Computing Semiparametric Efficiency Bounds in Discrete Choice Models with Strategic-interactions and Rational Expectations

Andrés Aradillas-López*

February 6, 2020

Abstract

This paper computes semiparametric efficiency bounds for finite-dimensional parameters in discrete choice models with nonparametric regressors in the form of conditional expectations. These can include expectations about exogenous events as well as expectations about the choices of other agents. Thus, the models studied here include incomplete-information games, social-interactions models as well as single-agent discrete choice models with uncertainty as special cases. Our bounds rely on the assumption of rational expectations and on regularity conditions of equilibrium beliefs. The paper focuses on binary-choice models but the derivation of the bounds illustrates how our approach can be extended to multinomial choice cases. Explicit efficiency bound expressions for the models examined here had not been derived before. Furthermore, since we also characterize the efficient influence functions, our results can also potentially be used to construct semiparametrically efficient estimators for these models.
JEL Codes: C14, C31, C25, C35, C57.

1 Introduction

This paper computes semiparametric efficiency bounds for finite-dimensional parameters in binary choice models that include nonparametric regressors. These regressors represent agents' expectations about exogenous events and/or about the expected choices of others. They include incomplete-information games and binary choice models under uncertainty as special cases. We compute efficiency bounds assuming rational expectations and the self-consistency conditions of equilibrium beliefs. As a result, the nonparametric regressors in the model are functionals of the unknown parameters (nonparametric distributions as well as the finite-dimensional parameters of interest) and the equilibrium properties of the model determine the structure of the efficiency bounds. In their computation we rely on the method of *representers* proposed by Severini and

*email: aaradill@psu.edu. Pennsylvania State University. Department of Economics, 518 Kern Graduate Building, University Park, PA 16802

Tripathi (2001). Even though the paper focuses on binary-choice models, the description of the approach used as well as the steps of the derivation should help to illuminate how to apply this method to analyze richer discrete choice models. Semiparametric efficiency with nonparametric or "generated" regressors has been studied in *linear* models, e.g, in Rilstone (1993) and Aradillas-López (2019), and for *partially linear models* in Li and Wooldridge (2002). To my knowledge, efficiency bounds with nonparametric regressors have not been computed in discrete-choice models, particularly in the case where the distribution of unobserved shocks is nonparametric. This paper will derive these bounds, along with the efficient influence function in models where there is also a strategic-interaction component. The models studied here will encompass social-interaction models, incomplete-information games and single-agent discrete choice models with uncertainty as special cases.

The paper proceeds as follows. Section 2 presents an overview of the method of *representers*, which we will apply to compute the efficiency bounds. Section 3 presents results for a global-interaction model which assumes that agents treat the choices of everybody else symmetrically. This model is a good starting point since the approach taken there to compute the bounds extends naturally to incomplete-information games, which are studied in Section 4. Section 5 discusses efficient estimation. While our results can enable researchers to apply existing methods based on the expression of the efficient influence function, we also sketch a possible estimation strategy based on a semi-empirical likelihood approach subject to conditional moment restrictions. Concluding remarks are included in Section 6. All proofs are included in the Appendix.

2 Computing efficiency bounds using representers: an outline

Our derivation of the bounds uses the *representer* approach proposed by Severini and Tripathi (2001) (henceforth ST). This method relies on the fact that, if the parameter of interest is a pathwise differentiable functional, the efficiency bound can be obtained by using the appropriate Hilbert space, invoking the Riesz-Fréchet Theorem and finding the *representer* described in that theorem. Our discussion here follows Section 2 in ST.

Notational conventions

We will let $\mathbb{S}(z)$ denote the support of a random variable z. λ will denote the Lebesgue measure and $L^2(S,\lambda)$, the set of all real-valued functions on S that are square integrable with respect to Lebesgue measure. For a random variable z, we will let $L^2(S,\lambda_z)$ denote the set of all functions defined on S which are square integrable with respect to the probability distribution of z.

Let $z_1,\ldots,z_n$ be $d\times 1$ iid random vectors with Lebesgue density $p_0(z)$. Assume for simplicity that p_0 has full support[1] on $\mathbb{R}^d$ and let us express $p_0(z)=\tau_0^2(z)$, with $\tau_0\in\Gamma$ and Γ is a subset

[1]As the discussion that follows will illustrate, bounded support or point-masses can be readily incorporated into

of the unit ball in $L^2(\mathbb{R}^d;\lambda)$. Assume for now that τ_0 is an unknown function and therefore an infinite-dimensional parameter. In the models we will study below, τ_0 itself will be a functional of other parameters, both finite and infinite-dimensional. Working with $\tau_0 = \sqrt{p_0}$ has the advantage that $\tau_0 \in L^2(\mathbb{R}^d;\lambda)$ while this may not be the case for p_0 itself.

Denote the parameter of interest as $\rho(\tau_0) \in \mathbb{R}$, where ρ is a pathwise differentiable functional and let $\nabla\rho$ denote the pathwise derivative of ρ. Ultimately, our focus will be a finite-dimensional parameter vector θ_0, in which case $\rho(\tau_0) = c'\theta_0$, where c is an arbitrary vector[2] . The objective is to obtain efficiency bounds for regular estimators of $\rho(\tau_0)$. Regular estimators are defined in Newey (1990, page 102). In essence, they require that the asymptotic distribution of the estimator be stable in a neighborhood of the true model (i.e, in a neighborhood of τ_0).

The method described in ST for computing efficiency bounds is based on the intuition provided by Stein (1956), who introduced the notion of efficiency bounds by noting that the problem of estimating a real-valued parameter with nonparametric components is at least as difficult (to first order of approximation) as any one-dimensional subproblem contained in it. Fix some $t_0 > 0$ and let $t \mapsto \tau_t$ denote a curve from $[0, t_0]$ on to Γ that passes through τ_0 at $t = 0$ (i.e, $\tau_t|_{t=0} = \tau_0$). Let $\dot{\tau}$ denote the slope of τ_t at $t = 0$. $\dot{\tau}$ is an element of the vector space $L^2(\mathbb{R}^d;\lambda)$ which is tangent[3] to Γ at τ_0. Let $T(\Gamma,\tau_0)$ denote the tangent cone that consists of all $\dot{\tau}$'s that are tangent to Γ at τ_0. Finally, let $\overline{lin\ \mathrm{T}(\Gamma,\tau_0)}$ denote the smallest closed (in the $L^2(\mathbb{R}^d;\lambda)$ norm) linear space containing $T(\Gamma,\tau_0)$.

Let $\ell_z(t) = \log \tau_t^2(z)$. The score and the Fisher information for estimating $t = 0$ are given, respectively by

$$S_0(z) = \left.\frac{d\ell_z(t)}{dt}\right|_{t=0} = \frac{2\dot{\tau}(z)}{\tau_0(z)} \quad \text{and} \quad i_F = \int_{\mathbb{R}^d} S_0^2(z)\tau_0^2(z)dz = 4\int_{\mathbb{R}^d} \dot{\tau}^2(z)dz.$$

ST equip $\overline{lin\ \mathrm{T}(\Gamma,\tau_0)}$ with the Fisher-information inner product $\langle\cdot,\cdot\rangle_F$ defined as

$$\langle\dot{\tau}_1,\dot{\tau}_2\rangle_F = 4\int_{\mathbb{R}^d} \dot{\tau}_1(z)\dot{\tau}_2(z)dz \quad \forall\ \dot{\tau}_1,\dot{\tau}_1 \in \overline{lin\ \mathrm{T}(\Gamma,\tau_0)}.$$

Let $\widehat{t}_n$ be any regular, $\sqrt{n}-$consistent estimator of $t = 0$ in the subproblem given by τ_t and let $asyvar\left\{\sqrt{n}\cdot\widehat{t}_n\right\}$ denote the asymptotic variance of $\sqrt{n}\cdot\widehat{t}_n$. The information inequality implies that $asyvar\left\{\sqrt{n}\cdot\widehat{t}_n\right\} \geq 1/i_F = \|\dot{\tau}\|_F^{-2}$. Next, since τ_t is ultimately a device to compute efficiency bounds, we should focus on subproblems that are informative about our parameter of interest $\rho(\tau_0)$.

the analysis.

[2]This shows that focusing on the case where $\rho(\tau_0)$ is a scalar can be done without loss of generality in the case that will preoccupy us.

[3]Suppose M is a subset of a normed vector space $(X, \|\cdot\|_X)$. Take a point $x_0 \in M$. We say that a vector $\dot{x} \in X$ is tangent to M at x_0 if there exists $t_0 > 0$ and a mapping $t \mapsto r_t$ into X satisfying $\|r_t\| = o(t)$ as $t \downarrow 0$, such that $x_t \equiv x_0 + t\dot{x} + r_t \in M\ \forall\ t \in [0, t_0]$. The curve $t \mapsto x_t$ passes through x_0 at $t = 0$ and $\dot{x}$ is the slope of this curve at $t = 0$.

To this end, normalize ρ and reparameterize τ_t so that $\rho(\tau_t) = t$ for $t \in [0, t_0]$. Thus, estimating $t = 0$ will be equivalent to estimating $\rho(\tau_0)$. It follows that, for all the subproblems of interest, $asyvar\left\{\sqrt{n}\left[\rho(\tau_{\widehat{t}_n}) - \rho(\tau_0)\right]\right\} = asyvar\left\{\sqrt{n} \cdot \widehat{t}_n\right\} \geq \|\dot{\tau}\|_F^{-2}$. Next, by definition, $\nabla\rho$ is a continuous linear functional[4] and, for the suproblems we are interested in, it satisfies $\nabla\rho(\dot{\tau}) = 1$ (implying that $\dot{\tau} \neq 0$). Refer to such $\dot{\tau}$'s as *feasible.*

Thus, in searching for the lower bound (*l.b.*), we would look to maximize $\|\dot{\tau}\|_F^{-2}$ over those $\dot{\tau}$'s in $\overline{lin\ \mathrm{T}(\Gamma, \tau_0)}$ that satisfy $\dot{\tau} \neq 0$ and $\nabla\rho(\dot{\tau}) = 1$. That is,

$$l.b. = sup\left\{\|\dot{\tau}\|_F^{-2} : \dot{\tau} \in \overline{lin\ \mathrm{T}(\Gamma, \tau_0)},\, \dot{\tau} \neq 0,\, \nabla\rho(\dot{\tau}) = 1\right\}.$$

Suppose $\nabla\rho(\dot{\tau})$ is a nonzero constant (a property shared by all feasible $\dot{\tau}$'s). Then, $\widetilde{\tau} \equiv \dot{\tau}/\nabla\rho(\dot{\tau}) \in \overline{lin\ \mathrm{T}(\Gamma, \tau_0)}$. In our search for *l.b.* we can focus on such $\widetilde{\tau}$'s. Since $\nabla\rho$ is a linear functional, we have $\nabla\rho(\widetilde{\tau}) = 1$ and therefore $\widetilde{\tau}$ is feasible. Furthermore, linearity of $\nabla\rho$ implies that

$$\|\widetilde{\tau}\|_F^{-1} = \left\|\frac{\dot{\tau}}{\nabla\rho(\dot{\tau})}\right\|_F^{-1} = \frac{|\nabla\rho(\dot{\tau})|}{\|\dot{\tau}\|_F} = \left|\nabla\rho\left(\frac{\dot{\tau}}{\|\dot{\tau}\|_F}\right)\right|$$

Obviously, we have $\left\|\frac{\dot{\tau}}{\|\dot{\tau}\|_F}\right\|_F = 1$. Therefore, going back to the notation of $\dot{\tau}$ instead of $\widetilde{\tau}$, the lower bound *l.b.* can be re-expressed as

$$l.b. = sup\left\{|\nabla\rho(\dot{\tau})|^2 : \dot{\tau} \in \overline{lin\ \mathrm{T}(\Gamma, \tau_0)},\, \dot{\tau} \neq 0,\, \|\dot{\tau}\|_F = 1\right\}.$$

Since $\nabla\rho$ is a continuous linear functional on the tangent space $\overline{lin\ \mathrm{T}(\Gamma, \tau_0)}$ equipped with $\|\cdot\|_F$, its norm (see Luenberger (1969, Section 5.2)) is given by

$$\|\nabla\rho\|_* = sup\left\{|\nabla\rho(\dot{\tau})| : \dot{\tau} \in \overline{lin\ \mathrm{T}(\Gamma, \tau_0)},\, \dot{\tau} \neq 0,\, \|\dot{\tau}\|_F = 1\right\}.$$

Therefore, $l.b. = \|\nabla\rho\|_*^2$. The key insight in ST is that the problem of computing *l.b* can be solved by invoking the Riesz-Fréchet Theorem (R-F Theorem henceforth) which states[5] that, since $\left(\overline{lin\ \mathrm{T}(\Gamma, \tau_0)}, \langle\cdot,\cdot\rangle_F\right)$ is a Hilbert space and $\nabla\rho$ is a continuous, linear functional defined in it, there exists a *unique* $\tau^* \in \overline{lin\ \mathrm{T}(\Gamma, \tau_0)}$ such that

$$\nabla\rho(\dot{\tau}) = \langle\tau^*, \dot{\tau}\rangle_F \;\; \forall\, \dot{\tau} \in \overline{lin\ \mathrm{T}(\Gamma, \tau_0)} \quad \text{and} \quad \|\nabla\rho\|_* = \|\tau^*\|_F\,. \tag{R-F}$$

τ^* is called the *representer* of the linear functional $\nabla\rho$. Thus, computing *l.b.* is done in two steps:

Step 1: Find the representer τ^* by solving the condition (R-F).

[4]Let M, $\dot{x}$ and x_t be as defined in footnote 3. A functional $\rho : M \to \mathbb{R}$ is said to be pathwise differentiable at x_0 if, for any x_t, there exists a continuous linear functional $\nabla : X \to \mathbb{R}$ such that $\left|\frac{\rho(x_t)-\rho(x_0)}{t} - \nabla\rho(\dot{x})\right| \to 0$ as $t \downarrow 0$.

[5]See Luenberger (1969, Section 5.3, Theorem 2) or Young (1988, Theorem 6.8).

