2 d\mu \\
&= \int \frac{1}{4}\Biggl(\bigl[P^{++}(x_1,x_2;\alpha_0)^{-1}+P^{--}(x_1,x_2;\alpha_0)^{-1}\bigr]\left(\frac{\partial G(x_1+\alpha_1,x_2+\alpha_2)}{\partial u}\right)^2 \\
&\quad + P^{-+}(x_1,x_2;\alpha_0)^{-1}\left(\frac{\partial G(x_1+\alpha_1,x_2)}{\partial u}\right)^2\Biggr) h_1^2(x_1+\alpha_1,x_2+\alpha_2)\,d\nu(x_1,x_2).
\end{aligned}
$$

We define the measure on Borel sets in $\mathbb{R}^2$ as

$$
\begin{aligned}
\pi_1(A) = \int_A \frac{1}{4}\Biggl(&\bigl[P^{++}(x_1,x_2;\alpha_0)^{-1}+P^{--}(x_1,x_2;\alpha_0)^{-1}\bigr]\left(\frac{\partial G(x_1+\alpha_1,x_2+\alpha_2)}{\partial u}\right)^2 \\
&+ P^{-+}(x_1,x_2;\alpha_0)^{-1}\left(\frac{\partial G(x_1+\alpha_1,x_2)}{\partial u}\right)^2\Biggr)\,d\nu(x_1-\alpha_1,x_2-\alpha_2),
\end{aligned}
$$

allowing us to characterize $I_{\lambda,\alpha_1} \leq \|h_1\|^2_{L_2(\pi_1)}$. Chamberlain (1986) demonstrates that the space of differentiable functions with compact support is dense in $L_2(\pi)$. Replicating the argument in the proof of zero information for the triangular system with complete information in the main paper, we can demonstrate that $\inf_{\lambda\in\tilde{\Lambda}} I_{\lambda,\alpha_1}=0$. Similarly, we can also show that $\inf_{\lambda\in\tilde{\Lambda}} I_{\lambda,\alpha_2}=0$. □

Our results fully illustrate why the zero Fisher information of the interaction parameters is a problem that is not related to the lack of their point identification or the multiplicity of equilibria. We have proven point identification of these parameters under general conditions and regarding multiplicity, we have explicitly completed the model using randomization of outcomes so that it is complete, yet we still cannot attain positive information. We conclude that the estimation and inference for the interaction parameters are nonstandard even in a simplified model.

Analysis of the game with incomplete information

We now modify the model to allow for incomplete information.

Economic model Our model is based on standard two-player game theory models with incomplete information. Game theoretical results have demonstrated that the introduction of what is referred to in that literature as payoff perturbations leads to a reduction in the number of equilibria.[3]

In the two-player game with incomplete information we again interpret the binary variables Y_1 and Y_2 as the actions of player 1 and player 2. Each player is characterized by the deterministic payoff (corresponding to linear indices X_1 and X_2), an interaction parameter, unobserved heterogeneity terms U and V, and what we refer to here as payoff perturbations, denoted by η_1 and η_2. The payoff of player 1 from action $Y_1=1$ can

[3] See the seminal work of Harsanyi (1995). Multiplicity of equilibria can still be an important issue in games of incomplete information as noted in Sweeting (2009) and de Paula and Tang (2012).

be represented as $Y_1^* = X_1 + \tilde{\alpha}_{10} Y_2 - U - \sigma\eta_1$, while the payoff from action $Y_1 = 0$ is normalized to 0.

In the economic model, player 1 observes X_1, X_2, U, V, a_1 but does not observed η_2; player 2 observes X_1, X_2, U, V, and η_2, but does not observe η_1.

This model is a generalization of the incomplete information model usually considered in empirical applications because we allow for the presence of unobserved heterogeneity components U and V, whose distribution we leave unspecified. We feel this is an important generalization, as most of the empirical results in the industrial organization literature devoted to the analysis of incomplete information games heavily rely on functional form assumptions regarding the distribution of unobserved heterogeneity. Hotz and Miller (1993), Bajari, Hong, and Ryan (2010), and Rust (1987) are just a few of many important examples.

In the economic model the strategy of player i is a mapping from the observable (to the agents) variables into actions: $(X_1, X_2, U, V, \eta_i) \mapsto \{0, 1\}$. Furthermore, player i forms beliefs regarding the action of his/her rival. Provided that η_1 and η_2 are independent, the beliefs are functions only of U, V, and linear indices. Thus, if $P_i(X_1, X_2, U, V)$ are players' beliefs regarding the actions of opponent players, then the strategy, for instance, of player 1 can be characterized as a random variable

$$\begin{aligned} Y_1 &= \mathbf{1}\{E[Y_1^* \mid X_1, X_2, U, V, \eta_1] > 0\} \\ &= \mathbf{1}\{X_1 - U + \tilde{\alpha}_{10} P_2(X_1, X_2, U, V) - \sigma\eta_1 > 0\}. \end{aligned} \tag{SA.3}$$

Similarly, the strategy of player 2 can be written as

$$Y_2 = \mathbf{1}\{X_2 - V + \tilde{\alpha}_{20} P_1(X_1, X_2, U, V) - \sigma\eta_2 > 0\}. \tag{SA.4}$$

We note the resemblance of equations (SA.3) and (SA.4) to the first equation of the triangular system with treatment uncertainty. As in that section, we alter the notation for the interaction parameters, as now they represent coefficients on what are different regressors in the incomplete model.

