ONLINE SUPPLEMENT TO IDENTIFICATION OF COMPLETE INFORMATION GAMES

BRENDAN KLINE
UNIVERSITY OF TEXAS AT AUSTIN

1. Basic model and assumptions

(1) $$u_{im}(0, y_{(-i)m}) = 0 \text{ and } u_{im}(1, y_{(-i)m}) = x_{im1} + \tilde{u}_i(x_{im(-1)}, \theta_i) + \sum_{j \neq i} g^w_{ijm} y_{jm} + \epsilon_{im}.$$

(2) $$g^w_{ijm} = s_i(g_m) \left(f_{ij}(z_{ijm}, \delta_{ij}) + \nu_{ijm} \right) g_{ijm}.$$

(3) $$g^w_{ijm} \equiv \Delta_{ij}.$$

Assumption 1.1 (Median independence of the unobservables)**.** *The following two median independence conditions hold:*

1. *For each role i, $\epsilon_i|x$ has zero median for all x in the support. Moreover, it holds that the cumulative distribution function $F_{\epsilon_i|x}(\cdot)$ is strictly increasing in a neighborhood of zero.*
2. *For all roles i and j, $(\epsilon_i + s_i(g)\nu_{ij})|(x, z, g_{ij} = 1, s_i(g))$ has zero median for all $(x, z, g_{ij} = 1, s_i(g))$ in the support. Moreover, it holds that the cumulative distribution function $F_{\epsilon_i + s_i(g)\nu_{ij}|x,z,g_{ij}=1,s_i(g)}(\cdot)$ is strictly increasing in a neighborhood of zero.*

Assumption 1.2 (Continuous unobservables)**.** *For each role i, the distribution of $\epsilon_i|x$ is continuous for all x in the support. Also, for all roles i and j, the distribution of $(\epsilon_i + s_i(g)\nu_{ij})|(x, z, g_{ij} = 1, s_i(g))$ is continuous for all $(x, z, g_{ij} = 1, s_i(g))$ in the support.*

Assumption 1.3 (Sufficient variation of the explanatory variables)**.** *For each role i, if $t_i \neq \theta_i$, then $P(\tilde{u}_i(\cdot, t_i) = \tilde{u}_i(\cdot, \theta_i)) < 1$. For all roles i and j, if $d_{ij} \neq \delta_{ij}$, then $P(f_{ij}(\cdot, d_{ij}) = f_{ij}(\cdot, \delta_{ij})) < 1$.*

Assumption 1.4 (Large support regressor)**.** *For any $x_{\cdot(-1)}$ in the support, $x_{\cdot 1}|x_{\cdot(-1)}$ has everywhere positive density with support on $\mathbb{R}^N$. Also, for any $(x_{\cdot(-1)}, z)$ in the support, $x_{\cdot 1}|(x_{\cdot(-1)}, z)$ has everywhere positive density with support on $\mathbb{R}^N$.*

Assumption 1.5 (Observed at random connections)**.** *It holds that $(\epsilon, \nu)|(x, z, g, \{s_i(g)\}_i) \sim (\epsilon, \nu)|(x, z, \hat{g}, g, \{s_i(g)\}_i)$ and, for all roles i and j, $(\epsilon, \nu)|(x, z, g_{ij} = 1, s_i(g)) \sim (\epsilon, \nu)|(x, z, \hat{g}_{ij} = 1, g_{ij} = 1, s_i(g))$.*

Assumption 1.6 (Existence of observed connections)**.** *For all roles i and j, there is a non-random constant $\overline{s}_{ij} > 0$ such that for any (x, z) in the support, $P(\hat{g}_{ij} = 1, s_i(g) = \overline{s}_{ij}|x, z) > 0$.*

Assumption 1.7 (Uniformly bounded interaction structure)**.** *For any $\delta < 1$, there is $G_\delta > 0$ such that, for any $(x, z, g, \{s_i(g)\}_i)$ in the support, $P(|g^w| \not\leq G_\delta|x, z, g, \{s_i(g)\}_i) \leq 1 - \delta$.*