Step 2: Compute $l.b. = \|\tau^*\|_F^2 = \langle \tau^*, \tau^* \rangle_F$

Note that, since

$$\begin{aligned} l.b = \|\tau^*\|_F^2 = 4\int_{\mathbb{R}^d} \tau^*(z)\tau^*(z)dz &= \int_{\mathbb{R}^d} \left(2 \cdot \frac{\tau^*(z)}{\tau_0(z)}\right) \cdot \left(2 \cdot \frac{\tau^*(z)}{\tau_0(z)}\right) \tau_0^2(z)dz \\ &= E\left[\left(2\frac{\tau^*(z)}{\tau_0(z)}\right) \cdot \left(2\frac{\tau^*(z)}{\tau_0(z)}\right)\right], \end{aligned}$$

it follows that the *efficient influence function* (Newey (1990, Section 3)) is given by $\frac{2\tau^*(z)}{\tau_0(z)}$ (the score function evaluated at the *representer*). See Severini and Tripathi (2001, page 30) for more details about the connection between the representer approach and the one described in Newey (1990).

3 A global interaction model with rational expectations

Our first model is a binary-choice problem faced by a symmetric population of agents, who, prior to making their choice, have to construct beliefs (conditional expectations) about the expected choice made by a "representative" agent and about the outcome of an exogenous (i.e, non-strategic) event. This example of a social-interaction model encompasses both *global interaction models* (Brock and Durlauf (2001, Section 2.2)) and single-agent *binary choice models under uncertainty* (Ahn and Manski (1993) and Ahn (1995)) as special cases. It also extends naturally to the type of incomplete-information games with asymmetric interaction effects that we will study in Section 4.

3.1 Description of the model

Consider a population of symmetric agents, each of which must make a binary decision, labeled $y \in \{0, 1\}$. Suppose the payoff of agent i of choosing $y_i = 0$ is normalized to zero, and the payoff of choosing $y_i = 1$ is given by

$$u_i(1) = x_i' b_0 + s' d_0 + a_0 \overline{m}(1) - \nu_i,$$

where $\overline{m}(1)$ denotes the proportion of agents in the population that choose $y = 1$. $(b_0', d_0', a_0)'$ represent payoff parameters parameters. Suppose agents must choose simultaneously and before knowing the realization of s, which represents the outcome of some exogenous (i.e, non-strategic) event. The realization of $(x_i', \nu_i)'$ is observed by agent i before making her choice. We will be interested in a setting where, after choices are made, the econometrician observes the realization of (y, x', s). Suppose ν_i is a payoff shifter independent across agents and independent of s, and suppose that each agent i conditions her beliefs on the realization of x_i. We can think of x_i as the *reference variables* used by i to construct her beliefs. Accordingly, suppose the decision rule for agent i is described by the model

$$y_i = \mathbb{1}\left\{x_i' b_0 + \widehat{E}_i[s|x_i]' d_0 + a_0 \widehat{E}_i[y|x_i] - \nu_i \geq 0\right\}.$$

$(b_0', d_0', a_0)'$ are unknown parameters, and $x_i \in \mathbb{R}^{d_x}$ and $s \in \mathbb{R}^{d_s}$. The notation $\widehat{E}_i$ denotes agent i's subjective beliefs, conditional on the realization of $x_i \in \mathbb{R}^{d_x}$. Our model nests two classes of models. First, the presence of $\widehat{E}_i[y|x_i]$ as one of the regressors includes *global interaction models* studied in Brock and Durlauf (2001, Section 2.2) as a special case. In global interaction models, agents assign an identical weight to the expected choice of every other member of the population, effectively playing a game against the representative agent. Second, the presence of agents' beliefs about the exogenous (non-strategic) event $\widehat{E}_i[s|x_i]$ includes the *binary choice model under uncertainty* analyzed in Ahn and Manski (1993) and Ahn (1995) also as a special case. This is the first paper to formally derive and present semiparametric efficiency bounds for either of these models. In particular, our results will leave the distributions of all covariates involved (including unobserved shocks) nonparametrically specified.

In the context of social-interactions models (see Manski (1993), Manski (1995, Section 7.2)), x_i can be viewed as *group reference variables* and s would denote a collection of group characteristics unobserved by agents at the time of making their choices. The parameters in d_0 would capture what Manski (1995, Section 7.2) calls *contextual effects* and a_0 would capture an *endogenous effect.* If the above decision rule arises from a game-theoretic model, a_0 would capture a strategic-interaction effect while s would denote a collection of non-strategic outcomes (states of the world, etc.) that are unobserved by agents at the time of making their choices. Finally, ν_i is latent variable unobserved by the econometrician.

Except for some index-exclusion restrictions (described below), the distribution of ε_i conditional on observables will be treated nonparametrically. Therefore, we will make the type of scale normalization typically assumed for identification purposes. Suppose we can partition $x = (x_1, x_2')'$ and, accordingly $b_0 = (b_{0,1}, b_{0,2}')'$, where an underlying theoretical model predicts[6] $b_{0,1} > 0$. Denote

$$\frac{1}{b_{0,1}} \cdot \left(x_i' b_0 + \widehat{E}_i[s|x_i]' d_0 + a_0 \widehat{E}_i[y|x_i] - \nu_i\right) = x_{1,i} + x_{2,i}' \beta_0 + \widehat{E}_i[s|x_i]' \gamma_0 + \alpha_0 \widehat{E}_i[y|x_i] - \varepsilon_i,$$

where $\beta_0 \equiv \frac{b_{0,2}}{b_{0,1}}'$, $\gamma_0 \equiv \frac{d_0}{b_{0,1}}$, $\alpha_0 \equiv \frac{a_0}{b_{0,1}}$ and $\varepsilon_i \equiv \frac{\nu_i}{b_{0,1}}$. Our model is observationally equivalent to the following,

$$y_i = \mathbb{1}\left\{x_{1,i} + x_{2,i}' \beta_0 + \widehat{E}_i[s|x_i]' \gamma_0 + \alpha_0 \widehat{E}_i[y|x_i] - \varepsilon_i \geq 0\right\}. \tag{1}$$

Assumption 1A (rational expectations) *Agents use rational expectations in the construction of their beliefs. For each agent i, we have* $\widehat{E}_i[s|x] = E[s|x] \equiv \mu_0(x)$ *(the true conditional expectation*

[6]It is enough to predict that $b_{0,1} \neq 0$, as its sign can be estimated at a fast rate. If it is negative, we would use $-x_{1,i}$ instead of $x_{1,i}$ as the regressor. Thus, without loss of generality we assume that $b_{0,1} > 0$.

of s given x).

We focus on settings where the econometrician observes $(y, x', s')'$, while $\mu_0(x)$ and $\widehat{E}_i[y|x]$ are treated as nonparametric regressors. Henceforth, denote

$$\theta_0 \equiv (\beta_0', \gamma_0', \alpha_0)' \quad \text{and} \quad v_0 \equiv x_1 + x_2'\beta_0 + \mu_0(x)'\gamma_0$$

Note that v_0 is a deterministic function of x. Our bounds will be derived under the following conditions.

Assumption 2A

(i) ***(Distributional properties of ε)****: $\varepsilon|x, s \sim \varepsilon|x$. Let $Pr(\varepsilon \le \epsilon|x) \equiv G_0(\epsilon|x)$, with corresponding conditional density given by $g_0^2(\epsilon|x)$. We assume that $G_0(\epsilon|x)$ is absolutely continuous w.p.1, with support $\mathbb{R}$. For identification purposes and as a location-normalization, we assume that there exists a constant c_κ and a known $\kappa \in (0,1)$ such that $G_0(c_\kappa|x) = \kappa$ w.p.1. For the existence of a $\sqrt{n}-$consistent estimator, we assume the single-index exclusion restriction that g_0^2 depends upon x only through the index v_0. That is, $g_0^2(\cdot|x) = g_0^2(\cdot|v_0)$.*

(ii) ***(Self-consistent beliefs, regularity and selection mechanism)****: For a given $\pi \in [0,1]$ let $\varphi(x,\pi) = \pi - G_0(v_0 + \alpha_0\pi|x)$. A solution (in π) to the fixed-point problem $\varphi(x,\pi) = 0$ is regular if it satisfies $\nabla_\pi\varphi(x,\pi) \neq 0$ (i.e, $1 - \alpha_0 g_0^2(v_0 + \alpha_0\pi|x) \neq 0$). We assume that each agent i selects, as their beliefs $\widehat{E}_i[y|x]$, a regular solution to the fixed-point problem $\varphi(x,\pi) = 0$ and we denote this solution as $\pi_0(x)$. We assume that, w.p.1, all agents use the same selection mechanism to choose $\pi_0(x)$, and that this mechanism is degenerate conditional on x (i.e, it selects a unique regular solution w.p.1). Regularity implies that, if we define $D_0(x) \equiv 1 - \alpha_0 g_0^2(v_0 + \alpha_0\pi_0(x)|x)$, then $|D_0(x)| \ge \underline{d} > 0$ w.p.1.*

(iii) ***(A full-rank condition)*** *Let $v \equiv (x_2', \mu_0(x)', \pi_0(x))' \in \mathbb{R}^d$. The support of v is not contained in any proper linear subspace of $\mathbb{R}^d$*

Identification, regular estimators and Assumptions 1A-2A

While the quantile invariance assumption is done for identification purposes, it would not, by itself, guarantee the existence of a $\sqrt{n}-$consistent estimator (see, e.g, Manski (1975), Kim and Pollard (1990)). To this end, we add the single-index exclusion restriction (see Klein and Spady (1993), Ahn (1995)). Together, Assumptions 1A and 2A (in particular, the degeneracy property of the belief-selection mechanism) imply that

$$\underbrace{Pr(y = 1|x)}_{=\pi_0(x)} = Pr\Big(y = 1\Big|\underbrace{x_1 + x_2'\beta_0 + E[s|x]'\gamma_0 + \alpha_0 Pr(y = 1|x)}_{=x_1 + x_2'\beta_0 + \mu(x)'\gamma_0 + \alpha_0\pi_0(x)}\Big). \tag{2}$$

Since beliefs are nonparametrically identified regressors given our assumptions, identification and a $\sqrt{n}-$consistent estimator of θ_0 can proceed from (2), for example, using single-index model semiparametric estimators (see Powell, Stock, and Stoker (1989), Klein and Spady (1993), Ichimura (1993)). The key feature of our model is the presence of nonparametric regressors in the linear index, and the purpose of the paper is to characterize how the semiparametric efficiency bound depends on the presence of these regressors. The efficiency bounds derived here will leave the distribution of shocks nonparametrically specified, and rely only on a quantile-independence restriction as described in part (i) of our assumption.

Existence and regularity of self-consistent beliefs and Assumption 2A

Existence of an equilibrium can be established as follows. Given that beliefs are well-defined probabilities, both π and G_0 belong to the $[0,1]$ interval. From here continuity of G_0 implies that we can invoke Brouwer's fixed-point theorem to prove the existence of a solution in π to the condition $\varphi(x,\pi) = 0$ for any x (see Mas-Colell, Whinston, and Green (1995, Theorem M.I.1)). Combined with the unbounded support property of G_0, we can show the existence of at least one regular solution[7] for any x. Uniqueness of a fixed-point solution and regularity will be guaranteed if $\alpha_0 \leq 0$, since $D_0(x) = 1 - \alpha_0 g_0^2(v_0 + \alpha_0\pi_0(x)|x)$. Thus, our assumptions about equilibrium selection are non-redundant only if $\alpha_0 > 0$. Assuming a degenerate equilibrium selection mechanism has been a commonly imposed assumption for identification in a number of previous papers involving incomplete-information games (see, e.g, Brock and Durlauf (2001), Pesendorfer and Schmidt-Dengler (2008), Aradillas-López (2012)).

The assumption of *regularity* of $\pi_0(x)$ will be key to our results. Using the implicit function theorem, regularity means that π_0 is a well-behaved functional of the unknown parameters of the model. In particular, regularity allows for the following. If we let $t \to \pi_t$ denote a curve from $[0, t_0]$ that passes through π_0 at $t = 0$ and we let $\dot{\pi}(x) = \frac{d\pi_t}{dt}\left\{\pi_t(x)\right\}\Big|_{t=0}$ denote its tangent vector, then $\dot{\pi}$ will be a well-behaved functional of the tangent vectors $(\dot{f}, \dot{g}, \dot{\theta})$. This relationship will be described in Section 3.2.1 below. Our definition of regularity (as a full-rank condition for the Jacobian of the equilibrium system) is entirely analogous to the definition of a regular Walrasian equilibrium price vector in a competitive economy (see Mas-Colell, Whinston, and Green (1995, Definition 17.D.1)).

A notational simplification based on Assumption 2A

The exclusion restriction $g_0^2(\cdot|x) = g_0^2(\cdot|v_0)$ in Assumption 2A implies $\pi_0(x) = \pi_0(v_0)$ and, denoting

$$u(v_0) \equiv v_0 + \alpha_0\pi(x) = v_0 + \alpha_0\pi(v_0),$$

[7] If $1 - \alpha_0 g_0^2(v_0 + \alpha_0\pi|x) \neq 0 \; \forall \; \pi \in [0,1]$, then all solutions must be regular, but this is a much stronger condition than we assume.

then we will have $Pr(y=1|x) = \pi(v_0) = G_0(u(v_0)|v_0)$. From now on, we will abbreviate

$$G_0(u(v_0)|v_0) \equiv G_0(v_0) \quad \text{and} \quad g_0^2(u(v_0)|v_0) \equiv g_0^2(v_0),$$

and we will denote $D_0(x)$ as $D_0(v_0)$. Therefore,

$$D_0(v_0) \equiv 1 - \alpha_0 g_0^2(v_0).$$

Going forward, we will use the above notation for convenience.