To characterize the Bayes–Nash equilibrium in the simultaneous move game of incomplete information, we consider a pair of strategies defined by (SA.3) and (SA.4). Moreover, the beliefs of players have to be consistent with their action probabilities conditional on the information set of their rival.

Taking into consideration the independence of player types η_i and the fact that their c.d.f. is known, we can characterize the pair of equilibrium beliefs as a solution to the system of nonlinear equations

$$\begin{aligned} \sigma\Phi^{-1}(P_1) &= x_1 - u + \tilde{\alpha}_{10} P_2, \\ \sigma\Phi^{-1}(P_2) &= x_2 - v + \tilde{\alpha}_{20} P_1. \end{aligned} \tag{SA.5}$$

Our informational assumption regarding the independence of the unobserved heterogeneity components U and V from payoff perturbations η_1 and η_2 enables us to define the game with a coherent equilibrium structure. This would not be the case if we allow correlation between the payoff-relevant unobservable variables of the two players, as their actions should reflect such correlation and the equilibrium beliefs should also be functions of the noise components.

On the other hand, given that the unobserved heterogeneity components U and V are correlated, the individual actions will be correlated. In other words, we consider the structure of the game where the actions of players are correlated without having to analyze a complicated equilibrium structure due to correlated unobserved player types. Multiple equilibria may arise here as well as the system of equations (SA.5) can have multiple solutions.[4] To resolve the uncertainty over equilibria and maintain symmetry with our discussion of games of complete information, we assume that uncertainty over multiple possible equilibrium beliefs is resolved by independent coin flips.

We note that the incomplete information model that we constructed embeds the complete information model in the previous section. When σ approaches 0, the payoffs in the incomplete information model are identical to those in the complete information model and are observable by both players. We illustrate the transition from the complete to the incomplete information environment in Figure SA.1. When $\sigma = 0$, the actions of the players are determined by U and V only. Figure SA.1(a) shows four regions, one for each possible pair of actions in the complete information model. There is a region in the middle where multiple pairs of actions are optimal, leading to multiple equilibria. With the introduction of uncertainty, we can only plot the probabilistic picture of players' actions (integrating over the payoff noise η_1 and η_2). We can then characterize the areas where specific action pairs are chosen with probability exceeding a given quantile $1 - q$. A decrease in the variance of payoff noise leads to the convergence of quantiles to the areas in the illustration of the complete information game in Figure SA.1(a).

Statistical model The econometrician observes a random sample of equilibrium outcomes and covariates. However, the econometrician does not observe realizations of U, V, η_1, and η_2, and knows the distributions of η_1 and η_2, but not of U and V. Under these assumptions as well as regularity conditions detailed below, we wish to determine the identification and information of the interaction parameters $\tilde{\alpha}_{10}$ and $\tilde{\alpha}_{20}$.

Our first result is that we establish the fact that the strategic interaction parameters $\tilde{\alpha}_{10}$ and $\tilde{\alpha}_{20}$ are identified. Note that X_1, X_2, U, and V enter the system of equations (SA.3) and (SA.4) in a way such that the equilibrium beliefs are functions of $X_1 - U$ and $X_2 - V$. Conditional on the realizations X_1, X_2, U, and V, the choices of the two players are also independent. On the other hand, given that the realizations of U and V are not observable to the econometrician, conditional on X_1 and X_2, the choices are correlated. The observed actions are binary and the distribution of the covariates is directly observed in the data (due to independence of the errors (η_1, η_2) and the unobserved heterogeneity (U, V) from the covariates). Thus, the information that the data contain regarding the model is fully summarized by the conditional expectations $E[Y_1|X_1, X_2]$, $E[Y_2|X_1, X_2]$, and $E[Y_1 Y_2|X_1, X_2]$.

[4] Sweeting (2009) considers a 2×2 game of incomplete information and gives examples of multiple equilibria in that game. Bajari, Hong, Krainer, and Nekipelov (2010) develop a class of algorithms for efficient computation of all equilibria in incomplete information games with logistically distributed noise components.