Assumption 1.8 (Conditional tail behavior of the unobservables)**.** *For all roles i and j, for any non-random function $\tau_i(\cdot)$ of x of the form that $\tau_i(x) \equiv ax_{i1} + \tilde{\tau}_i(x_{i(-1)})$, where $a = 1$ or $a = -1$, and any sequence of x that varies only in $x_{\cdot 1}$ such that $\tau_i(x) \to \infty$, it holds that $F_{\epsilon_i|x}(\tau_i(x)) \to 1$ and $F_{\epsilon_i|x,z,g_{ij}=1,s_i(g)}(\tau_i(x)) \to 1$ and $F_{\epsilon_i|x,z,g,\{s_i(g)\}_i}(\tau_i(x)) \to 1$; and, for any sequence of x that varies only in $x_{\cdot 1}$ such that $\tau_i(x) \to -\infty$, it holds that $F_{\epsilon_i|x}(\tau_i(x)) \to 0$ and $F_{\epsilon_i|x,z,g_{ij}=1,s_i(g)}(\tau_i(x)) \to 0$ and $F_{\epsilon_i|x,z,g,\{s_i(g)\}_i}(\tau_i(x)) \to 0$.*

Assumption 1.9 (Non-flat unobservables)**.** *The following two conditions hold:*

(1) For each role i, and each $(x_{(-i)(-1)}, x_i)$ in the support, and $t \neq 0$, it holds that $\limsup_{x_{(-i)1}} F_{\epsilon_i|x}(t) \neq \frac{1}{2}$ and $\liminf_{x_{(-i)1}} F_{\epsilon_i|x}(t) \neq \frac{1}{2}$, where the limits are along any sequence of $x_{(-i)1}$ such that, for each $j \neq i$, either $x_{j1} \to \infty$ or $x_{j1} \to -\infty$.

(2) For all roles i and j, and each $(x_{(-i)(-1)}, x_i)$ and z and $s_i(g)$ in the support, and $t \neq 0$, it holds that $\limsup_{x_{(-i)1}} F_{\epsilon_i+s_i(g)\nu_{ij}|x,z,g_{ij}=1,s_i(g)}(t) \neq \frac{1}{2}$ and $\liminf_{x_{(-i)1}} F_{\epsilon_i+s_i(g)\nu_{ij}|x,z,g_{ij}=1,s_i(g)}(t) \neq \frac{1}{2}$, where the limits are along any sequence of $x_{(-i)1}$ such that, for each $k \neq i$, either $x_{k1} \to \infty$ or $x_{k1} \to -\infty$.

2. Extension: Identification of the distribution of the unobservables

It is also possible to point identify the distribution of the unobservables. This is useful because the distribution of the unobservables is an ingredient in marginal effects and other counterfactuals.

Assumption 2.1 (Observed network). $\hat{g} \equiv g$

Assumption 2.2 (Interaction effect sign homogeneity). *For all roles i and j, it holds that* $\text{sgn}\,(f_{ij}(z_{ij}, \delta_{ij})) = \text{sgn}\,(f_{ji}(z_{ji}, \delta_{ji}))$ *for any (z_{ij}, z_{ji}) in the support. Also, using the notation that $\Delta_{ij}(z_{ij}) \equiv \text{sgn}\,(f_{ij}(z_{ij}, \delta_{ij}))$, the support of $\nu|(x, z, g)$ is sufficiently small so that $P(g^w_{ij} \geq 0$ if $\Delta_{ij}(z_{ij}) \geq 0$ and $g^w_{ij} \leq 0$ if $\Delta_{ij}(z_{ij}) \leq 0 (\forall i, j)|x, z, g) = 1$ for all (x, z, g) in the support.*

Assumption 2.2 requires, first, that the sign of the "observable" part of the interaction effect is symmetric within any pair of agents. However, this sign can vary as a function of z, and can vary across pairs of agents. Also, second, this assumption requires that the sign of the "overall" interaction effect g^w_{ij} is the same as the sign of the "observable" part.