3.2 Computing the efficiency bound in the global interaction model

Let

$$\begin{aligned}\mathcal{G} = \Bigg\{ & g \in L^2(\mathbb{R} \times \mathbb{R}^{d_x}; \lambda \times \lambda_x) : \ g(\varepsilon|x) = g(\varepsilon|v_0), \ \int_{\mathbb{R}} g^2(\varepsilon|x) d\varepsilon = 1, \ g^2(\epsilon|x) > 0, \ g(\epsilon|x) \text{ is bounded} \\ & \text{and continuous, and } \int_{-\infty}^{c_\kappa} g^2(\varepsilon|x) d\varepsilon = \kappa \ \text{ w.p.1.} \Bigg\}. \end{aligned} \tag{3}$$

Then, Assumption 2A implies that $g_0 \in \mathcal{G}$.
Let $f_0^2(\cdot|x)$ denote the conditional density of s given x. And let $m_0^2(\cdot)$ denote the marginal density of x. Define

$$\begin{aligned}\mathcal{F} = & \left\{ f \in L^2(\mathbb{R}^{d_s} \times \mathbb{R}^{d_x}; \lambda \times \lambda_x) : \ f(s|x) > 0, \ \int_{\mathbb{R}^{d_s}} f^2(s|x) ds = 1 \text{ w.p.1.} \right\}; \\ \mathcal{M} = & \left\{ m \in L^2(\mathbb{R}^{d_x}; \lambda) : \ m(x) > 0, \ \int_{\mathbb{R}^{d_x}} m^2(x) dx = 1 \right\}.\end{aligned} \tag{4}$$

Then, $f_0 \in \mathcal{F}$ and $m_0 \in \mathcal{M}$. For simplicity, on the case where both s and (in particular) x are continuously distributed. The steps in the proof of our main result will show how to extend this to cases where these regressors have point masses. The unknown parameters of the model are $\tau_0 = (\theta_0, g_0, f_0, m_0)$. The nonparametric regressors μ_0 are functionals of f_0 and the nonparametric regressors π_0 are functionals of f_0 and g_0.
Using the same arguments as Lemmas B.1 and B.2 in ST, the tangent spaces for $\mathcal{G}$, $\mathcal{F}$ and $\mathcal{M}$ can

be shown to be as follows,

$$
\begin{aligned}
\overline{lin\ T(\mathcal{G},g_0)} &= \Big\{ \dot{g} \in L^2(\mathbb{R}\times\mathbb{R}^{d_x};\lambda\times\lambda_x)\colon\ \dot{g}(\varepsilon|x)=\dot{g}(\varepsilon|v_0),\ \int_{\mathbb{R}} \dot{g}(\epsilon|x)g_0(\epsilon|x)d\epsilon = 0, \\
&\qquad \text{and} \quad \int_{-\infty}^{c_\kappa} \dot{g}(\epsilon|x)g_0(\epsilon|x)d\epsilon = 0 \ \ \text{w.p.1.}\Big\} \\
\overline{lin\ T(\mathcal{F},f_0)} &= \Big\{ \dot{f} \in L^2(\mathbb{R}^{d_s}\times\mathbb{R}^{d_x};\lambda\times\lambda_x)\colon\ \int_{\mathbb{R}^{d_s}} \dot{f}(s|x)f_0(s|x)ds = 0 \ \ \text{w.p.1.}\Big\} \\
\overline{lin\ T(\mathcal{M},m_0)} &= \Big\{ \dot{m} \in L^2(\mathbb{R}^{d_x};\lambda)\colon\ \int_{\mathbb{R}^{d_x}} \dot{m}(x)m_0(x)dx = 0 \Big\}.
\end{aligned}
\tag{5}
$$

Let $\dot{\tau} = (\dot{\theta}, \dot{g}, \dot{f}, \dot{m})$. This vector belongs to the product tangent space

$$
\dot{\mathscr{T}} = \mathbb{R}^d \times \overline{lin\ T(\mathcal{G},g_0)} \times \overline{lin\ T(\mathcal{H},h_0)} \times \overline{lin\ T(\mathcal{F},f_0)} \times \overline{lin\ T(\mathcal{M},m_0)}.
$$

3.2.1 Tangent vectors for the nonparametric regressors in the global interaction model

The distinguishing feature of our model is the presence of two types of nonparametric regressors: μ_0 and π_0. These are functionals of the unknown parameters of the model and, as a result, their tangent vectors $\dot{\mu}$ and $\dot{\pi}$ will be functionals of $\dot{\tau}$. While an expression for $\dot{\mu}$ will follow straightforwardly from the definition of μ_0, the characterization of $\dot{\pi}$ will rely on the regularity and self-consistency conditions of Assumption 2A.

From our previous assumptions, the conditional pmf of $y|x$, denoted as $p_0^2(y|x)$ is given by $p_0^2(y|x) = [G_0(v_0)]^y\,[1-G_0(v_0)]^{1-y}$ and the joint density of (y,s,x) is $p_0^2(y|x)f_0^2(s|x)m_0^2(x)$. For some $t_0 > 0$ let $t \mapsto \underbrace{(\theta_t, g_t, f_t, m_t)}_{\equiv \tau_t}$ be a curve from $[0,t_0]$ into $\mathbb{R}^d \times \mathcal{G}\times\mathcal{F}\times\mathcal{M}$. Let $\mu_t(x) = \int_{\mathbb{R}^{d_s}} s f_t^2(s|x)ds$ and let $\pi_t(x)$ be defined implicitly as

$$
\pi_t(x) = \int_{-\infty}^{x_1 + x_2'\beta_t + \mu_t(x)'\gamma_t + \alpha_t\pi_t(x)} g_t^2(\varepsilon|x)d\varepsilon. \tag{6}
$$

Let $\dot{\mu}(x) = \frac{d}{dt}\left\{\mu_t(x)\right\}\big|_{t=0}$. Then, $\dot{\mu}(x) = 2\int_{\mathbb{R}^{d_s}} s\dot{f}(s|x)f_0(s|x)ds$. Next, let $\dot{\pi}(x) = \frac{d}{dt}\left\{\pi_t(x)\right\}\big|_{t=0}$. We can obtain an expression for $\dot{\pi}(x)$ from (6) combined with our previous assumptions. First, define[8]

$$
\dot{\Delta}(v_0) = 2\int_{-\infty}^{u(v_0)} \dot{g}(\varepsilon|v_0)g_0(\varepsilon|v_0)d\varepsilon, \quad A_0(x) \equiv g_0^2(v_0)\cdot v', \quad B_0(v_0) \equiv g_0^2(v_0)\cdot\gamma_0'
$$

[8]Recall that we defined previously $v_0 \equiv x_1 + x_2'\beta_0 + \mu_0(x)'\gamma_0$ and $v \equiv (x_2', \mu_0(x)', \pi_0(x))'$.

From (6) and the regularity conditions in Assumption 2A we obtain

$$\dot{\pi}(x) = D_0(v_0)^{-1}\left(A_0(x)\dot{\theta} + B_0(v_0)\dot{\mu}(x) + \dot{\Delta}(v_0)\right) \tag{7}$$

3.2.2 Characterizing the score and the Fisher-information inner product in the global interaction model

As a function of t, the conditional pmf of $y|x$ is

$$p_t^2(y|x) = \left[\int_{-\infty}^{x_1+x_2'\beta_t+\mu_t(x)'\gamma_t+\alpha_t\pi_t(x)} g_t^2(\varepsilon|x)d\varepsilon\right]^y \left[1 - \int_{-\infty}^{x_1+x_2'\beta_t+\mu_t(x)'\gamma_t+\alpha_t\pi_t(x)} g_t^2(\varepsilon|x)d\varepsilon\right]^{1-y}$$

and the joint density of (y, s, x) is $q_t(y, s, x) = p_t^2(y|x)f_t^2(s|x)m_t^2(x)$. Denote

$$T_0(v_0) \equiv G_0(v_0)\left[1 - G_0(v_0)\right].$$

The score for estimating $t = 0$ in this model becomes

$$S_0 = \left(\frac{y - G_0(v_0)}{T_0(v_0)}\right) \cdot \left[g_0^2(v_0)\left(v'\dot{\theta} + \gamma_0'\dot{\mu}(x) + \alpha_0\dot{\pi}(x)\right) + \dot{\Delta}(v_0)\right] + 2\frac{\dot{f}(s|x)}{f_0(s|x)} + 2\frac{\dot{m}(x)}{m_0(x)}$$

Using the expression in (7), the score simplifies to

$$\begin{aligned} S_0 = \left(\frac{y - G_0(v_0)}{T_0(v_0)}\right) \cdot \Bigg[& g_0^2(v_0) \cdot \left(v' + \alpha_0 D_0(v_0)^{-1}A_0(x)\right)\dot{\theta} \\ & + g_0^2(v_0) \cdot \left(\gamma_0' + \alpha_0 D_0(v_0)^{-1}B_0(v_0)\right)\dot{\mu}(x) \\ & + \left(g_0^2(v_0)\alpha_0 D_0(v_0)^{-1} + 1\right)\dot{\Delta}(v_0)\Bigg] + 2\frac{\dot{f}(s|x)}{f_0(s|x)} + 2\frac{\dot{m}(x)}{m_0(x)} \end{aligned} \tag{8}$$

Let

$$\begin{aligned} \underbrace{M_\theta(x)}_{d\times 1} &= \left(v + A_0(x)'D_0(v_0)^{-1}\alpha_0\right) \cdot \frac{g_0^2(v_0)}{\sqrt{T_0(v_0)}} \\ \underbrace{M_\mu(v_0)}_{d_s\times 1} &= \left(\gamma_0 + B_0(v_0)'D_0(v_0)^{-1}\alpha_0\right) \cdot \frac{g_0^2(v_0)}{\sqrt{T_0(v_0)}} \\ \underbrace{M_\Delta(v_0)}_{1\times 1} &= \left(g_0^2(v_0)D_0(v_0)^{-1}\alpha_0 + 1\right) \cdot \frac{1}{\sqrt{T_0(v_0)}} \end{aligned} \tag{9}$$

Using iterated expectations, we have

$$E\left[S_0^2\right] = E\left[\left(M_\theta(x)'\dot{\theta} + M_\mu(x)'\dot{\mu}(x) + M_\Delta(v_0)\dot{\Delta}(v_0)\right)^2\right] + 4E\left[\int_{\mathbb{R}^{d_s}} \dot{f}(s|x)^2 ds\right] + 4\int_{\mathbb{R}^{d_x}} \dot{m}(x)^2 dx$$

And the Fisher-information inner product is therefore given by

$$
\begin{aligned}
&\langle \dot\tau_1, \dot\tau_2 \rangle_F = \\
&E\left[\Big(M_\theta(x)'\dot\theta_1 + M_\mu(x)'\dot\mu_1(x) + M_\Delta(v_0)\dot\Delta_1(v_0)\Big)\cdot\Big(M_\theta(x)'\dot\theta_2 + M_\mu(x)'\dot\mu_2(x) + M_\Delta(v_0)\dot\Delta_2(v_0)\Big)\right] \\
&+ 4E\left[\int_{\mathbb{R}^{d_s}} \dot f_1(s|x)\dot f_2(s|x)ds\right] + 4\int_{\mathbb{R}^{d_x}} \dot m_1(x)\dot m_2(x)dx
\end{aligned} \tag{10}
$$

From here, using the representer method outlined in Section 2 we compute the semiparametric efficiency bound for $\sqrt{n}-$consistent, regular estimators of $c'\theta_0$, for an arbitrary c, by finding the representer $\tau^* \in \dot{\mathscr{T}}$ that satisfies $\langle \tau^*, \dot\tau \rangle_F = c'\dot\theta \;\forall\; \dot\tau \in \dot{\mathscr{T}}$. Once the representer is characterized, the efficiency bound is given by $l.b = \|\tau^*\|_F^2 = \langle \tau^*, \tau^* \rangle_F$. Since c is arbitrary, the efficiency bound for $\sqrt{n}-$consistent, regular estimators of θ_0 will be obtained from here.

The efficient influence function for for $\sqrt{n}-$consistent, regular estimators of $c'\theta_0$ is given by the score function evaluated at the *representer*. Using (8) and (9), once the representer τ^* is computed, the efficient influence function is given by

$$
\varphi = \left(\frac{y - G_0(v_0)}{\sqrt{T_0(v_0)}}\right)\cdot\Big[M_\theta(x)'\theta^* + M_\mu(v_0)'\mu^*(x) + M_\Delta(v_0)\Delta^*(v_0)\Big] + 2\frac{f^*(s|x)}{f_0(s|x)} + 2\frac{m^*(x)}{m_0(x)} \tag{11}
$$

Since c is arbitrary, the efficient influence function for $\sqrt{n}-$consistent, regular estimators of θ_0 will be obtained from here.

3.2.3 Efficiency bound for the global interaction model

Proposition 1 *Let M_θ and M_μ be as defined in* (9) *and let*

$$
\begin{aligned}
\underbrace{\Gamma(x)}_{1\times 1} &= \big(1 + M_\mu(v_0)'Var[s|x]M_\mu(v_0)\big)^{-1}, \\
\underbrace{\Phi^*(x)}_{d\times 1} &= \Big(M_\theta(x) - E[M_\theta(x)\Gamma(x)|v_0]\cdot E\left[\Gamma(x)|v_0\right]^{-1}\Big)\Gamma(x)
\end{aligned}
$$

and

$$
\Sigma_\theta^* = E\left[\Phi^*(x)\Phi^*(x)'\right] + E\left[\Phi^*(x)M_\mu(v_0)'Var[s|x]M_\mu(v_0)\Phi^*(x)'\right].
$$

If Σ_θ^* is invertible, the semiparametric efficiency bound for $\sqrt{n}$-consistent, regular estimators of θ_0 in Model (1) *under Assumptions 1A-2A is well-defined and is equal to Σ_θ^{*-1}. And the efficient influence function is given by*

$$
\psi(y, s, x, \theta_0) = \Sigma_\theta^{*-1}\Phi^*(x)\cdot\Big[T_0(v_0)^{-1/2}\big(y - G_0(v_0)\big) - M_\mu(v_0)'\big(s - E[s|x]\big)\Big].
$$

Invertibility of Σ_θ^* is a key condition for $\sqrt{n}-$consistency and a finite efficiency bound. Note that the expressions in Proposition 1 simplify to

$$\Gamma(x) = \frac{D_0(v_0)\sqrt{T_0(v_0)}}{T_0(v_0)\cdot D_0(v_0)^2 + \gamma_0' Var[s|x]\gamma_0 \cdot g_0^4(v_0)},$$
$$\Phi^*(x) = \left(v - \frac{E\left[v\cdot\Gamma(x)|v_0\right]}{E\left[\Gamma(x)|v_0\right]}\right)\cdot\Gamma(x) g_0^2(v_0),$$

Proof: In the appendix. ■

4 A binary choice game with incomplete information

Here we extend the global interaction model of the previous section to an incomplete-information game with (potentially) asymmetric interaction effects. As our results will show, the efficiency bound can be described as a generalization of the expression derived in Proposition 1.