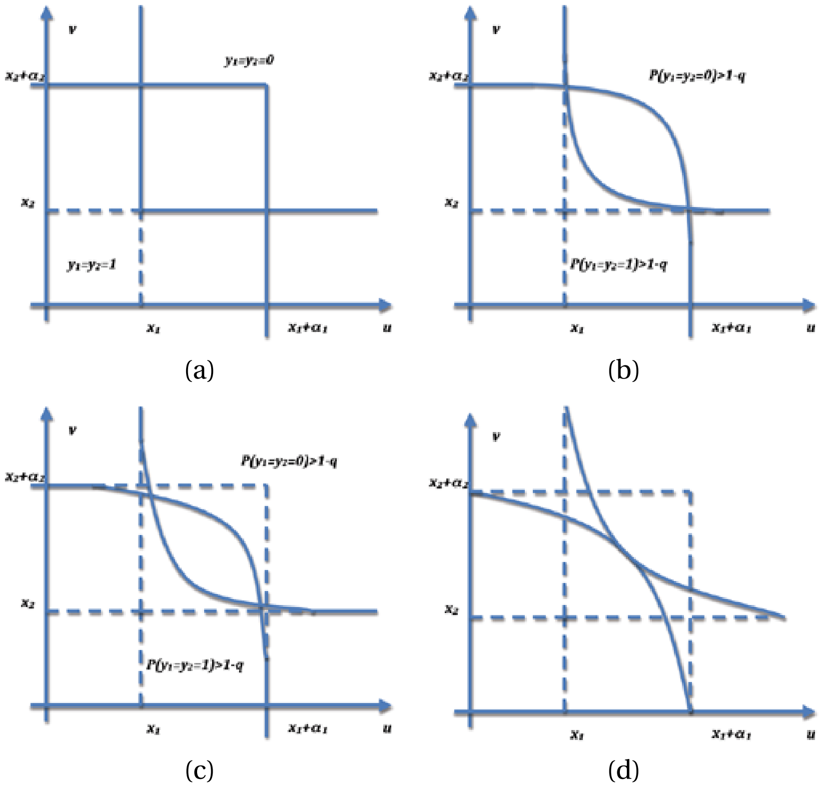

FIGURE SA.1. Incomplete information game.

The identification argument will then have two parts. First, one needs to solve system (SA.5) to obtain mappings $P_1(X_1 - U, X_2 - V)$ and $P_2(X_1 - U, X_2 - V)$. Second, one can relate these mappings to the observable probabilities of actions. Although, with continuous distribution of the noise η_1 and η_2 the considered model has an equilibrium, the system of equilibrium choice probabilities can have multiple solutions. We approach cases of multiple equilibria by resolving the uncertainty via coin flips. We thus make the following three assumptions for the statistical model.

ASSUMPTION SA.3. *Suppose that $\eta \perp (U, V)$ and $\eta \perp (X_1, X)$. The distribution of η has a differentiable density with the full support on $\mathbb{R}$ and a c.d.f. $\Phi(\cdot)$ that is known by the economic agent and the econometrician. In addition, we assume that the density of $\phi(\cdot)$ has "regular" tail behavior, such that there exists $\Delta > 0$ such that for all x for which either $\Phi(x) < \Delta$ or $\Phi(x) > 1 - \Delta$, the density $\phi(\cdot)$ is monotone in x.*

ASSUMPTION SA.4. *Suppose that η_1 and η_2 are privately observed by the two players, meaning player 1 observes η_1 but not η_2, and analogously for player 2. We assume $\eta_1 \perp \eta_2$ and both satisfy Assumption SA.3.*

Our next assumption pertains to the possibility of multiple equilibria that may arise when there are multiple solutions to the following system of nonlinear equations.

Assumption SA.5. *If for some point $(X_1 - U, X_2 - V)$, the system of equations* (SA.5) *has multiple solutions, then the uncertainty regarding the realization of an equilibrium is resolved via a uniform distribution over those solutions.*

In the proof of Theorem SA.3 below we show that the set of solutions is finite. As a result, the observed choice probabilities will correspond to the average value of the mappings P_1 and P_2 over the set of possible values for each given $X_1 - U$ and $X_2 - V$ for given values of covariates.

We then proceed with showing that there exists a set of values of observable covariates (X_1, X_2) of strictly positive measure such that the mapping from observed choice probabilities into the strategic interaction parameters $(\tilde{\alpha}_{10}, \tilde{\alpha}_{20})$ is univalent. In other words, we can identify those parameters from observed data. Once the strategic interaction parameters are identified, the joint distribution of unobserved shocks is identified by the conditional covariance function for observed choices (Y_1, Y_2) given (X_1, X_2).

The following theorem summarizes our identification result.

Theorem SA.3. *Suppose that Assumptions* SA.2, SA.3, SA.4, *and* SA.5 *are satisfied. Then the strategic interaction terms $\tilde{\alpha}_{10}$ and $\tilde{\alpha}_{20}$ in the model defined by* (SA.3) *and* (SA.4) *are identified.*

Proof. Let $P(Y_1, Y_2|X_1, X_2)$ be the observable conditional probability of player's actions conditional on the covariates.