Assumption 2.3 (Space of unobservables). *$(\epsilon, \nu)|(x, z, g)$ is conditionally independent from $x_{\cdot 1}$. The distribution of $\epsilon|(x, z, g)$ is continuous, and the distribution of $(\epsilon, \nu)|(x, z, g)$ varies continuously with (x, z). Also, $(\epsilon \perp \nu)|(x, z, g)$ and the non-redundant components of $\nu|(x, z, g)$ are mutually independent.*[1] *Finally, unless $\Delta_{ij}(z_{ij}) \leq 0$ for all i and j, and all z in the support, the distribution of $\epsilon|(x, z, g)$ is uniquely characterized by its bivariate marginals.*

The first part of assumption 2.3 (or equivalent) is necessary for identification, as with single-agent discrete choice. The second part rules out pathologies with discontinuous distributions. The third part is necessary for identification because just a few linear combinations of the unobservables determine the distribution of the data, so the entire correlation structure of the unobservables cannot be point identified.[2] If $\nu \equiv 0$, as in the standard model in equation 3, this part is satisfied. The last part is satisfied by many standard distributions, including the normal distribution

[1] ν_{ij} and ν_{ji} are redundant if $\nu_{ij} \equiv \nu_{ji}$ by assumption. The "non-redundant components" includes just one of these. Essentially, this part of the assumption is used to imply that the marginals of ν uniquely characterize the joint distribution. This can happen either because the components of ν are independent, or are equal to each other.

[2] The following elaborates on this claim. Suppose for simplicity that $N = 2$, $g_{ij} \equiv 1$, $f_{ij}(\cdot, \delta_{ij}) = \Delta_{ij}$ and $s_i(g) \equiv 1$; drop conditioning on x and z from the notation, suppose that $\Delta_{12} = \Delta_{21} \leq 0$, and maintain assumption 2.2. In order to focus on the distribution of the unobservables, suppose that the selection mechanism is known to choose $(0, 1)$ in the region of multiple equilibria. (If the unobservables cannot be identified with these assumptions, they cannot be identified without them.) The only distributions of the unobservables that are relevant to determining the distribution of the utility functions are (ϵ_1, ϵ_2), $(\epsilon_1, \epsilon_2 + \nu_{21})$, $(\epsilon_1 + \nu_{12}, \epsilon_2)$, and $(\epsilon_1 + \nu_{12}, \epsilon_2 + \nu_{21})$ corresponding to the outcome $(0, 0)$, $(1, 0)$, $(0, 1)$, and $(1, 1)$ respectively. Moreover, since probabilities sum to 1, actually (ϵ_1, ϵ_2), $(\epsilon_1 + \nu_{12}, \epsilon_2)$, and $(\epsilon_1 + \nu_{12}, \epsilon_2 + \nu_{21})$ uniquely determine the distribution of the data. So, for example, ν_{21} only appears in the form of $\epsilon_2 + \nu_{21}$, implying that it is not possible to identify the joint distribution of (ϵ_2, ν_{21}) without additional assumptions.

and other elliptical distributions. It is also satisfied if the components of $\epsilon|(x,z,g)$ are mutually independent, as in some of the related work. Any "entry game" with negative interaction effects satisfies this part of the assumption. And, this part of the assumption is satisfied if $N=2$.

Assumption 2.4 (Continuous utility functions)**.** *For all roles i and j, $\tilde{u}_i(\cdot,\theta_i)$ and $f_{ij}(\cdot,\delta_{ij})$ are continuous.*

Assumption 2.5 (Characteristic functions)**.** *One of the following conditions holds:*

(1) For all roles i and j, the characteristic functions of $\epsilon_i|(x,z,g)$ and $\nu_{ij}|(x,z,g)$ are non-zero on dense subsets of the real line.
(2) For all roles i and j, the characteristic functions of $\nu_{ij}|(x,z,g)$ and $(\lambda_i\epsilon_i+\lambda_j\epsilon_j)|(x,z,g)$ for any numbers (λ_i,λ_j) are analytic.
(3) For all roles i and j, $\nu_{ij}|(x,z,g)\equiv 0$.