4.1 Description of the game

The game is played between P players, labeled $q = 1, \ldots, P$, each of which must choose a binary action $y_q \in \{0,1\}$. We will follow the notational convention of letting the subscript '$-q$' refer to all players except q. In particular, $y_{-q} \equiv (y_1, \ldots, y_{q-1}, y_{q+1}, \ldots, y_P)'$. Choices are made simultaneously. Let now i refer to the i^{th} game. Suppose the payoff of player q in game i of choosing $y_q = 0$ is normalized to zero, and the payoff of choosing $y_q = 1$ is given by

$$u_{q,i}(1) = x_{q,i}' b_{q0} + s' d_{q0} + y_{-q,i}' a_{q0} - \nu_{q,i}.$$

The strategic interaction parameters for player q are given by

$$a_{q0} = \left(a_{q0}^1, \ldots, a_{q0}^{q-1}, a_{q0}^{q+1}, \ldots, a_{q0}^P\right).$$

a_{q0}^r captures the strategic effect of player r on player q. We allow for pairwise asymmetries in these effects: we can have $a_{q0}^r \neq a_{q0}^{r'}$ and/or $a_{q0}^r \neq a_{r0}^q$. The overall strategic interaction effect on player q is summarized by the term

$$y_{-q,i}' a_{q0} = \sum_{r\neq q} y_{r,i}\cdot a_{q0}^r.$$

Once again, suppose agents must make their choices simultaneously and before observing the realization of s. Group $\underset{\sim}{x} \equiv \underset{q}{\cup}\, x_q \in \mathbb{R}^{d_x}$. Suppose that, prior to making her choice, every player q observes the realization of $\underset{\sim}{x}$ and of ν_q, but that the latter is private information for player q. Suppose ν_q is independent of ν_{-q} and that agents condition their subjective beliefs on the realization of $\underset{\sim}{x}$. Once again, agents maximize their expected payoff given their beliefs. Accordingly, the decision rule for player q in game i is described by the model

$$y_{q,i} = \mathbb{1}\{x_{q,i}'b_{q0} + \widehat{E}_{q,i}[s|\underset{\sim}{x}_i]'d_{q0} + \widehat{E}_{q,i}[y_{-q}|\underset{\sim}{x}_i]'a_{q0} - \nu_{q,i} \geq 0\}.$$

As before, $\widehat{E}_{q,i}$ denotes the subjective expectation of player q, conditional on $\underset{\sim}{x}_i$. As in the global-interaction model we need a scale normalization. Partition $x_q = (x_{1q}, x_{2q}')'$ and $b_{q0} = (b_{q0,1}, b_{q0,2}')'$, where an underlying theoretical model predicts[9] $b_{q0,1} > 0$. Let $\beta_{q0} \equiv \frac{b_{q0,2}'}{b_{q0,1}}$, $\gamma_{q0} \equiv \frac{d_{q0}'}{b_{q0,1}}$, $\alpha_{q0} \equiv \frac{a_{q0}'}{b_{q0,1}}$ and $\varepsilon_{q,i} \equiv \frac{\nu_{q,i}}{b_{q0,1}}$. This model is observationally equivalent to

$$y_{q,i} = \mathbb{1}\left\{x_{1q,i} + x_{2q,i}'\beta_{q0} + \widehat{E}_{q,i}\left[s|\underset{\sim}{x}_i\right]'\gamma_{q0} + \widehat{E}_{q,i}\left[y_{-q}|\underset{\sim}{x}_i\right]'\alpha_{q0} - \varepsilon_{q,i} \geq 0\right\}. \tag{12}$$

The decision rule in (12) is a direct counterpart of Equation (1) in the global interactions model studied previously. We will let

$$\theta_{q0} \equiv \left(\beta_{q0}', \gamma_{q0}', \alpha_{q0}'\right)' \in \mathbb{R}^{d_q}, \quad \text{and} \quad \theta_0 \equiv \left(\theta_{10}', \theta_{20}', \ldots, \theta_{P0}'\right)' \in \mathbb{R}^d.$$

Note that

$$\alpha_{q0} = \left(\alpha_{q0}^1, \ldots, \alpha_{q0}^{q-1}, \alpha_{q0}^{q+1}, \ldots, \alpha_{q0}^P\right).$$

where α_{q0}^r captures the (normalized) strategic effect of player r on player q. As we did in the global interaction case, we will maintain rational expectations and regularity of equilibrium beliefs.

Assumption 1B ***(rational expectations in the game)*** *We assume that $\varepsilon_q \perp \varepsilon_r | \underset{\sim}{x}$ for all $q \neq r$, and the realization of $\underset{\sim}{x}$ is public information. Accordingly, players condition their beliefs on $\underset{\sim}{x}$, as described in* (12). *Players use rational expectations in the construction of their beliefs. Thus, $\widehat{E}_{q,i}\left[s|\underset{\sim}{x}_i\right] = E\left[s|\underset{\sim}{x}_i\right] \equiv \mu_0(\underset{\sim}{x})$ (the true conditional expectation of s given $\underset{\sim}{x}$).*

Similarly to the global interaction model studied previously, we will define

$$v_{q0} = x_{1q} + x_{2q}'\beta_{q0} + \mu_0(x)'\gamma_{q0}.$$

In what follows, it will be useful to define

$$\overline{\alpha}_{q0} = \left(\alpha_{q0}^1, \ldots, \alpha_{q0}^{q-1}, 0, \alpha_{q0}^{q+1}, \ldots, \alpha_{q0}^P\right),$$

so we can express $\widehat{E}_{q,i}\left[y_{-q}|\underset{\sim}{x}_i\right]'\alpha_{q0} = \widehat{E}_{q,i}\left[y|\underset{\sim}{x}_i\right]'\overline{\alpha}_{q0}$ and therefore by Assumption 1B, Equation (12) becomes

$$y_{q,i} = \mathbb{1}\left\{v_{q0,i} + \widehat{E}_{q,i}\left[y|\underset{\sim}{x}_i\right]'\overline{\alpha}_{q0} - \varepsilon_{q,i} \geq 0\right\}.$$

[9]See footnote 6.

As we did in the global interaction model, we will obtain efficiency bounds here assuming *regularity* of the solution to the equilibrium conditions of the game. To this end we will extend the conditions in Assumption 2A to this setting.

Assumption 2B

(i) ***(Distributional properties of ε_q):*** *$\varepsilon_q|\underset{\sim}{x}, s \sim \varepsilon_q|\underset{\sim}{x}$. Let $Pr(\varepsilon_q \leq \epsilon|\underset{\sim}{x}) \equiv G_{q0}(\epsilon|\underset{\sim}{x})$, with corresponding density given by $g^2_{q0}(\epsilon|\underset{\sim}{x})$. We assume that $G_{q0}(\epsilon|\underset{\sim}{x})$ is absolutely continuous w.p.1, with support $\mathbb{R}$. For identification purposes and as a location-normalization, we assume that there exists a constant c_κ and a known $\kappa \in (0,1)$ such that $G_{q0}(c_\kappa|\underset{\sim}{x}) = \kappa$ w.p.1. For the existence of a $\sqrt{n}-$consistent estimator, we will assume a multiple-index exclusion restriction. Let $\underset{\sim}{v}_0 \equiv (v_{10}, v_{20}, \ldots, v_{P0})'$. Then each g^2_{q0} depends on $\underset{\sim}{x}$ only through $\underset{\sim}{v}_0$. That is, $g^2_{q0}(\cdot|\underset{\sim}{x}) = g^2_{q0}(\cdot|\underset{\sim}{v}_0)$ for each $q = 1, \ldots, P$.*

(ii) ***(Equilibrium beliefs, regularity and selection mechanism):***
For a given $\underset{\sim}{\pi} \equiv (\pi_1, \ldots, \pi_P)' \in [0,1]^P$, let

$$\begin{aligned}\underset{\sim}{H}(\underset{\sim}{x}, \underset{\sim}{\pi}) &= \left(G_{10}(v_{10} + \underset{\sim}{\pi}'\overline{\alpha}_{10}|\underset{\sim}{x}),\, G_{20}(v_{20} + \underset{\sim}{\pi}'\overline{\alpha}_{20}|\underset{\sim}{x}), \ldots, G_{P0}(v_{P0} + \underset{\sim}{\pi}'\overline{\alpha}_{P0}|\underset{\sim}{x})\right)', \\ \underset{\sim}{\varphi}(\underset{\sim}{x}, \underset{\sim}{\pi}) &= \underset{\sim}{\pi} - \underset{\sim}{H}(\underset{\sim}{x}, \underset{\sim}{\pi}).\end{aligned}$$

A solution (in $\underset{\sim}{\pi}$) to the fixed-point problem $\underset{\sim}{\varphi}(\underset{\sim}{x}, \underset{\sim}{\pi}) = 0$ is regular if the Jacobian $\nabla_{\underset{\sim}{\pi}}\underset{\sim}{\varphi}(\underset{\sim}{x}, \underset{\sim}{\pi})$ is invertible. We assume that players select, as their beliefs $\widehat{E}_q[y|\underset{\sim}{x}]$, a regular solution to the fixed point problem $\underset{\sim}{\varphi}(\underset{\sim}{x}, \underset{\sim}{\pi}) = 0$ and we denote this solution as $\underset{\sim}{\pi}_0(\underset{\sim}{x})$. We assume that, w.p.1, players use the same selection mechanism to choose $\underset{\sim}{\pi}_0(\underset{\sim}{x})$, and that this mechanism is degenerate conditional on $\underset{\sim}{x}$ (i.e, it selects a unique regular solution w.p.1). Regularity implies that $\nabla_{\underset{\sim}{\pi}}\underset{\sim}{\varphi}(\underset{\sim}{x}, \underset{\sim}{\pi}_0(\underset{\sim}{x})) \equiv \underset{\sim}{D}_0(\underset{\sim}{x})$ is invertible w.p.1.

(iii) ***(A full-rank condition)*** *For each q let $v_q \equiv (x'_{2q}, \mu_0(\underset{\sim}{x})', \underset{\sim}{\pi}_0(\underset{\sim}{x})') \in \mathbb{R}^{d_q}$. The support of v_q is not contained in any proper linear subspace of $\mathbb{R}^{d_q}$.*

Identification, regular estimators and Assumptions 1B- 2B

Assumptions 1B and 2B (in particular, the degeneracy property of the equilibrium selection mechanism) combined imply that

$$Pr(y_q = 1|\underset{\sim}{x}) = Pr\left(y_q = 1 \;\middle|x_{1q} + x'_{2q}\beta_{q0} + E[s|\underset{\sim}{x}]'\gamma_{q0} + E[y_{-q}|\underset{\sim}{x}]'\alpha_{q0}\right) \quad \forall\, q. \tag{13}$$

Since the nonparametric regressors that appear in (13) are identified, identification of θ_0 and a $\sqrt{n}-$consistent estimator can follow from here, for example, using procedures designed for multiple-index models (see Ichimura and Lee (1991), Donkers and Schafgans (2008), Ahn, Ichimura, Powell, and Ruud (2018)). The main feature of our model is that these indices depend on nonparametric regressors (beliefs) and the goal of the paper is to characterize how the presence of these nonparametric regressors affects the semiparametric efficiency bounds. As in the global interaction model studied previously, the efficiency bounds we will derive for the incomplete-information game leave the distribution of shocks nonparametrically specified, relying only on a quantile-independence restriction[10].

Existence and regularity of equilibrium beliefs and Assumption 2B

As in the global interactions model, existence of equilibrium beliefs in our game follows from Brouwer's fixed point theorem, since both $\underset{\sim}{\pi}$ and $\underset{\sim}{H}(\underset{\sim}{x}, \underset{\sim}{\pi})$ belong in $[0,1]^P$ and $\underset{\sim}{H}$ is continuous. A sufficient condition for all equilibria to be regular would be for the Jacobian $\nabla_{\underset{\sim}{\pi}} \underset{\sim}{\varphi}(\underset{\sim}{x}, \underset{\sim}{\pi})$ to be invertible for all $\underset{\sim}{\pi} \in [0,1]^P$, but this is much stronger than what we need.

A notational simplification based on Assumption 2B

The exclusion restriction $g^2_{q0}(\cdot|\underset{\sim}{x}) = g^2_{q0}(\cdot|\underset{\sim}{v}_0)$ implies that equilibrium beliefs $\underset{\sim}{\pi}_0(\underset{\sim}{x})$ are functionals of $\underset{\sim}{v}_0$. That is, $\underset{\sim}{\pi}_0(\underset{\sim}{x}) = \underset{\sim}{\pi}_0(\underset{\sim}{v}_0)$. Extending the notation we used previously in the global interactions model we will denote from now on

$$\begin{aligned} u_{q0}(\underset{\sim}{v}_0) &\equiv v_{q0} + \underset{\sim}{\pi}_0(\underset{\sim}{v}_0)' \overline{\underset{\sim}{\alpha}}_{q0} \\ g^2_{q0}(\underset{\sim}{v}_0) &\equiv g^2_{q0}(\underset{\sim}{v}_0|\underset{\sim}{v}_0), \\ G_{q0}(\underset{\sim}{v}_0) &\equiv G_{q0}(\underset{\sim}{v}_0|\underset{\sim}{v}_0), \end{aligned}$$

And letting $\underset{\sim}{I}_P$ be the $P \times P$ identity matrix, the Jacobian described in part (ii) of Assumption 2B can be expressed as a functional of $\underset{\sim}{v}_0$,

$$\underset{\sim}{D}_0(\underset{\sim}{v}_0) \equiv \nabla_{\underset{\sim}{\pi}} \underset{\sim}{\varphi}(\underset{\sim}{x}, \underset{\sim}{\pi}_0(\underset{\sim}{v}_0)) = \underset{\sim}{I}_P - \underbrace{\begin{pmatrix} g^2_{10}(\underset{\sim}{v}_0) \cdot \overline{\underset{\sim}{\alpha}}'_{10} \\ g^2_{20}(\underset{\sim}{v}_0) \cdot \overline{\underset{\sim}{\alpha}}'_{20} \\ \vdots \\ g^2_{P0}(\underset{\sim}{v}_0) \cdot \overline{\underset{\sim}{\alpha}}'_{P0} \end{pmatrix}}_{P \times P}$$

Regularity implies that $\underset{\sim}{D}_0(\underset{\sim}{v}_0)$ is invertible w.p.1.