Consider the system of equilibrium cutoff strategies responses (SA.3) and (SA.4) with belief functions determined by (SA.5). Take the sequence of covariates $x^l_{2n} \to -\infty$ (e.g., one can take $x^l_{2n} = -n$) for $n = 1, 2, \ldots$. Denote

$$t^{2l}_n(x_1, u, v; \eta_2) = \mathbf{1}\{x^l_{2n} - v + \tilde{\alpha}_{20} P_1(x_1, x^l_{2n}, u, v) - \sigma\eta_2 > 0\}.$$

Note that $|t^l_n(x_1, u, v; \eta_2)| \le 1$; moreover, for functions

$$\tau^{l\pm}_n(v; \eta_2) = \mathbf{1}\{x^l_{2n} - v \pm |\tilde{\alpha}_{20}| - \sigma\eta_2 > 0\}$$

we have

$$\tau^{l+}_n(v; \eta_2) \le t^{2l}_n(x_1, u, v; \eta_2) \le \tau^{l-}_n(v; \eta_2).$$

We note that for all v and η_2,

$$\lim_{n\to\infty} \tau^{l+}_n(v; \eta_2) = 0$$

and

$$\lim_{n\to\infty} \tau^{l-}_n(v; \eta_2) = 0.$$

Therefore, due to the inequality above, $\lim_{n\to\infty} t^{2l}_n(x_1, u, v; \eta_2) = 0$. In equilibrium

$$P_2(x_1, x^l_{2n}, u, v) = \int_{\eta_2} t^{2l}_n(x_1, u, v; \eta_2) f_\eta(\eta_2)\, d\eta_2.$$

Thus, by dominated convergence theorem

$$\lim_{n\to\infty} P_2\big(x_1, x^l_{2n}, u, v\big) = 0.$$

For the first player, we consider the function

$$t^{1l}_n(x_1, u, v; \eta_1) = \mathbf{1}\big\{x_1 - u + \tilde{\alpha}_{10} P_2\big(x_1, x^l_{2n}, u, v\big) - \sigma\eta_1 > 0\big\}.$$

Note that

$$\lim_{n\to\infty} t^{1l}_n(x_1, u, v; \eta_1) = \mathbf{1}\{x_1 - u - \sigma\eta_1 > 0\}.$$

We note that $|t^{1l}_n(x_1, u, v; \eta_1)| \leq 1$ and, thus, we can apply the dominated convergence theorem to find that

$$\lim_{n\to\infty} P_1\big(x_1, x^l_{2n}, u, v\big) = \Phi\bigg(\frac{x_1 - u}{\sigma}\bigg).$$

Thus we conclude that

$$\lim_{n\to\infty} P\big(Y_1 = 1|x_1, x^l_{2n}\big) = \int \Phi\bigg(\frac{x_1 - u}{\sigma}\bigg) g_u(u)\, du. \tag{SA.6}$$

Next we take the sequence $x^r_{2n} \to +\infty$ for $n = 1, 2, \ldots$. Denote

$$t^{2r}_n(x_1, u, v; \eta_2) = \mathbf{1}\big\{x^r_{2n} - v + \tilde{\alpha}_{20} P_1\big(x_1, x^r_{2n}, u, v\big) - \sigma\eta_2 > 0\big\}.$$

As before, $|t^r_n(x_1, u, v; \eta_2)| \leq 1$ and

$$\tau^{r+}_n(v; \eta_2) \leq t^{2r}_n(x_1, u, v; \eta_2) \leq \tau^{r-}_n(v; \eta_2),$$

where now

$$\tau^{r\pm}_n(v; \eta_2) = \mathbf{1}\big\{x^r_{2n} - v \pm |\tilde{\alpha}_{20}| - \sigma\eta_2 > 0\big\}.$$

Note that

$$\lim_{n\to\infty} \tau^{r\pm}_n(v; \eta_2) = 1.$$

This means that

$$\lim_{n\to\infty} t^{2r}_n(x_1, u, v; \eta_2) = 1.$$

Thus, by the dominated convergence theorem

$$\lim_{n\to\infty} P_2\big(x_1, x^r_{2n}, u, v\big) = 1.$$

Repeating our analysis for the first player, we conclude that

$$\lim_{n\to\infty} P_1\big(x_1, x^r_{2n}, u, v\big) = \Phi\bigg(\frac{x_1 + \tilde{\alpha}_{10} - u}{\sigma}\bigg)$$

and

$$\lim_{n\to\infty} P(Y_1 = 1|x_1, x_{2n}^r) = \int \Phi\left(\frac{x_1 + \tilde{\alpha}_{10} - u}{\sigma}\right) g_u(u)\, du. \tag{SA.7}$$

Thus for each x' and x'' such that

$$\lim_{n\to\infty} P(Y_1 = 1|x', x_{2n}^l) = \lim_{n\to\infty} P(Y_1 = 1|x'', x_{2n}^r),$$

the interaction parameter is identified as $\tilde{\alpha}_{10} = x'' - x'$. The argument for identification of $\tilde{\alpha}_{20}$ can be expressed analogously. □

We note that the proof of identification here relies on extreme values of X_1 and X_2, as was used for the identification result for the complete information game. However, for the incomplete information game it is not necessary to rely on limiting values to attain point identification of the interaction parameters. Consequently, as we now show, we can attain positive Fisher information for the interaction parameters in the incomplete information game.