This assumption is used to guarantee that a "pseudo-deconvolution" step of the identification strategy is valid. Many standard distributions satisfy the first condition because infinitely divisible distributions have characteristic functions without zeros. The second condition allows many zeros, but requires moment generating functions.

Assumption 2.6 (Large support regressor)**.** *For any $(x_{\cdot(-1)},z,g)$ in the support, $x_{\cdot 1}|(x_{\cdot(-1)},z,g)$ has everywhere positive density with support on $\mathbb{R}^N$.*

The following theorem shows that under these additional assumptions, the distribution of the unobservables is point identified. These assumptions are more restrictive than the assumptions used to point identify the parameters $\theta=(\theta_i)_i$ and $\delta=(\delta_{ij})_{ij}$, reflecting the additional complications involved in point identifying the distribution of the unobservables. Because of the importance of point identifying the distribution of the unobservables for some uses of the model, for example counterfactuals, it nevertheless seems worthwhile to report the result about point identification under these stronger assumptions.

Theorem 2.1. *Suppose that the model of the utility functions is given in equations 1 and 2, and suppose that there is Nash equilibrium play, in pure strategies. Suppose that θ and δ are point identified. Under assumptions 1.7, 1.8, 2.1, 2.2, 2.3, 2.4, 2.5, and 2.6, the distribution of $(\epsilon,\tilde{\nu})|(x,z,g)$ is point identified on the support of (x,z,g), where $\tilde{\nu}$ is the sub-vector of ν corresponding to connected agents according to g. If* *there is only level-2 rational play, allowing for mixed strategies, dropping assumption 2.2, then the distributions of $\epsilon_i|(x,z,g)$, and $\nu_{ij}|(x,g)$ for $g_{ij}=1$, are point identified on the support of (x,z,g) for all roles i and j.*

Corollary 2.1. *Under the same conditions, except replacing Nash equilibrium play with level-2 rational play, allowing for mixed strategies, dropping assumption 2.2, and adding the condition that $(\epsilon_1 \perp \cdots \perp \epsilon_N)|(x,z,g)$, the distribution of $(\epsilon,\tilde{\nu})|(x,z,g)$ is point identified on the support of (x,z,g).*

Remark 2.1 (Intuition for identification strategy)**.** Pure strategy Nash equilibrium implies, essentially, that "partial correlation" in the outcomes that is not due to the explanatory variables is due to correlation in the unobservables. If there is not pure strategy Nash equilibrium play, then identification of the correlation of the unobservables is "confounded" by the possibility of mis-coordinated play. Identification of the marginal distributions is possible under level-2 rationality.

Remark 2.2 (Equilibrium existence)**.** A pure strategy Nash equilibrium exists, as long as g^w is symmetric, because then the game is a potential game with potential $P(y) \equiv \sum_{i\in N}\left(u_i(y_i) + \frac{y_i}{2}\sum_{j\neq i} g^w_{ij} y_j\right)$ where $u_i(0)=0$ and $u_i(1) = x_{i1} + \tilde{u}_i(x_{i(-1)},\theta_i) + \epsilon_i$. Then see Monderer and Shapley (1996). A pure strategy Nash equilibrium can exist even if g^w is asymmetric, for example: if $N=2$, or if there is a non-negative interaction effect (i.e., Topkis (1979)), among other conditions.

Remark 2.3 (Identification of ν)**.** Only $\tilde{\nu}$ is identified; it is not possible to identify $\nu_{ij}|(x,z,g)$ when $g_{ij}=0$, because such ν_{ij} does not enter the model.

3. Proofs

Proof of theorem 2.1. Due to the independence conditions in assumption 2.3 it is enough to point identify the distribution of $\epsilon|(x,z,g)$, and of $\nu_{ij}|(x,z,g)$ when $g_{ij}=1$. There are two cases for a given (x,z,g): $\Delta_{ij}(z_{ij}) \leq 0$ for all i and j, and $\Delta_{ij}(z_{ij}) \geq 0$ for some i and j.

The first case is $\Delta_{ij}(z_{ij}) \leq 0$ for all i and j.