4.2 Computing the efficiency bound in the incomplete-information game

The conditions in Assumption 2B imply that each g_{q0} belongs in the space $\mathcal{G}$ described in Equation (3). Furthermore, if we denote $dim(\underset{\sim}{x}) \equiv d_x$ and $dim(s) \equiv d_s$ we let $f^2_0(\cdot|\underset{\sim}{x})$ denote the conditional

[10]Identification in incomplete-information games has been achieved under alternative assumptions, such as excluded-regressor restrictions in Lewbel and Tang (2015).

density of s given $\underset{\sim}{x}$ and we let $m_0^2(\underset{\sim}{x})$ denote the marginal density of $\underset{\sim}{x}$, then $f_0 \in \mathcal{F}$ and $m_0 \in \mathcal{M}$, as described in Equation (4). Accordingly, the corresponding tangent spaces $\overline{lin\ T(\mathcal{G}, g_{q0})}$, $\overline{lin\ T(\mathcal{F}, f_0)}$ and $\overline{lin\ T(\mathcal{M}, m_0)}$ are as described in Equation 5. The unknown parameters of this model are $\tau_0 = \big(\theta_0, \{g_{q0}\}_{q=1}^P, f_0, m_0\big) \in \underset{\sim}{\Gamma}$. The nonparametric regressors μ_0 and $\underset{\sim}{\pi}_0$ are a functionals of f_0 and $\{g_{q0}\}_{q=1}^P$, respectively. As we have done before, fix some $t_0 > 0$ and let $t \mapsto \tau_t$ denote a curve from $[0, t_0]$ to $\underset{\sim}{\Gamma}$ that passes through τ_0 at $t = 0$ and let $\dot{\tau} = (\dot{\theta}, \{\dot{g}_q\}_{q=1}^P, \dot{f}, \dot{m})$ denote the slope of τ_t at $t = 0$. This vector belongs to the product tangent space

$$\dot{\mathscr{T}} = \mathbb{R}^d \times \left(\times_{q=1}^P \overline{lin\ T(\mathcal{G}, g_{q0})}\right) \times \overline{lin\ T(\mathcal{H}, h_0)} \times \overline{lin\ T(\mathcal{F}, f_0)} \times \overline{lin\ T(\mathcal{M}, m_0)}.$$

4.2.1 Tangent vectors for the nonparametric regressors in the incomplete-information game

Our model has once again two types of nonparametric regressors. The rational expectations μ_0 for the non-strategic outcome s and the equilibrium beliefs $\underset{\sim}{\pi}_0$. These are functionals of the unknown parameters in the model and their tangent vectors $\dot{\mu}$ and $\dot{\underset{\sim}{\pi}}$ are functionals of $\dot{\tau}$. As in the global interaction model, $\dot{\mu}$ will follow straightforwardly from the definition of μ_0, while $\dot{\underset{\sim}{\pi}}$ will be derived invoking the assumed regularity properties of equilibrium beliefs.

By definition, $\dot{\mu}(\underset{\sim}{x}) = \frac{d}{dt}\left\{\mu_t(\underset{\sim}{x})\right\}\big|_{t=0}$ and $\dot{\underset{\sim}{\pi}}(\underset{\sim}{x}) = \frac{d}{dt}\left\{\underset{\sim}{\pi}_t(\underset{\sim}{x})\right\}\big|_{t=0}$. Since $\mu_t(\underset{\sim}{x}) = \int_{\mathbb{R}^{d_s}} s f_t^2(s|\underset{\sim}{x})ds$, we have $\dot{\mu}(\underset{\sim}{x}) = 2\int_{\mathbb{R}^{d_s}} s\dot{f}(s|\underset{\sim}{x})f_0(s|\underset{\sim}{x})ds$. Next, $\underset{\sim}{\pi}_t(\underset{\sim}{x})$ can be defined implicitly by the system

$$\underset{\sim}{\pi}_t(\underset{\sim}{x}) = \begin{pmatrix} \pi_{1t}(\underset{\sim}{x}) \\ \pi_{2t}(\underset{\sim}{x}) \\ \vdots \\ \pi_{Pt}(\underset{\sim}{x}) \end{pmatrix} = \begin{pmatrix} \int_{-\infty}^{x_{11}+x_{21}'\beta_{1t}+\mu_t(\underset{\sim}{x})'\gamma_{1t}+\underset{\sim}{\pi}_t(\underset{\sim}{x})'\overline{\alpha}_{1t}} g_{1t}^2(\varepsilon|\underset{\sim}{x})d\varepsilon \\ \int_{-\infty}^{x_{12}+x_{22}'\beta_{2t}+\mu_t(\underset{\sim}{x})'\gamma_{2t}+\underset{\sim}{\pi}_t(\underset{\sim}{x})'\overline{\alpha}_{2t}} g_{2t}^2(\varepsilon|\underset{\sim}{x})d\varepsilon \\ \vdots \\ \int_{-\infty}^{x_{1P}+x_{2P}'\beta_{Pt}+\mu_t(\underset{\sim}{x})'\gamma_{Pt}+\underset{\sim}{\pi}_t(\underset{\sim}{x})'\overline{\alpha}_{Pt}} g_{Pt}^2(\varepsilon|\underset{\sim}{x})d\varepsilon \end{pmatrix}$$

where $\overline{\alpha}_{qt} = \left(\alpha_{qt}^1, \ldots, \alpha_{qt}^{q-1}, 0, \alpha_{qt}^{q+1}, \ldots, \alpha_{qt}^P\right)'$. Letting

$$\dot{\Delta}_q(\underset{\sim}{v}_0) = 2\int_{-\infty}^{u_q(\underset{\sim}{v}_0)} \dot{g}_q(\varepsilon|\underset{\sim}{v}_0)\cdot g_{q0}(\varepsilon|\underset{\sim}{v}_0)d\varepsilon,$$

we have

$$\dot{\underset{\sim}{\pi}}(\underset{\sim}{x}) = \begin{pmatrix} \dot{\pi}_1(\underset{\sim}{x}) \\ \dot{\pi}_2(\underset{\sim}{x}) \\ \vdots \\ \dot{\pi}_P(\underset{\sim}{x}) \end{pmatrix} = \begin{pmatrix} g_{10}^2(\underset{\sim}{v}_0) \cdot \left(v_1' \dot{\theta}_1 + \dot{\mu}(\underset{\sim}{x})' \gamma_{10} + \dot{\underset{\sim}{\pi}}(\underset{\sim}{x})' \overline{\alpha}_{10} \right) + \dot{\Delta}_1(\underset{\sim}{v}_0) \\ g_{20}^2(\underset{\sim}{v}_0) \cdot \left(v_2' \dot{\theta}_2 + \dot{\mu}(\underset{\sim}{x})' \gamma_{20} + \dot{\underset{\sim}{\pi}}(\underset{\sim}{x})' \overline{\alpha}_{20} \right) + \dot{\Delta}_2(\underset{\sim}{v}_0) \\ \vdots \\ g_{P0}^2(\underset{\sim}{v}_0) \cdot \left(v_P' \dot{\theta}_P + \dot{\mu}(\underset{\sim}{x})' \gamma_{P0} + \dot{\underset{\sim}{\pi}}(\underset{\sim}{x})' \overline{\alpha}_{P0} \right) + \dot{\Delta}_P(\underset{\sim}{v}_0) \end{pmatrix}$$

$$\Longrightarrow \underbrace{\begin{pmatrix} \underset{\sim}{I}_P - \begin{matrix} g_{10}^2(\underset{\sim}{v}_0) \cdot \overline{\alpha}_{10}' \\ g_{20}^2(\underset{\sim}{v}_0) \cdot \overline{\alpha}_{20}' \\ \vdots \\ g_{P0}^2(\underset{\sim}{v}_0) \cdot \overline{\alpha}_{P0}' \end{matrix} \end{pmatrix}}_{=\underset{\sim}{D}_0(\underset{\sim}{v}_0)} \cdot \begin{pmatrix} \dot{\pi}_1(\underset{\sim}{x}) \\ \dot{\pi}_2(\underset{\sim}{x}) \\ \vdots \\ \dot{\pi}_P(\underset{\sim}{x}) \end{pmatrix} = \begin{pmatrix} g_{10}^2(\underset{\sim}{v}_0) \cdot v_1' \dot{\theta}_1 \\ g_{20}^2(\underset{\sim}{v}_0) \cdot v_2' \dot{\theta}_2 \\ \vdots \\ g_{P0}^2(\underset{\sim}{v}_0) \cdot v_P' \dot{\theta}_P \end{pmatrix} + \begin{pmatrix} g_{10}^2(\underset{\sim}{v}_0) \cdot \gamma_{10}' \dot{\mu}(\underset{\sim}{x}) \\ g_{20}^2(\underset{\sim}{v}_0) \cdot \gamma_{20}' \dot{\mu}(\underset{\sim}{x}) \\ \vdots \\ g_{P0}^2(\underset{\sim}{v}_0) \cdot \gamma_{P0}' \dot{\mu}(\underset{\sim}{x}) \end{pmatrix} + \begin{pmatrix} \dot{\Delta}_1(\underset{\sim}{v}_0) \\ \dot{\Delta}_2(\underset{\sim}{v}_0) \\ \vdots \\ \dot{\Delta}_P(\underset{\sim}{v}_0) \end{pmatrix}$$

Thus, if we define

$$\underset{\sim}{A}_0(\underset{\sim}{x}) = \underbrace{\begin{pmatrix} g_{10}^2(\underset{\sim}{v}_0) \cdot v_1' & 0' & \cdots & 0' \\ 0' & g_{20}^2(\underset{\sim}{v}_0) \cdot v_2' & \cdots & 0' \\ \vdots & \vdots & \ddots & \vdots \\ 0' & 0' & 0' & g_{P0}^2(\underset{\sim}{v}_0) \cdot v_P' \end{pmatrix}}_{P \times d}, \quad \dot{\underset{\sim}{\Delta}}(\underset{\sim}{v}_0) = \underbrace{\left(\dot{\Delta}_1(\underset{\sim}{v}_0),\, \dot{\Delta}_2(\underset{\sim}{v}_0), \ldots, \dot{\Delta}_P(\underset{\sim}{v}_0) \right)'}_{P \times 1}$$

$$\underset{\sim}{B}_0(\underset{\sim}{v}_0) = \underbrace{\begin{pmatrix} g_{10}^2(\underset{\sim}{v}_0) \cdot \gamma_{10}' \\ g_{20}^2(\underset{\sim}{v}_0) \cdot \gamma_{20}' \\ \vdots \\ g_{P0}^2(\underset{\sim}{v}_0) \cdot \gamma_{P0}' \end{pmatrix}}_{P \times d_s}, \quad \dot{\theta} = \underbrace{\left(\dot{\theta}_1,\, \dot{\theta}_2, \ldots, \dot{\theta}_P \right)'}_{d \times 1}$$

Then, by the regularity condition in Assumption 2B we can express $\underset{\sim}{\pi}_t(\underset{\sim}{x})$ compactly as

$$\dot{\underset{\sim}{\pi}}(\underset{\sim}{x}) = \underset{\sim}{D}_0(\underset{\sim}{v}_0)^{-1} \cdot \left(\underset{\sim}{A}_0(\underset{\sim}{x}) \dot{\theta} + \underset{\sim}{B}_0(\underset{\sim}{v}_0) \dot{\mu}(\underset{\sim}{x}) + \dot{\underset{\sim}{\Delta}}(\underset{\sim}{v}_0) \right). \tag{14}$$

This is a direct generalization of Equation (7) in the global interactions model studied previously.

4.2.2 Characterizing the score and the Fisher-information inner product in the incomplete-information game

As a function of t, the conditional pmf of $y_q|\underset{\sim}{x}$ is

$$p^2_{qt}(y_q|\underset{\sim}{x}) =$$
$$\left[\int_{-\infty}^{x_{1q}+x'_{2q}\beta_{qt}+\mu_t(\underset{\sim}{x})'\gamma_{qt}+\underset{\sim}{\pi}_t(\underset{\sim}{y})'\overline{\alpha}_{qt}} g_{qt}(\varepsilon|\underset{\sim}{x})d\varepsilon\right]^{y_q}\left[1-\int_{-\infty}^{x_{1q}+x'_{2q}\beta_{qt}+\mu_t(\underset{\sim}{x})'\gamma_{qt}+\underset{\sim}{\pi}_t(\underset{\sim}{y})'\overline{\alpha}_{qt}} g_{qt}(\varepsilon|\underset{\sim}{x})d\varepsilon\right]^{1-y_q}$$

and the joint density of $(\underset{\sim}{y}, s, \underset{\sim}{x})$ is $\prod_{q=1}^{P} p^2_{qt}(y_q|\underset{\sim}{x})f^2_t(s|\underset{\sim}{x})m^2_t(\underset{\sim}{x})$. Denote

$$T_{q0}(\underset{\sim}{v}_0) = G_{q0}(\underset{\sim}{v}_0)[1-G_{q0}(\underset{\sim}{v}_0)].$$

The score for estimating $t=0$ in this model becomes

$$S_0 = \sum_{q=1}^{P}\left(\frac{y_q - G_{q0}(\underset{\sim}{v}_0)}{T_{q0}(\underset{\sim}{v}_0)}\right)\cdot\left[g^2_{q0}(\underset{\sim}{v}_0)\cdot\left(v'_q\dot{\theta}_q+\gamma'_{q0}\dot{\mu}(\underset{\sim}{x})+\overline{\alpha}'_{q0}\dot{\underset{\sim}{\pi}}(\underset{\sim}{x})\right)+\dot{\Delta}_q(\underset{\sim}{v}_0)\right]+2\frac{\dot{f}(s|\underset{\sim}{x})}{f_0(s|\underset{\sim}{x})}+2\frac{\dot{m}(\underset{\sim}{x})}{m_0(\underset{\sim}{x})}$$

Let $\underbrace{\iota_q}_{P\times 1} \equiv (\underbrace{0,0,\ldots,0}_{q-1\text{ zeros}},1,\underbrace{0,\ldots,0}_{P-q\text{ zeros}})'$ (the unit vector in $\mathbb{R}^P$ with zeros everywhere and 1 in the q^{th} position). Recall that $dim(v_q) = dim(\theta_{q0}) \equiv d_q$ and $dim(\theta_0) = \sum_{q=1}^{P} d_q \equiv d$. Define

$$\underbrace{\overline{v}_q}_{d\times 1} \equiv (\underbrace{0',0',\ldots,0'}_{\substack{\sum_{r=1}^{q-1} d_r\\ \text{zeros}}}, v'_q, \underbrace{0',\ldots,0'}_{\substack{\sum_{r=q+1}^{P} d_r\\ \text{zeros}}})'$$