Specifically, we find that for any finite variance of noise σ^2 (which can be arbitrarily small), the information in the model of the incomplete information game is strictly positive. We also provide a result that characterizes the Fisher information for the strategic interaction parameters as the variance of players' privately observed payoff shocks approaches 0. As in the incomplete information triangular model, the Fisher information of those parameters approaches 0.

Theorem SA.4. *Suppose that Assumptions SA.2, SA.3, SA.4, and SA.5 are satisfied.*

(i) *For any $\sigma > 0$, the information corresponding to parameters $(\tilde{\alpha}_{10}, \tilde{\alpha}_{20})$ in the incomplete information game defined by (SA.3) and (SA.4) is strictly positive.*

(ii) *As $\sigma \to 0$, the information corresponding to parameters $(\tilde{\alpha}_{10}, \tilde{\alpha}_{20})$ in the incomplete information game defined by (SA.3) and (SA.4) approaches 0.*

Proof of Theorem SA.4. *Proof of result (i).* We start the proof with the following lemma that demonstrates that the addition of our equilibrium selection mechanism does not affect the smoothness properties of the semiparametric likelihood function.

Lemma SA.1. *The set of values of strategic interaction parameters α_1 and α_2 in the static game of incomplete information for which the game has multiple equilibria is a closed connected set with a differentiable boundary $S^m(\alpha_1, \alpha_2)$.*

Proof. Given the continuous differentiability of the distribution of random perturbations, we can characterize the boundary of the set of multiple equilibria as the set of points on $\mathbb{R}^2$ where the curves corresponding to the best responses of the players to their beliefs regarding their opponents touch for the first time. This corresponds to the set of points on $\mathbb{R}^2$ where

$$\sigma\Phi^{-1}(P_i) = q_i + \alpha_i P_j, \qquad \alpha_i \phi\left(\frac{1}{\sigma}(q_i + \alpha_i P_j)\right) = \left(\alpha_j \phi\left(\Phi^{-1}(P_j)\right)\right)^{-1}, \quad i, j = 1, 2, i \neq j.$$

For given parameters α_1 and α_2, this defines a mapping from the set of covariates q_1, q_2 to the beliefs. This mapping reduces the dimensionality of the overall mapping by 2, as it incorporates the original system of equations for the beliefs and the restriction on the derivatives of the belief functions. It will be a one-dimensional closed curve $e(q_1, q_2) = 0$. This curve will be differentiable in the strategic interaction parameters due to continuous differentiability of the density of the payoff noise. This curve represents the boundary of the set of multiple equilibria, which we denote $S^m(\alpha_1, \alpha_2)$. □

We now use the constructed set of parameters leading to multiple equilibria to form the likelihood function of the models. The likelihood of the model can then be characterized by four objects:

$$
\begin{aligned}
E[Y_1 Y_2 | x_1, x_2] &= P_{11}(x_1, x_2; \alpha) \\
&= \int \Phi\left(\frac{x_1 - u + \alpha_1 P_2(x_1 - u, x_2 - v)}{\sigma}\right) \\
&\quad \times \Phi\left(\frac{x_2 - v + \alpha_2 P_2(x_1 - u, x_2 - v)}{\sigma}\right) g(u, v)\, du\, dv, \\
E[Y_1 | x_1, x_2] &= Q_1(x_1, x_2; \alpha) \\
&= \int \Phi\left(\frac{x_1 - u + \alpha_1 P_2(x_1 - u, x_2 - v)}{\sigma}\right) g(u, v)\, du\, dv, \\
E[Y_2 | x_1, x_2] &= P_1(x_1, x_2; \alpha) \\
&= \int \Phi\left(\frac{x_2 - v + \alpha_2 P_1(x_1 - u, x_2 - v)}{\sigma}\right) g(u, v)\, du\, dv,
\end{aligned}
$$

$$
\begin{aligned}
&\Pr\big((X_1 - U, X_2 - V) \in S^m(\alpha_1, \alpha_2) | x_1, x_2\big) \\
&= \Delta(x_1, x_2; \alpha) \\
&= \int \mathbf{1}\big\{(x_1 - u, x_2 - v) \in S^m(\alpha_1, \alpha_2)\big\} g(u, v)\, du\, dv.
\end{aligned}
$$