First, since the interaction effects are non-positive w.p.1 under assumption 2.2, the event $x_{i1} + \tilde{u}_i(x_{i(-1)},\theta_i) + \epsilon_i \leq 0$ for all i is equivalent to the event that the unique pure strategy Nash equilibrium is $(0,\ldots,0)$.[3] Thus, $P(y=(0,\ldots,0)|x,z,g) = P(\epsilon_i \leq$

[3]This ignores the probability zero event that $x_{i1} + \tilde{u}_i(x_{i(-1)},\theta_i) + \epsilon_i = 0$ for any i, by assumption 2.3. Under that condition on the utility function, action 0 is a dominant strategy for all agents,

$-x_{i1} - \tilde{u}_i(x_{i(-1)}, \theta_i)(\forall i)|x, z, g)$. $P(y = (0, \ldots, 0)|x, z, g)$ is observed on the support of (x, z, g), since the distribution of $\epsilon|(x, z, g)$ is continuous in (x, z) by assumption 2.3, the utility functions are continuous by assumption 2.4, and g is discrete. Since $\epsilon|(x, z, g)$ is conditionally independent from $x_{\cdot 1}$, by assumption 2.3, the entire distribution of $\epsilon|(x, z, g)$ can be recovered by varying $x_{\cdot 1}$, by assumption 2.6, since $-x_{i1} - \tilde{u}_i(x_{i(-1)}, \theta_i)$ is known given that θ is point identified. That follows because, by the above, the value of the cumulative distribution function of $\epsilon|(x, z, g)$ is observed at each value of its argument. Thus, the distribution of $\epsilon|(x, z, g)$ is point identified.

Second, by arguments similar to those used in the proof of the main point identification result, by taking $x_{k1} \to -\infty$ for all $k \neq i, j$, $P(y = (1, 1, 0, \ldots, 0)|x, z, g) \to P(\epsilon_i + s_i(g)\nu_{ij}g_{ij} \geq -x_{i1} - \tilde{u}_i(x_{i(-1)}, \theta_i) - s_i(g)f_{ij}(z_{ij}, \delta_{ij})g_{ij}, \epsilon_j + s_j(g)\nu_{ji}g_{ji} \geq -x_{j1} - \tilde{u}_j(x_{j(-1)}, \theta_j) - s_j(g)f_{ji}(z_{ji}, \delta_{ji})g_{ji}|x, z, g)$, where the 1s correspond to roles i and j. Since the distribution of $\epsilon|(x, z, g)$ is continuous, and $(\epsilon \perp \nu)|(x, z, g)$, by assumption 2.3, the distribution of $(\epsilon_i + s_i(g)\nu_{ij}g_{ij}, \epsilon_j + s_j(g)\nu_{ji}g_{ji})|(x, z, g)$ is continuous, so agents are indifferent among their strategies to a pure strategy of the other agent with zero probability, so the boundary conditions can be ignored. Since $(\epsilon, \nu)|(x, z, g)$ is conditionally independent from $x_{\cdot 1}$, by assumption 2.3, it is possible to recover the entire distribution of $(\epsilon_i + s_i(g)\nu_{ij}g_{ij}, \epsilon_j + s_j(g)\nu_{ji}g_{ji})|(x, z, g)$ by varying $x_{\cdot 1}$, by assumption 2.6, by the same arguments as above.[4] Thus, that distribution is point identified. In particular, the distribution of $(\epsilon_i + s_i(g)\nu_{ij})|(x, z, g)$ is point identified when $g_{ij} = 1$. Since $(\epsilon_i \perp \nu_{ij})|(x, z, g)$ by assumption 2.3, the characteristic functions satisfy $\varphi_{\epsilon_i + s_i(g)\nu_{ij}|x,z,g}(t) = \varphi_{\epsilon_i|x,z,g}(t)\varphi_{\nu_{ij}|x,z,g}(s_i(g)t)$. This implies by assumption 2.5 that the characteristic function of $\nu_{ij}|(x, z, g)$ when $g_{ij} = 1$ is point identified, since the distribution of $\epsilon_i|(x, z, g)$ is point identified from the previous paragraph, so the distribution of $\nu_{ij}|(x, z, g)$ when $g_{ij} = 1$ is point identified.[5] Therefore the distribution of the unobservables is point identified.

since $\Delta_{ij}(z_{ij}) \leq 0$ and using assumption 2.2, so $(0, \ldots, 0)$ is the unique Nash equilibrium outcome. Conversely, if $(0, \ldots, 0)$ is the pure strategy Nash equilibrium outcome, by definition that condition on the utility function must hold.