Plugging the expression in (14) and grouping terms, the score simplifies to

$$\begin{aligned} S_0 = \sum_{q=1}^{P}\left(\frac{y_q - G_{q0}(\underset{\sim}{v}_0)}{T_{q0}(\underset{\sim}{v}_0)}\right)\cdot\Bigg[& g^2_{q0}(\underset{\sim}{v}_0)\cdot\left(\overline{v}'_q+\overline{\alpha}'_{q0}\underset{\sim}{D}_0(\underset{\sim}{v}_0)^{-1}\underset{\sim}{A}_0(\underset{\sim}{x})\right)\dot{\theta} \\ & + g^2_{q0}(\underset{\sim}{v}_0)\cdot\left(\underset{\sim}{\gamma}'_{q0}+\overline{\alpha}'_{q0}\underset{\sim}{D}_0(\underset{\sim}{v}_0)^{-1}\underset{\sim}{B}_0(\underset{\sim}{v}_0)\right)\dot{\mu}(\underset{\sim}{x}) \\ & + \left(g^2_{q0}(\underset{\sim}{v}_0)\cdot\overline{\alpha}'_{q0}\underset{\sim}{D}_0(\underset{\sim}{v}_0)^{-1}+\iota'_q\right)\dot{\underset{\sim}{\Delta}}(\underset{\sim}{v}_0)\Bigg]+2\frac{\dot{f}(s|\underset{\sim}{x})}{f_0(s|\underset{\sim}{x})}+2\frac{\dot{m}(\underset{\sim}{x})}{m_0(\underset{\sim}{x})} \end{aligned} \tag{15}$$

Let

$$
\begin{aligned}
\underbrace{\underset{\sim}{M}_{q\theta}(\underset{\sim}{x})}_{d\times 1} &= \left(\overline{v}_q + \underset{\sim}{A}_0(\underset{\sim}{x})'\left(\underset{\sim}{D}_0(\underset{\sim}{v}_0)^{-1}\right)'\overline{\alpha}_{q0}\right)\cdot \frac{g_{q0}^2(\underset{\sim}{v}_0)}{\sqrt{T_{q0}(\underset{\sim}{v}_0)}},\\
\underbrace{\underset{\sim}{M}_{q\mu}(\underset{\sim}{v}_0)}_{d_s\times 1} &= \left(\underset{\sim}{\gamma}_{q0} + \underset{\sim}{B}_0(\underset{\sim}{v}_0)'\left(\underset{\sim}{D}_0(\underset{\sim}{v}_0)^{-1}\right)'\overline{\alpha}_{q0}\right)\cdot \frac{g_{q0}^2(\underset{\sim}{v}_0)}{\sqrt{T_{q0}(\underset{\sim}{v}_0)}},\\
\underbrace{\underset{\sim}{M}_{q\Delta}(\underset{\sim}{v}_0)}_{P\times 1} &= \left(g_{q0}^2(\underset{\sim}{v}_0)\left(\underset{\sim}{D}_0(\underset{\sim}{v}_0)^{-1}\right)'\overline{\alpha}_{q0} + \iota_q\right)\cdot \frac{1}{\sqrt{T_{q0}(\underset{\sim}{v}_0)}},\\
\underbrace{\underset{\sim}{M}_{\theta}(\underset{\sim}{x})}_{d\times P} &= \left(M_{1\theta}(\underset{\sim}{x}),\ldots,M_{P\theta}(\underset{\sim}{x})\right),\\
\underbrace{\underset{\sim}{M}_{\mu}(\underset{\sim}{v}_0)}_{d_s\times P} &= \left(M_{1\mu}(\underset{\sim}{v}_0),\ldots,M_{P\mu}(\underset{\sim}{v}_0)\right),\\
\underbrace{\underset{\sim}{M}_{\Delta}(\underset{\sim}{v}_0)}_{P\times P} &= \left(M_{1\Delta}(\underset{\sim}{v}_0),\ldots,M_{P\Delta}(\underset{\sim}{v}_0)\right)
\end{aligned}
\tag{16}
$$

Using iterated expectations[11],

$$
\begin{aligned}
&E\left[S_0^2\right]\\
&=\sum_{q=1}^{P} E\left[\left(M_{q\theta}(\underset{\sim}{x})'\dot{\theta} + M_{q\mu}(\underset{\sim}{v}_0)'\dot{\mu}(\underset{\sim}{x}) + M_{q\Delta}(\underset{\sim}{v}_0)'\dot{\underset{\sim}{\Delta}}(\underset{\sim}{v}_0)\right)^2\right] + 4E\left[\int_{\mathbb{r}^{d_s}} \dot{f}(s|\underset{\sim}{x})^2 ds\right] + 4\int_{\mathbb{R}_{d_x}} \dot{m}(x)^2 dx\\
&= E\left[\left(\dot{\theta}'\underset{\sim}{M}_{\theta}(\underset{\sim}{x}) + \dot{\mu}(\underset{\sim}{x})'\underset{\sim}{M}_{\mu}(\underset{\sim}{v}_0) + \dot{\underset{\sim}{\Delta}}(\underset{\sim}{v}_0)'\underset{\sim}{M}_{\Delta}(\underset{\sim}{v}_0)\right)^2\right] + 4E\left[\int_{\mathbb{r}^{d_s}} \dot{f}(s|\underset{\sim}{x})^2 ds\right] + 4\int_{\mathbb{R}_{d_x}} \dot{m}(x)^2 dx
\end{aligned}
$$

And from here, the Fisher-information inner product becomes

$$
\begin{aligned}
&\langle \dot{\tau}_1, \dot{\tau}_2 \rangle_F = \\
&E\left[\left(\dot{\theta}_1'\underset{\sim}{M}_{\theta}(\underset{\sim}{x}) + \dot{\mu}_1(\underset{\sim}{x})'\underset{\sim}{M}_{\mu}(\underset{\sim}{v}_0) + \dot{\underset{\sim}{\Delta}}_1(\underset{\sim}{v}_0)'\underset{\sim}{M}_{\Delta}(\underset{\sim}{v}_0)\right)\cdot\left(\underset{\sim}{M}_{\theta}(\underset{\sim}{x})'\dot{\theta}_2 + \underset{\sim}{M}_{\mu}(\underset{\sim}{v}_0)'\dot{\mu}_2(\underset{\sim}{x}) + \underset{\sim}{M}_{\Delta}(\underset{\sim}{v}_0)\dot{\underset{\sim}{\Delta}}_2(\underset{\sim}{v}_0)\right)'\right]\\
&+4E\left[\int_{\mathbb{R}^{d_s}} \dot{f}_1(s|\underset{\sim}{x})\dot{f}_2(s|\underset{\sim}{x})ds\right] + 4\int_{\mathbb{R}^{d_x}} \dot{m}_1(\underset{\sim}{x})\dot{m}_2(\underset{\sim}{x})dx
\end{aligned}
\tag{17}
$$

This is a direct generalization of the expression in (10) for the global interaction model. Once again, the bound is computed by finding the representer $\tau^* \in \dot{\mathscr{T}}$ that satisfies $\langle \tau^*, \dot{\tau}\rangle_F = c'\dot{\theta}\ \forall\ \dot{\tau} \in \dot{\mathscr{T}}$, where c is an arbitrary vector.

[11] Note that $M_{q\theta}(\underset{\sim}{x})'\dot{\theta} + M_{q\mu}(\underset{\sim}{v}_0)'\dot{\mu}(\underset{\sim}{x}) + M_{q\Delta}(\underset{\sim}{v}_0)'\dot{\underset{\sim}{\Delta}}(\underset{\sim}{v}_0)$ is a scalar.

4.2.3 Efficiency bound for the incomplete-information game

Proposition 2 *Once again, let $\underset{\sim}{I}_P$ denote the $P\times P$ identity matrix. Assume $\underset{\sim}{I}_P+\underset{\sim}{M}_\mu(\underset{\sim}{v}_0)'Var[s|\underset{\sim}{x}]\underset{\sim}{M}_\mu(\underset{\sim}{v}_0)$ is invertible w.p.1 and define*

$$\underbrace{\underset{\sim}{\Gamma}(\underset{\sim}{x})}_{P\times P} = \left[\underset{\sim}{I}_P + \underset{\sim}{M}_\mu(\underset{\sim}{v}_0)'Var[s|\underset{\sim}{x}]\underset{\sim}{M}_\mu(\underset{\sim}{v}_0)\right]^{-1},$$

$$\underbrace{\underset{\sim}{\Phi}^*(\underset{\sim}{x})}_{d\times P} = \left(\underset{\sim}{M}_\theta(\underset{\sim}{x}) - E\left[\underset{\sim}{M}_\theta(\underset{\sim}{x})\underset{\sim}{\Gamma}(\underset{\sim}{x})|\underset{\sim}{v}_0\right]\cdot E\left[\underset{\sim}{\Gamma}(\underset{\sim}{x})|\underset{\sim}{v}_0\right]^{-1}\right)\underset{\sim}{\Gamma}(\underset{\sim}{x})$$

and

$$\underset{\sim}{\Sigma}^*_\theta = E\left[\underset{\sim}{\Phi}^*(\underset{\sim}{x})\underset{\sim}{\Phi}^*(\underset{\sim}{x})'\right] + E\left[\underset{\sim}{\Phi}^*(\underset{\sim}{x})\underset{\sim}{M}_\mu(\underset{\sim}{v}_0)'Var[s|\underset{\sim}{x}]\underset{\sim}{M}_\mu(\underset{\sim}{v}_0)\underset{\sim}{\Phi}^*(\underset{\sim}{x})'\right].$$

If $\underset{\sim}{\Sigma}^*_\theta$ is invertible, the semiparametric efficiency bound for $\sqrt{n}-$consistent, regular estimators of θ_0 in Model (12) *under Assumptions 1B-2B is well-defined and is equal to $\underset{\sim}{\Sigma}^{*-1}_\theta$. And the efficient influence function is given by*

$$\underset{\sim}{\psi}(\underset{\sim}{y}, s, \underset{\sim}{x}, \theta_0) = \underset{\sim}{\Sigma}^{*-1}_\theta \underset{\sim}{\Phi}^*(\underset{\sim}{x})\cdot\left[diag(\underset{\sim}{T}_0(\underset{\sim}{v}_0)^{-1/2})\cdot\left(\underset{\sim}{y} - \underset{\sim}{G}_0(\underset{\sim}{v}_0)\right) - \underset{\sim}{M}_\mu(\underset{\sim}{v}_0)'\left(s - E[s|\underset{\sim}{x}]\right)\right],$$

where

$$\underbrace{diag(\underset{\sim}{T}_0(\underset{\sim}{v}_0)^{-1/2})}_{P\times P} = \begin{pmatrix} T_{10}(\underset{\sim}{v}_0)^{-1/2} & 0 & \cdots & 0 \\ 0 & T_{20}(\underset{\sim}{v}_0)^{-1/2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & T_{P0}(\underset{\sim}{v}_0)^{-1/2} \end{pmatrix},$$

$$\underbrace{\left(\underset{\sim}{y} - \underset{\sim}{G}_0(\underset{\sim}{v}_0)\right)}_{P\times 1} = \begin{pmatrix} y_1 - G_{10}(\underset{\sim}{v}_0) \\ y_2 - G_{20}(\underset{\sim}{v}_0) \\ \vdots \\ y_P - G_{P0}(\underset{\sim}{v}_0) \end{pmatrix}$$

Proof: In the appendix. ■

Invertibility of $\underset{\sim}{\Sigma}^*_\theta$ is a key condition for $\sqrt{n}-$consistency and a finite efficiency bound. A quick comparison shows that the efficiency bound for the incomplete-information game is a generalization of the one we found for the global interaction model in Proposition 1. This is not surprising because the latter is a special case of a game that each individual plays against the "representative agent". In both models examined here, rational expectations and the assumption of regularity of equilibrium beliefs play a crucial role in the computation of the bounds.

5 Constructing efficient estimators

To my knowledge, no efficient estimators have been proposed for the type of models studied here; particularly for the case where the distribution of unobserved disturbances is nonparametric (the case that we focused on). Existing methods, both for binary choice models with uncertainty (Ahn and Manski (1993) and Ahn (1995)) and for incomplete-information games (Aradillas-López (2012)) where the distribution of unobserved payoff shocks is unknown, rely on two-step procedures where the nonparametric regressors are estimated separately in a first step and then plugged in a second step into a criterion function which is then optimized to estimate the parameters of the model. By their two-step nature, these estimators can be inefficient in some econometric models.

While the purpose of this paper is to compute efficiency bounds in these models, we can potentially use our results to construct semiparametrically efficient estimators. For example, we can apply our results to existing methods based on the features of the efficient influence function. We will begin by describing these procedures and we will conclude with an outline of a possible likelihood-based approach where the parameter θ along with the nonparametric regressors (beliefs) are estimated simultaneously, subject to the rational-expectations and equilibrium-belief conditions.

5.1 Estimators based on the efficient influence function

Having an analytical expression of the efficient influence function and a $\sqrt{n}-$consistent estimator allows us, under proper conditions, to construct a semiparametrically efficient estimator. Newey (1990, Section 5) describes fundamental results on the construction of efficient estimators based on the efficient influence function. The results described there are based on Bickel (1982), Schick (1986) and Klaasen (1987). Their estimators are of the form

$$\widetilde{\theta} = \widehat{\theta} + \frac{1}{n}\sum_{i=1}^{n} \widehat{\psi}(y_i, s_i, x_i, \widehat{\theta}), \tag{18}$$

where $\widehat{\psi}(\cdot)$ is an estimator of the efficient influence function and $\widehat{\theta}$ is a first-step estimator such that $\sqrt{n}(\widehat{\theta} - \theta_0)$ is bounded in probability. The methods proposed by the aforementioned authors and described in Newey (1990, Section 5) rely on discretization of the parameter space and sample splitting as ways to ensure that $\widehat{\psi}(y_i, s_i, x_i, \theta)$ converges at a rate faster than $1/\sqrt{n}$. An inspection of the efficient influence functions found in this paper shows that these can be estimated by methods such as kernels or series. Properties of efficient influence functions estimated by kernels have been studied, e.g, in Bickel (1982) and Bickel, Klaasen, Ritov, and Wellner (1998), while series-based estimators have been studied, e.g, in Newey (2004). Whatever method is used, the results in this paper provide a way to construct semiparametrically efficient estimators in strategic-interaction models with rational expectations. To our knowledge, these appear to be the first such results for these types of models; in particular, for constructing efficient estimators of incomplete-information games.