We assume that $\alpha_1, \alpha_2 > 0$ without loss of generality. We construct the probabilities corresponding to observed equilibrium outcomes as

$$
\begin{aligned}
P^{++}(x_1, x_2; \alpha) &= P_{11}(x_1, x_2; \alpha) - \frac{1}{2}\Delta(x_1, x_2; \alpha), \\
P^{-+}(x_1, x_2; \alpha) &= P_1(x_1, x_2; \alpha) - P_{11}(x_1, x_2; \alpha) + \frac{1}{2}\Delta(x_1, x_2; \alpha), \\
P^{+-}(x_1, x_2; \alpha) &= Q_1(x_1, x_2; \alpha) - P_{11}(x_1, x_2; \alpha) + \frac{1}{2}\Delta(x_1, x_2; \alpha),
\end{aligned}
$$

and

$$
P^{--}(x_1, x_2; \alpha) = 1 - P_1(x_1, x_2; \alpha) - Q_1(x_1, x_2; \alpha) + P_{11}(x_1, x_2; \alpha) - \frac{1}{2}\Delta(x_1, x_2; \alpha).
$$

Denote the gradients

$$D_1(x_1,x_2;\alpha)=\frac{\partial}{\partial\alpha'}\left(P_{11}(x_1,x_2;\alpha)-\frac{1}{2}\Delta(x_1,x_2;\alpha)\right),$$
$$D_2(x_1,x_2;\alpha)=\frac{\partial}{\partial\alpha'}P^{-+}(x_1,x_2;\alpha),$$

and

$$D_3(x_1,x_2;\alpha)=\frac{\partial}{\partial\alpha'}P^{+-}(x_1,x_2;\alpha).$$

We focus on the square root of the density corresponding to the likelihood of the model:

$$\begin{aligned}r(y_1,y_2|x_1,x_2;\alpha)^{1/2}&=y_1y_2P^{++}(x_1,x_2;\alpha)^{1/2}+(1-y_1)y_2P^{-+}(x_1,x_2;\alpha)^{1/2}\\&\quad+y_1(1-y_2)P^{+-}(x_1,x_2;\alpha)^{1/2}+(1-y_1)(1-y_2)P^{--}(x_1,x_2;\alpha)^{1/2}.\end{aligned}$$

Then we can express the mean-square gradient of this density as

$$\begin{aligned}\psi_\alpha(x_1,x_2)&=\frac{1}{2}\left\{y_1y_2P^{++}(x_1,x_2;\alpha)^{-1/2}-(1-y_1)(1-y_2)P^{--}(x_1,x_2;\alpha)^{-1/2}\right\}D_1(x_1,x_2;\alpha)\\&\quad+\frac{1}{2}\left\{(1-y_1)y_2P^{-+}(x_1,x_2;\alpha)^{-1/2}-(1-y_1)(1-y_2)P^{--}(x_1,x_2;\alpha)^{-1/2}\right\}D_2(x_1,x_2;\alpha)\\&\quad+\frac{1}{2}\left\{y_1(1-y_2)P^{+-}(x_1,x_2;\alpha)^{-1/2}-(1-y_1)(1-y_2)P^{--}(x_1,x_2;\alpha)^{-1/2}\right\}D_3(x_1,x_2;\alpha).\end{aligned}$$

We note that the corresponding score has mean zero and that, conditional on the covariates, the terms in this expression are positively correlated. Then by definition, $I_\alpha=4\int\psi_\alpha(x_1,x_2)\psi_\alpha(x_1,x_2)'\,d\mu$. Thus, if ν is the measure on $\mathbb{R}^2$ corresponding to the distribution of x_1 and x, following the approach in the derivation of information of the complete information model, we define the measures on Borel subsets of $\mathbb{R}^2$:

$$\pi_1(A)=\int_A\frac{1-P^{-+}(x_1,x_2;\alpha_0)-P^{+-}(x_1,x_2;\alpha_0)}{P^{++}(x_1,x;\alpha_0)P^{--}(x_1,x;\alpha_0)}\,d\nu(x_1,x),$$
$$\pi_2(A)=\int_A\frac{1-Q_1(x_1,x_2;\alpha_0)}{P^{-+}(x_1,x;\alpha_0)P^{--}(x_1,x;\alpha_0)}\,d\nu(x_1,x),$$

and

$$\pi_3(A)=\int_A\frac{1-P_1(x_1,x_2;\alpha_0)}{P^{-+}(x_1,x;\alpha_0)P^{--}(x_1,x;\alpha_0)}\,d\nu(x_1,x).$$

Due to discovered positive correlation between the components of the mean-square gradient, we can evaluate the information as

$$I_\alpha\geq\left\|D_1(x_1,x_2;\alpha_0)\right\|^2_{L_2(\pi_1)}+\left\|D_2(x_1,x_2;\alpha_0)\right\|^2_{L_2(\pi_2)}+\left\|D_3(x_1,x_2;\alpha_0)\right\|^2_{L_2(\pi_3)}.$$