[4]Technically, this follows because any other distribution of the unobservables would have $\tilde{P}(\epsilon_i + s_i(g)\nu_{ij}g_{ij} \geq t_i, \epsilon_j + s_j(g)\nu_{ji}g_{ji} \geq t_j|x, z, g) \neq P(\epsilon_i + s_i(g)\nu_{ij}g_{ij} \geq t_i, \epsilon_j + s_j(g)\nu_{ji}g_{ji} \geq t_j|x, z, g)$ for an open set of (t_i, t_j) since the distributions are continuous. Consequently, the limit described in the first part of this paragraph would converge to a different place for different distributions of the unobservables, for a positive probability of the observables.

[5]If $\varphi_{\epsilon_i|x,z,g}(t) \neq 0$ on a dense subset, then divide by $\varphi_{\epsilon_i|x,z,g}(t)$ for t where it is non-zero to recover $\varphi_{\nu_{ij}|x,z,g}(s_i(g)t)$ using the fact that continuous functions are characterized by their values on dense subsets. Otherwise, since characteristic functions are non-zero in some neighborhood of zero, divide by $\varphi_{\epsilon_i|x,z,g}(t)$ in that neighborhood of zero. If $\varphi_{\nu_{ij}|x,z,g}(s_i(g)t)$ is analytic, the value of the characteristic function in a neighborhood of zero uniquely characterizes the characteristic function, so $\varphi_{\nu_{ij}|x,z,g}(s_i(g)t)$ is point identified (i.e, Ushakov (1999, Theorem 1.7.7)).

The second case is that $\Delta_{ij}(z_{ij}) \geq 0$ for some (but not necessarily all) i and j.

First, consider i and j such that $\Delta_{ij}(z_{ij}) \geq 0$. As before, by taking $x_{k1} \to -\infty$ for all $k \neq i, j$, $P(y = (1, 0, 0, \ldots, 0)|x, z, g) \to P(\epsilon_i \geq -x_{i1} - \tilde{u}_i(x_{i(-1)}, \theta_i), \epsilon_j + s_j(g)\nu_{ji}g_{ji} \leq -x_{j1} - \tilde{u}_j(x_{j(-1)}, \theta_j) - s_j(g)f_{ji}(z_{ji}, \delta_{ji})g_{ji}|x, z, g)$ and $P(y = (0, 1, 0, \ldots, 0)|x, z, g) \to P(\epsilon_i + s_i(g)\nu_{ij}g_{ij} \leq -x_{i1} - \tilde{u}_i(x_{i(-1)}, \theta_i) - s_i(g)f_{ij}(z_{ij}, \delta_{ij})g_{ij}, \epsilon_j \geq -x_{j1} - \tilde{u}_j(x_{j(-1)}, \theta_j)|x, z, g)$, and so $(\epsilon_i, \epsilon_j + s_j(g)\nu_{ji}g_{ji})|(x, z, g)$ and $(\epsilon_i + s_i(g)\nu_{ij}g_{ij}, \epsilon_j)|(x, z, g)$ are point identified. So in particular $\epsilon_i|(x, z, g)$ and $\epsilon_j|(x, z, g)$ are point identified. Then, since $(\epsilon_i \perp \nu_{ij})|(x, z, g)$ by assumption 2.3, the characteristic functions satisfy $\varphi_{\epsilon_i + s_i(g)\nu_{ij}g_{ij}|x,z,g}(t) = \varphi_{\epsilon_i|x,z,g}(t)\varphi_{\nu_{ij}|x,z,g}(s_i(g)t)$ when $g_{ij} = 1$. As before, this implies by assumption 2.5 that the characteristic function of $\nu_{ij}|(x, z, g)$ when $g_{ij} = 1$ is point identified, so the distribution of $\nu_{ij}|(x, z, g)$ when $g_{ij} = 1$ is point identified. Now, since $(\epsilon_i + s_i(g)\nu_{ij}g_{ij}, \epsilon_j)|(x, z, g)$ is point identified, in particular any weighted sum $(\lambda_i(\epsilon_i + s_i(g)\nu_{ij}g_{ij}) + \lambda_j\epsilon_j)|(x, z, g)$ is point identified. Again by the independence conditions, the characteristic functions satisfy $\varphi_{\lambda_i s_i(g)\nu_{ij}g_{ij} + (\lambda_i\epsilon_i + \lambda_j\epsilon_j)|x,z,g}(t) = \varphi_{\lambda_i s_i(g)\nu_{ij}g_{ij}|x,z,g}(t)\varphi_{\lambda_i\epsilon_i + \lambda_j\epsilon_j|x,z,g}(t)$. By the same arguments as before, the characteristic function of $(\lambda_i\epsilon_i + \lambda_j\epsilon_j)|(x, z, g)$ is point identified, so the distribution of $(\lambda_i\epsilon_i + \lambda_j\epsilon_j)|(x, z, g)$ is point identified. Since distributions of random vectors are uniquely determined by the distributions of all linear combinations of the components, this implies that the distribution of $(\epsilon_i, \epsilon_j)|(x, z, g)$ is point identified.