5.2 Possible likelihood-based approaches

An efficient, semiparametric likelihood-based approach would estimate the parameter θ along with the nonparametric regressors (equilibrium beliefs) *simultaneously.* One potentially viable path would be to extend the empirical likelihood method proposed in Kitamura, Tripathi, and Ahn (2004) to our models. Their method is designed to construct efficient estimators for finite-dimensional parameters subject to *conditional* moment restrictions. In our case, these restrictions would correspond to the equilibrium conditions of beliefs. Our models would involve, in addition to a finite-dimensional parameter, the equilibrium beliefs as nonparametric regressors. The procedure could be designed to estimate, simultaneously, the parameters along with the equilibrium beliefs, subject to the conditional moment restrictions implied by self-consistency of beliefs. A detailed analysis of such an estimator and the conditions under which it would be asymptotically efficient are beyond the scope of this paper and are left for future research.

6 Concluding remarks

Discrete choice models with conditional expectations as nonparametric regressors include examples such as incomplete-information games, social-interactions models as well as single-agent discrete choice models with uncertainty. Assuming rational expectations and equilibrium beliefs implies that these regressors are functionals of the unknown parameters of the model. Therefore, the semiparametric efficiency bound for $\sqrt{n}-$consistent, regular estimators of the finite-dimensional parameters in these models is affected by the properties of the nonparametric regressors. Using the method of *representers* in the tangent space proposed by Severini and Tripathi (2001) we computed, for the first time, efficiency bounds for binary-choice examples of these models under the assumptions of rational expectations and regularity of equilibrium beliefs. Even though the goal of the paper is the derivation of the efficiency bounds, our results can potentially be applied to construct efficient estimators. We discussed how they can be combined with existing methods based on the features of the efficient influence function and we outlined a possible semi-empirical likelihood based approach. Finally, the steps and arguments used in the derivation of our bounds should hopefully provide a roadmap to help the reader extend the results beyond binary-choice models.

A Appendix

A.1 Proof of Proposition 1

We are interested in computing the efficiency bound for $\rho(\tau_0) = c'\theta_0$ for an arbitrary c. Our goal is to find the representers $\tau^* \equiv (\theta^*, g^*, f^*, m^*) \in \dot{\mathscr{T}}$ that satisfy the R-F condition $\nabla\rho(\dot{\tau}) = \langle \tau^*, \dot{\tau}\rangle_F \ \forall\ \dot{\tau} \in \dot{\mathscr{T}}$, with $\nabla\rho(\dot{\tau}) = c'\dot{\theta}$ in this case. Firstly, we will set $m^* = 0$, as this representer will be ancillary to our problem. Secondly, it will be convenient to look for a solution to the R-F condition where we can express $f^*(s|x) = \theta^{*\prime} \underbrace{t^*(s|x)}_{d\times 1}$ and $g^*(\varepsilon|v_0) = \theta^{*\prime} \underbrace{\lambda^*(\varepsilon|x)}_{d\times 1}$, where $t^* \in \overline{lin\ T(\mathcal{F}, f_0)}$ and $\lambda^* \in \overline{lin\ T(\mathcal{G}, g_0)}$ element-wise (note that any linear combination of elements in these tangent spaces are, by definition, elements of the tangent spaces, too). By the R-F Theorem, if such a solution exists, it is the unique solution to the R-F condition. As our proof will show, this is the case. From the above expressions, we have $\mu^*(x) = \delta^*(x)\theta^*$, where $\underbrace{\delta^*(x)}_{d_s\times d} = 2\int_{\mathbb{R}^{d_s}} s\cdot t^*(s|x)' f_0(s|x)ds$ and $\Delta^*(v_0) = \theta^{*\prime}\eta^*(v_0)$, where $\underbrace{\eta^*(v_0)}_{d\times 1} = 2\int_{-\infty}^{u(v_0)} \lambda^*(\varepsilon|v_0)g_0(\varepsilon|v_0)d\varepsilon$. Using the objects defined in (9), the expression of the Fisher-information inner product in (10) becomes

$$
\begin{aligned}
&\langle \tau^*, \dot{\tau}\rangle_F = \\
&\theta^{*\prime} E\left[\left(M_\theta(x) + \delta^*(x)' M_\mu(v_0) + \eta^*(v_0) M_\Delta(v_0)\right)\cdot\left(M_\theta(x)'\dot{\theta} + M_\mu(v_0)'\dot{\mu}(x) + M_\Delta(v_0)\cdot\dot{\Delta}(v_0)\right)\right] \\
&+\theta^{*\prime} E\left[4\int_{\mathbb{R}^{d_s}} t^*(s|x) f_0(s|x)ds\right]
\end{aligned}
$$

Let

$$\Phi^*(x) \equiv M_\theta(x) + \delta^*(x)' M_\mu(v_0) + \eta^*(v_0) M_\Delta(v_0).$$

Grouping terms, we obtain

$$
\begin{aligned}
\langle \tau^*, \dot{\tau}\rangle_F = &\underbrace{\theta^{*\prime} E\left[\Phi^*(x) M_\theta(x)'\right]\dot{\theta}}_{(A1A)} + \underbrace{\theta^{*\prime} E\left[\Phi^*(x) M_\Delta(v_0)\dot{\Delta}(v_0)\right]}_{(A1B)} \\
&+ \underbrace{\theta^{*\prime} E\left[\int_{\mathbb{R}^{d_s}} \left\{2\Phi^*(x) M_\mu(v_0)' s f_0(s|x) + 4t^*(s|x)\right\}\dot{f}(s|x)ds\right]}_{(A1C)}
\end{aligned}
\tag{A1}
$$

to find the representer τ^* that satisfies the R-F condition, we will first find the one that makes both $(A1B)$ and $(A1C)$ equal to zero. To make $(A1C)$ equal to zero, we choose the representer

$$t^*(s|x) = -\frac{1}{2}\Phi^*(x) M_\mu(v_0)'\left(s - E[s|x]\right) f_0(s|x). \tag{A2}$$

Note first that with this choice, $(A1C)$ becomes

$$2\theta^{*\prime}E\left[\Phi^*(x)M_\mu(v_0)'E[s|x]\underbrace{\int_{\mathbb{R}^{d_s}}\dot{f}(s|x)f_0(s|x)ds}_{=0\ \forall\ \dot{f}\in\overline{lin\ T(\mathcal{F},f_0)}}\right]=0.$$

Next we need to verify that $t^*\in\overline{lin\ T(\mathcal{F},f_0)}$ element-wise. From (5), we need to show that $\int_{\mathbb{R}^{d_s}}t^*(s|x)f_0(s|x)ds=0$ w.p.1. For any x, we have

$$\int_{\mathbb{R}^{d_s}}t^*(s|x)f_0(s|x)ds=-\frac{1}{2}\Phi^*(x)M_\mu(v_0)'\underbrace{\int_{\mathbb{R}^{d_s}}(s-E[s|x])\,f_0^2(s|x)ds}_{=0}=0,$$

and therefore $t^*\in\overline{lin\ T(\mathcal{F},f_0)}$ element-wise. Recall that we defined $\delta^*(x)=2\int_{\mathbb{R}^{d_s}}st^*(s|x)'f_0(s|x)ds$. With our choice of representer t^*, we obtain

$$\begin{aligned}\delta^*(x)&=-\int_{\mathbb{R}^{d_s}}s\cdot(s-E[s|x])'\,f_0^2(s|x)dsM_\mu(v_0)\Phi^*(x)'\\&=-\int_{\mathbb{R}^{d_s}}(s-E[s|x])\cdot(s-E[s|x])'\,f_0^2(s|x)dsM_\mu(v_0)\Phi^*(x)'\\&=-Var[s|x]M_\mu(v_0)\Phi^*(x)'\end{aligned}$$

Next recall that we defined $\Phi^*(x)=M_\theta(x)+\delta^*(x)'M_\mu(v_0)+\eta^*(v_0)M_\Delta(v_0)$. Denoting $\Gamma(x)=\left(1+M_\mu(v_0)'Var[s|x]M_\mu(v_0)\right)^{-1}$ and using the above expression for $\delta^*(x)$, we obtain

$$\Phi^*(x)=(M_\theta(x)+\eta^*(v_0)M_\Delta(v_0))\cdot\Gamma(x). \tag{A3}$$

Our next step is to make (A1B) equal to zero. That is, we want $\theta^{*\prime}E\left[\Phi^*(x)M_\Delta(v_0)\dot{\Delta}(v_0)\right]=0$. We will accomplish this by forcing $E\left[\Phi^*(x)|v_0\right]=0$ w.p.1. From the above expression for $\Phi^*(x)$, this will be done if we make $\eta^*(v_0)=-E\left[M_\theta(x)\Gamma(x)|v_0\right]E[\Gamma(x)|v_0]^{-1}M_\Delta(v_0)^{-1}$. Since $\eta^*(v_0)=2\int_{-\infty}^{u(v_0)}\lambda^*(\varepsilon|v_0)g_0(\varepsilon|v_0)d\varepsilon$, we need to find a representer $\lambda^*\in\overline{lin\ T(\mathcal{G},g_0)}$ that satisfies

$$\int_{-\infty}^{u(v_0)}\lambda^*(\varepsilon|v_0)g_0(\varepsilon|v_0)d\varepsilon=-\frac{1}{2}E\left[M_\theta(x)\Gamma(x)|v_0\right]E[\Gamma(x)|v_0]^{-1}M_\Delta(v_0)^{-1}. \tag{A4}$$

From (5), $\lambda^* \in \overline{lin\ T(\mathcal{G}, g_0)}$ (element-wise) requires the conditions $\int_{\mathbb{R}} \lambda^*(\varepsilon|v_0)g_0(\varepsilon|v_0)d\varepsilon = 0$ and $\int_{-\infty}^{c_\kappa} \lambda^*(\varepsilon|v_0)g_0(\varepsilon|v_0)d\varepsilon = 0$ w.p.1. Our choice for λ^* is

$$\lambda^*(\varepsilon|v_0) = -E\left[M_\theta(x)\Gamma(x)|v_0\right]E[\Gamma(x)|v_0]^{-1}M_\Delta(v_0)^{-1} \times \left[\mathbb{1}\left\{u(v_0) \leq c_\kappa\right\} \cdot \left(\frac{\mathbb{1}\{\varepsilon \leq u(v_0)\}}{\int_{-\infty}^{u(v_0)} g_0^2(\varepsilon|v_0)d\varepsilon}\right.\right.$$
$$\left.\left. - \frac{\mathbb{1}\{u(v_0) < \varepsilon \leq c_\kappa\}}{\int_{u(v_0)}^{c_\kappa} g_0^2(\varepsilon|v_0)d\varepsilon}\right) + \mathbb{1}\left\{u(v_0) > c_\kappa\right\} \cdot \left(\frac{\mathbb{1}\{c_\kappa < \varepsilon \leq u(v_0)\}}{\int_{c_\kappa}^{u(v_0)} g_0^2(\varepsilon|v_0)d\varepsilon} - \frac{\mathbb{1}\{\varepsilon > u(v_0)\}}{\int_{u(v_0)}^{\infty} g_0^2(\varepsilon|v_0)d\varepsilon}\right)\right] \cdot g_0(\varepsilon|v_0) \tag{A5}$$

It is easy to verify that the representer λ^* satisfies (A4), as well as $\int_{\mathbb{R}} \lambda^*(\varepsilon|v_0)g_0(\varepsilon|v_0)d\varepsilon = 0$ and $\int_{-\infty}^{c_\kappa} \lambda^*(\varepsilon|v_0)g_0(\varepsilon|v_0)d\varepsilon = 0$ w.p.1, and therefore $\lambda^* \in \overline{lin\ T(\mathcal{G}, g_0)}$. Thus, we now have

$$\Phi^*(x) = \Big(M_\theta(x) - E[M_\theta(x)\Gamma(x)|v_0] \cdot E\left[\Gamma(x)|v_0\right]^{-1}\Big)\,\Gamma(x). \tag{A6}$$

And our representers in (A2) and (A5) make both $(A1B)$ and $(A1C)$ equal to zero, and therefore (A1) becomes $\langle\tau^*, \dot{\tau}\rangle_F = \theta^{*\prime}E\left[\Phi^*(x)M_\theta(x)'\right]\dot{\theta} \quad \forall\, \dot{\theta} \in \dot{\mathscr{T}}$. From (A6) we have $E[\Phi^*(x)|v_0] = 0$. Therefore,

$$\begin{aligned}
E\left[\Phi^*(x)M_\theta(x)'\right] &= E\Big[\Phi^*(x)\underbrace{\Big(M_\theta(x) - E[M_\theta(x)\Gamma(x)|v_0] \cdot E\left[\Gamma(x)|v_0\right]^{-1}\Big)'}_{=(\Phi^*(x)\cdot\Gamma(x)^{-1})'}\Big] \\
&= E\left[\Phi^*(x) \cdot \left(\Phi^*(x) \cdot \Gamma(x)^{-1}\right)'\right] \\
&= E\left[\Phi^*(x)\left(\Gamma(x)^{-1}\right)\Phi^*(x)'\right] = E\Big[\Phi^*(x)\underbrace{\left(1 + M_\mu(v_0)'Var[s|x]M_\mu(v_0)\right)}_{=\Gamma(x)^{-1}}\Phi^*(x)'\Big] \\
&= E\left[\Phi^*(x)\Phi^*(x)'\right] + E\left[\Phi^*(x)M_\mu(v_0)'Var[s|x]M_\mu(v_0)\Phi^*(x)'\right] \equiv \Sigma_\theta^*
\end{aligned}$$