Then we can construct the measure π^* that minorizes the Radon–Nikodym density of measures π_1 and π_2, meaning that $\frac{d\pi^*}{d\nu}=\min\{\frac{d\pi_1}{d\nu},\frac{d\pi_2}{d\nu}\}$. Based on this structure of the measure, we can write

$$I_\alpha \geq \left\|D_1(x_1, x_2; \alpha_0)\right\|^2_{L_2(\pi^*)} + \left\|D_2(x_1, x_2; \alpha_0)\right\|^2_{L_2(\pi^*)} + \left\|D_3(x_1, x_2; \alpha_0)\right\|^2_{L_2(\pi_3)}.$$

By combining the triangle inequality and taking into account the nonnegativity of the square, we can evaluate

$$I_\alpha \geq \left\|D_1(x_1, x_2; \alpha_0) + D_2(x_1, x_2; \alpha_0)\right\|^2_{L_2(\pi^*)}.$$

Then we note that

$$D_1(x_1, x_2; \alpha_0) + D_2(x_1, x_2; \alpha_0) = \int \phi\left(\frac{1}{\sigma}\Phi^{-1}(P_1)\right)\left(\alpha_1 \frac{\partial P_2}{\partial \alpha} + (P_2, 0)'\right) g(u, v)\, du\, dv.$$

We denote $t_1 = (x_1 - u)/\sigma$ and $t_2 = (x - v)/\sigma$. Then

$$\begin{aligned} &D_1(x_1, x_2; \alpha_0) + D_2(x_1, x_2; \alpha_0) \\ &\quad = \sigma^2 \int \phi\left(\frac{1}{\sigma}\Phi^{-1}\big(P_1(\sigma t_1, \sigma t_2)\big)\right) \\ &\qquad \times \left(\alpha_1 \frac{\partial P_2(\sigma t_1, \sigma t_2)}{\partial \alpha} + \big(P_2(\sigma t_1, \sigma t_2), 0\big)'\right) g(\sigma t_1 + x_1, \sigma t_2 + x_2)\, dt_1\, dt_2. \end{aligned}$$

Denote

$$w(t_1, t_2) = \phi\left(\frac{1}{\sigma}\Phi^{-1}\big(P_1(\sigma t_1, \sigma t_2)\big)\right)\left(\alpha_1 \frac{\partial P_2(\sigma t_1, \sigma t_2)}{\partial \alpha} + \big(P_2(\sigma t_1, \sigma t_2), 0\big)'\right).$$

Then we can express

$$D_1(x_1, x_2; \alpha_0) + D_2(x_1, x_2; \alpha_0) = \sigma^2 \int w(t_1, t_2) g(x_1 + \sigma t_1, x_2 + \sigma t_2)\, dt_1\, dt_2.$$

Suppose that $S \subset \mathbb{R}^2$ is a compact set such that $\pi^*(S) > C$. Then given that $g(\cdot, \cdot)$ is continuous and strictly positive, there exists

$$M(t_1, t_2) = \inf_{(x_1, x_2) \in S} \left|g(x_1 + \sigma t_1, x_2 + \sigma t_2)\right|$$

that is not equal to zero for at least some $(t_1, t_2) \in \mathbb{R}^2$. We take

$$\sqrt[4]{\varepsilon} = \sup_{t \in [-B, B] \times [-B, B]} \left|M(t)\right|,$$

where B is selected such that $[-B, B] \times [-B, B]$ contains at least one point where $M(t) \neq 0$. Suppose that the supremum is attained at point (t_1^*, t_2^*). By continuity, there exists some neighborhood of (t_1^*, t_2^*) where $M(t) > \sqrt{\varepsilon/2}$. Denote the size of this neighborhood by R. By construction, $w(t_1, t_2)$ is a continuous function that is not equal to zero (given that we assumed that $\alpha_1, \alpha_2 > 0$, we have $\alpha_1 \frac{\partial P_2}{\partial \alpha_1} > 0$). Moreover, this function has a well defined limit as $\sigma \to 0$. Thus, this function attains its lower bound in every compact set and that lower bound is above zero:

$$\inf_{(t_1,t_2)\in B_R(t_1^*,t_2^*)} \|w(t_1,t_2)\| = A\sqrt[4]{\varepsilon} > 0.$$

We substitute our evaluations into the bound for the information,

$$I_\alpha \geq \|D_1(x_1,x_2;\alpha_0)+D_2(x_1,x_2;\alpha_0)\|^2_{L_2(\pi^*)} \geq \|\big(D_1(x_1,x_2;\alpha_0)+D_2(x_1,x_2;\alpha_0)\big)\mathbf{1}_S\|^2_{L_2(\pi^*)}$$
$$\geq CA^2\sigma^2\sqrt{\varepsilon}\left\|\int_{\mathbb{R}} M(t_1,t_2)\,dt\right\|^2 I_{2\times 2} \geq \frac{1}{2}CA^2R^2\sigma^2\varepsilon I_{2\times 2} > 0,$$

where $I_{2\times 2}$ is the identity matrix. Therefore, the information corresponding to parameters α_1 and α_2 is strictly positive.