Second, consider i and j such that $\Delta_{ij}(z_{ij}) \leq 0$. By taking $x_{k1} \to -\infty$ for all $k \neq i, j$, $P(y = (1, 1, 0, \ldots, 0)|x, z, g) \to P(\epsilon_i + s_i(g)\nu_{ij}g_{ij} \geq -x_{i1} - \tilde{u}_i(x_{i(-1)}, \theta_i) - s_i(g)f_{ij}(z_{ij}, \delta_{ij})g_{ij}, \epsilon_j + s_j(g)\nu_{ji}g_{ji} \geq -x_{j1} - \tilde{u}_j(x_{j(-1)}, \theta_j) - s_j(g)f_{ji}(z_{ji}, \delta_{ji})g_{ji}|x, z, g)$ and $P(y = (0, 0, 0, \ldots, 0)|x, z, g) \to P(\epsilon_i \leq -x_{i1} - \tilde{u}_i(x_{i(-1)}, \theta_i), \epsilon_j \leq -x_{j1} - \tilde{u}_j(x_{j(-1)}, \theta_j)|x, z, g)$, so $(\epsilon_i + s_i(g)\nu_{ij}g_{ij}, \epsilon_j + s_j(g)\nu_{ji}g_{ji})|(x, z, g)$ and $(\epsilon_i, \epsilon_j)|(x, z, g)$ are point identified. And so, by the last condition of assumption 2.3, the distribution of $\epsilon|(x, z, g)$ is point identified. Then, since $(\epsilon_i \perp \nu_{ij})|(x, z, g)$ by assumption 2.3, the characteristic functions satisfy $\varphi_{\epsilon_i + s_i(g)\nu_{ij}g_{ij}|x,z,g}(t) = \varphi_{\epsilon_i|x,z,g}(t)\varphi_{\nu_{ij}|x,z,g}(s_i(g)t)$ when $g_{ij} = 1$. As before, this implies by assumption 2.5 that the characteristic function of $\nu_{ij}|(x, z, g)$ when $g_{ij} = 1$ is point identified, so the distribution of $\nu_{ij}|(x, z, g)$ when $g_{ij} = 1$ is point identified. So, the distribution of the unobservables is point identified.

By similar arguments, except taking limits for $k \neq i$, not $k \neq i, j$, the claim about point identification under level-2 rationality is established. □

References

Monderer, D., Shapley, L. S., 1996. Potential games. Games and Economic Behavior 14 (1), 124–143.

Topkis, D. M., 1979. Equilibrium points in nonzero-sum n-person submodular games. SIAM Journal of Control and Optimization 17 (6), 773–787.

Ushakov, N. G., 1999. Selected Topics in Characteristic Functions. VSP.