Proof of result (ii). Suppose that measure π^{**} is such that its Radon–Nikodym density is constructed as

$$\frac{d\pi^{**}}{d\nu} = \max\left\{\frac{d\pi_1}{d\nu}, \frac{d\pi_2}{d\nu}\right\}.$$

Then we can see that

$$I_\alpha \leq \|D_1(x_1,x_2;\alpha_0)\|^2_{L_2(\pi^{**})} + \|D_2(x_1,x_2;\alpha_0)\|^2_{L_2(\pi^{**})} + \|D_3(x_1,x_2;\alpha_0)\|^2_{L_2(\pi^{**})}$$
$$+ 2\|D_1D_2\|^2_{L_2(\pi^{**})} + 2\|D_1D_3\|^2_{L_2(\pi^{**})} + 2\|D_2D_3\|^2_{L_2(\pi^{**})}.$$

Consider the change of variables

$$t_1 = \Phi^{-1}\big(P_1(x_1-u,x_2-v)\big)$$

and

$$t_2 = \Phi^{-1}\big(P_2(x_1-u,x_2-v)\big).$$

Thus, we can write (denoting $a_i = \phi(\Phi^{-1}(P_i))$ for $i = 1,2$)

$$\big|D_1(x_1,x_2;\alpha)\big| \leq \int \begin{pmatrix} \dfrac{1+\alpha_2 a_2}{a_2} \\ \dfrac{1+\alpha_1 a_1}{a_1} \end{pmatrix} \big|a_1a_2\alpha_1^2\alpha_2^2 - 1\big|$$
$$\times P_1P_2 g(x_1+\alpha_1P_1-\sigma t_1, x_2+\alpha_2P_2-\sigma t_2)\,dt_1\,dt_2$$
$$\leq \sigma^2\begin{pmatrix}\alpha_2\\ \alpha_1\end{pmatrix}\int \phi\big(\Phi^{-1}(P_1)\big)\phi\big(\Phi^{-1}(P_2)\big)$$
$$\times P_1P_2 g(x_1+\alpha_1P_1-\sigma t_1, x_2+\alpha_2P_2-\sigma t_2)\,dt_1\,dt_2 + o\big(\sigma^2\big)$$
$$\leq \sigma^2\bar{\phi}^2\begin{pmatrix}\alpha_2\\ \alpha_1\end{pmatrix} + o\big(\sigma^2\big),$$

provided that $\phi(\cdot) \leq \bar{\phi}$ and $P_1, P_2 \leq 1$. The same evaluation can be written for other components $D_i(x_1, x_2; \alpha)$ with $i = 1, 2, 3$. We evaluate the information as

$$I_\alpha \leq \sigma^2 \bar{\phi}^2 A + o(\sigma^2)$$

for a fixed matrix A (determined by coefficients α_1 and α_2). When $\sigma \to 0$, this upper bound approaches 0. Thus, the resulting information converges to zero. □

References

Bajari, P., H. Hong, J. Krainer, and D. Nekipelov (2010), "Computing equilibria in static games of incomplete information using the all-solution homotopy." Working paper. [10]

Bajari, P., H. Hong, and S. Ryan (2010), "Identification and estimation of discrete games of complete information." *Econometrica*, 78, 1529–1568. [9]

Bjorn, P. and Q. Vuong (1985), "Simultaneous equations models for dummy endogenous variables: A game theoretic formulation with an application to labor force participation." Caltech Working Paper 537. [2]

Chamberlain, G. (1986), "Asymptotic efficiency in semi-parametric models with censoring." *Journal of Econometrics*, 32 (2), 189–218. [5, 8]

de Paula, A. and X. Tang (2012), "Inference of signs of interaction effects in simultaneous games with incomplete information." *Econometrica*, 80, 143–172. [8]

Harsanyi, J. C. (1995), "A new theory of equilibrium selection for games with complete information." *Games and Economic Behavior*, 8 (1), 91–122. [8]

Hotz, J. and R. Miller (1993), "Conditional choice probabilities and the estimation of dynamic models." *Review of Economic Studies*, 60, 497–529. [9]

Rust, J. (1987), "Optimal replacement of GMC bus engines: An empirical model of Harold Zurcher." *Econometrica*, 55, 999–1033. [9]

Sweeting, A. (2009), "The strategic timing of radio commercials: An empirical analysis using multiple equilibria." *RAND Journal of Economics*, 40, 710–742. [8, 10]

Tamer, E. (2003), "Incomplete bivariate discrete response model with multiple equilibria." *Review of Economic Studies*, 70, 147–165. [2, 3, 4, 5]

Co-editors Orazio Attanasio and Rosa L. Matzkin handled this manuscript.

Manuscript received 6 June, 2012; final version accepted 3 October, 2017; available online 5 December, 2017.