Therefore, (A1) becomes $\langle\tau^*, \dot{\tau}\rangle_F = \theta^{*\prime}\Sigma_\theta^*\dot{\theta} \quad \forall\, \dot{\theta} \in \dot{\mathscr{T}}$, and the R-F condition becomes $c'\dot{\theta} = \theta^{*\prime}\Sigma_\theta^*\dot{\theta}$ $\forall\, \dot{\theta} \in \dot{\mathscr{T}}$. Assuming that Σ_θ^* is invertible, the representer that solves this condition is $\theta^* = \Sigma_\theta^{*-1} \cdot c$. Our representers then become

$$\tau^* = (\theta^*, g^*, f^*, m^*) = \left(\theta^*, \theta^{*\prime}\lambda^*, \theta^{*\prime}t^*, 0\right), \quad \text{with} \quad \theta^* = \Sigma_\theta^{*-1} \cdot c,$$

and t^*, λ^* as described in Equations (A2) and (A5). From here, invoking the R-F Theorem as described in Section 2 we obtain the efficiency bound for $\sqrt{n}-$consistent, regular estimators of $\rho(\theta_0) = c'\theta_0$,

$$l.b = \langle\tau^*, \tau^*\rangle_F = \theta^{*\prime}\Sigma_\theta^*\theta^* = c'\Sigma_\theta^{*-1}\Sigma_\theta^*\Sigma_\theta^{*-1} \cdot c = c'\Sigma_\theta^{*-1}c.$$

Since c is arbitrary, the lower bound for the asymptotic variance of $\sqrt{n}-$consistent, regular estimators of θ_0 is given by Σ_θ^{*-1}, which concludes the proof of Proposition 1. ■

A.2 Proof of Proposition 2

The steps of the proof completely mirror those of Proposition 1. As we did there, the first step is to set $m^* = 0$ and to look for a solution to the R-F condition where we can express the remaining representers as $f^*(s|\underset{\sim}{x}) = \theta^{*\prime} \underbrace{t^*(s|\underset{\sim}{x})}_{d\times 1}$ and $g_q^*(\varepsilon|\underset{\sim}{v}_0) = \theta^{*\prime} \underbrace{\lambda_q^*(\varepsilon|\underset{\sim}{v}_0)}_{d\times 1}$, where $t^* \in \overline{lin\ T(\mathcal{F}, f_0)}$ and $\lambda_q^* \in \overline{lin\ T(\mathcal{G}, g_0)}$ element-wise. Accordingly, if we define

$$\underbrace{\eta_q^*(\underset{\sim}{v}_0)}_{d\times 1} = 2\int_{-\infty}^{u_{q0}(\underset{\sim}{v}_0)} \lambda_q^*(\varepsilon|\underset{\sim}{v}_0) g_{q0}(\varepsilon|\underset{\sim}{v}_0) d\varepsilon \quad \text{and let} \quad \underbrace{\underset{\sim}{\eta}^*(\underset{\sim}{v}_0)}_{d\times P} \equiv (\eta_1^*(\underset{\sim}{v}_0), \ldots, \eta_P^*(\underset{\sim}{v}_0)), \tag{A7}$$

we can express $\underbrace{\underset{\sim}{\Delta}^*(\underset{\sim}{v}_0)}_{P\times 1} = \underset{\sim}{\eta}^*(\underset{\sim}{v}_0)'\theta^*$. Similarly, defining $\underbrace{\delta^*(\underset{\sim}{x})}_{d_s\times d} = 2\int_{\mathbb{R}^{d_s}} st^*(s|\underset{\sim}{x})' f_0(s|\underset{\sim}{x}) ds$, we have $\mu^*(\underset{\sim}{x}) = \delta^*(\underset{\sim}{x})\theta^*$. Finally, letting

$$\underbrace{\underset{\sim}{\Phi}^*(\underset{\sim}{x})}_{d\times P} = \underset{\sim}{M}_\theta(\underset{\sim}{x}) + \delta^*(\underset{\sim}{x})' \underset{\sim}{M}_\mu(\underset{\sim}{v}_0) + \underset{\sim}{\eta}^*(\underset{\sim}{v}_0) \underset{\sim}{M}_\Delta(\underset{\sim}{v}_0),$$

then using the Fisher-information inner product expression in (17), we have

$$\begin{aligned} \langle \tau^*, \dot{\tau} \rangle_F = {} & \underbrace{\theta^{*\prime} E\left[\underset{\sim}{\Phi}^*(\underset{\sim}{x}) \underset{\sim}{M}_\theta(\underset{\sim}{x})'\right] \dot{\theta}}_{(A8A)} + \underbrace{\theta^{*\prime} E\left[\underset{\sim}{\Phi}^*(\underset{\sim}{x}) \underset{\sim}{M}_\Delta(\underset{\sim}{v}_0) \dot{\underset{\sim}{\Delta}}(\underset{\sim}{v}_0)\right]}_{(A8B)} \\ & + \underbrace{\theta^{*\prime} E\left[\int_{\mathbb{R}^{d_s}} \left\{2\underset{\sim}{\Phi}^*(\underset{\sim}{x}) \underset{\sim}{M}_\mu(\underset{\sim}{v}_0)' s f_0(s|\underset{\sim}{x}) + 4t^*(s|\underset{\sim}{x})\right\} \dot{f}(s|\underset{\sim}{x}) ds\right]}_{(A8C)} \end{aligned} \tag{A8}$$

The expression in (A8) is a direct generalization of (A1) in the proof of Proposition 1. Like we did there, our goal is to find representers that make (A8B) and (A8C) equal to zero. Analogous steps to those in the proof of Proposition 1 yield the representer

$$t^*(s|\underset{\sim}{x}) = -\frac{1}{2} \underset{\sim}{\Phi}^*(\underset{\sim}{x}) \underset{\sim}{M}_\mu(\underset{\sim}{v}_0)' \left(s - E[s|\underset{\sim}{x}]\right) f_0(s|\underset{\sim}{x}) \tag{A9}$$

which is analogous to the representer in (A2), in the proof of Proposition 1. Note that $\int_{\mathbb{R}^{d_s}} t^*(s|\underset{\sim}{x}) f_0(s|\underset{\sim}{x}) ds = 0$ w.p.1, and therefore $t^* \in \overline{lin\ T(\mathcal{F}, f_0)}$ element-wise. Next let

$$\underset{\sim}{\Gamma}(\underset{\sim}{x}) = \left(\underset{\sim}{I}_P + \underset{\sim}{M}_\mu(\underset{\sim}{v}_0)' Var[s|\underset{\sim}{x}] \underset{\sim}{M}_\mu(\underset{\sim}{v}_0)\right)^{-1}.$$

From (A9), analogous steps to those leading to (A3) in the proof of Proposition 1 yield

$$\underset{\sim}{\Phi}^*(\underset{\sim}{x}) = \left(\underset{\sim}{M}_\theta(\underset{\sim}{x}) + \underset{\sim}{\eta}^*(\underset{\sim}{v}_0) \underset{\sim}{M}_\Delta(\underset{\sim}{v}_0)\right) \underset{\sim}{\Gamma}(\underset{\sim}{x}). \tag{A10}$$

This is directly analogous to the expression in (A3), in the proof of Proposition 1. As we did there, our next step is to choose $\underset{\sim}{\eta}^*$ to ensure that $E[\underset{\sim}{\Phi}^*(\underset{\sim}{x})|\underset{\sim}{v}_0] = 0$ w.p.1. This requires setting

$$\underbrace{\underset{\sim}{\eta}^*(\underset{\sim}{v}_0)}_{d\times P} = -E\left[\underset{\sim}{M}_\theta(\underset{\sim}{x})\underset{\sim}{\Gamma}(\underset{\sim}{x})|\underset{\sim}{v}_0\right] E\left[\underset{\sim}{\Gamma}(\underset{\sim}{x})|\underset{\sim}{v}_0\right]^{-1} \underset{\sim}{M}_\Delta(\underset{\sim}{v}_0)^{-1} \tag{A11}$$

Let $\underset{\sim}{\eta}^*_{[q]}(\underset{\sim}{v}_0)$ denote the q^{th} column of $\underset{\sim}{\eta}^*(\underset{\sim}{v}_0)$, as defined above and recall that we defined $\underset{\sim}{\eta}^*_{[q]}(\underset{\sim}{v}_0) = 2\int_{-\infty}^{u_{q0}(\underset{\sim}{v}_0)} \lambda_q^*(\varepsilon|\underset{\sim}{v}_0)g_{q0}(\varepsilon|\underset{\sim}{v}_0)d\varepsilon$ (see (A7)). Parallel to the proof of Proposition 1, the representers λ_q^* are chosen from here. Our choice for λ_q^* is

$$\begin{aligned}\lambda_q^*(\varepsilon|\underset{\sim}{v}_0) = \frac{1}{2}\cdot\underset{\sim}{\eta}^*_{[q]}(\underset{\sim}{v}_0)\times\Bigg[\mathbb{1}\left\{u_{q0}(\underset{\sim}{v}_0)\le c_\kappa\right\}\cdot\Bigg(\frac{\mathbb{1}\{\varepsilon\le u_{q0}(\underset{\sim}{v}_0)\}}{\int_{-\infty}^{u_{q0}(\underset{\sim}{v}_0)} g_{q0}^2(\varepsilon|\underset{\sim}{v}_0)d\varepsilon} \\ -\frac{\mathbb{1}\{u_{q0}(\underset{\sim}{v}_0)<\varepsilon\le c_\kappa\}}{\int_{u_{q0}(\underset{\sim}{v}_0)}^{c_\kappa} g_{q0}^2(\varepsilon|\underset{\sim}{v}_0)d\varepsilon}\Bigg)+\mathbb{1}\left\{u_{q0}(\underset{\sim}{v}_0)>c_\kappa\right\}\cdot\Bigg(\frac{\mathbb{1}\{c_\kappa<\varepsilon\le u_{q0}(\underset{\sim}{v}_0)\}}{\int_{c_\kappa}^{u_{q0}(\underset{\sim}{v}_0)} g_{q0}^2(\varepsilon|\underset{\sim}{v}_0)d\varepsilon}-\frac{\mathbb{1}\{\varepsilon>u_{q0}(\underset{\sim}{v}_0)\}}{\int_{u_{q0}(\underset{\sim}{v}_0)}^{\infty} g_{q0}^2(\varepsilon|\underset{\sim}{v}_0)d\varepsilon}\Bigg)\Bigg]\cdot g_{q0}(\varepsilon|\underset{\sim}{v}_0)\end{aligned} \tag{A12}$$

This ensures $2\int_{-\infty}^{u_{q0}(\underset{\sim}{v}_0)} \lambda_q^*(\varepsilon|\underset{\sim}{v}_0)g_{q0}(\varepsilon|\underset{\sim}{v}_0)d\varepsilon = \underset{\sim}{\eta}^*_{[q]}(\underset{\sim}{v}_0)$ w.p.1. In addition, $\int_{\mathbb{R}} \lambda_q^*(\varepsilon|\underset{\sim}{v}_0)g_{q0}(\varepsilon|\underset{\sim}{v}_0)d\varepsilon = 0$ and $\int_{-\infty}^{c_\kappa} \lambda_q^*(\varepsilon|\underset{\sim}{v}_0)g_{q0}(\varepsilon|\underset{\sim}{v}_0)d\varepsilon = 0$ w.p.1, and therefore $\lambda_q^* \in \overline{lin\ T(\mathcal{G}, g_0)}$. And we have

$$\underset{\sim}{\Phi}^*(\underset{\sim}{x}) = \left(\underset{\sim}{M}_\theta(\underset{\sim}{x}) - E[\underset{\sim}{M}_\theta(\underset{\sim}{x})\underset{\sim}{\Gamma}(\underset{\sim}{x})|\underset{\sim}{v}_0]\cdot E\left[\underset{\sim}{\Gamma}(x)|\underset{\sim}{v}_0\right]^{-1}\right)\underset{\sim}{\Gamma}(\underset{\sim}{x}). \tag{A13}$$

And (A8) simplifies to $\langle\tau^*, \dot{\tau}\rangle_F = \theta^{*\prime}E\left[\underset{\sim}{\Phi}^*(\underset{\sim}{x})\underset{\sim}{M}_\theta(\underset{\sim}{x})'\right]\dot{\theta}\ \ \forall\,\dot{\theta}\in\dot{\mathscr{T}}$. The rest of the steps are entirely analogous to those of the proof of Proposition 1. From (A13), we have $E\left[\underset{\sim}{\Phi}^*(\underset{\sim}{x})|\underset{\sim}{v}_0\right] = 0$ w.p.1. Therefore,

$$\begin{aligned}E\left[\underset{\sim}{\Phi}^*(\underset{\sim}{x})\underset{\sim}{M}_\theta(\underset{\sim}{x})'\right] &= E\Big[\underset{\sim}{\Phi}^*(\underset{\sim}{x})\underbrace{\left(\underset{\sim}{M}_\theta(\underset{\sim}{x}) - E\left[\underset{\sim}{M}_\theta(\underset{\sim}{x})\underset{\sim}{\Gamma}(\underset{\sim}{x})|\underset{\sim}{v}_0\right]E\left[\underset{\sim}{\Gamma}(\underset{\sim}{x})|\underset{\sim}{v}_0\right]^{-1}\right)'}_{=(\underset{\sim}{\Phi}^*(\underset{\sim}{x})\cdot\underset{\sim}{\Gamma}(\underset{\sim}{x})^{-1})'}\Big] \\ &= E\left[\underset{\sim}{\Phi}^*(\underset{\sim}{x})(\underset{\sim}{\Phi}^*(\underset{\sim}{x})\cdot\underset{\sim}{\Gamma}(\underset{\sim}{x})^{-1})'\right] \\ &= E\left[\underset{\sim}{\Phi}^*(\underset{\sim}{x})\left(\underset{\sim}{\Gamma}(\underset{\sim}{x})^{-1}\right)'\underset{\sim}{\Phi}^*(\underset{\sim}{x})'\right] = E\Big[\underset{\sim}{\Phi}^*(\underset{\sim}{x})\underbrace{\left(\underset{\sim}{I}_P + \underset{\sim}{M}_\mu(\underset{\sim}{v}_0)'Var[s|\underset{\sim}{x}]\underset{\sim}{M}_\mu(\underset{\sim}{v}_0)\right)}_{=(\underset{\sim}{\Gamma}(\underset{\sim}{x})^{-1})'}\underset{\sim}{\Phi}^*(\underset{\sim}{x})'\Big] \\ &= E\left[\underset{\sim}{\Phi}^*(\underset{\sim}{x})\underset{\sim}{\Phi}^*(\underset{\sim}{x})'\right] + E\left[\underset{\sim}{\Phi}^*(\underset{\sim}{x})\underset{\sim}{M}_\mu(\underset{\sim}{v}_0)'Var[s|\underset{\sim}{x}]\underset{\sim}{M}_\mu(\underset{\sim}{v}_0)\underset{\sim}{\Phi}^*(\underset{\sim}{x})'\right] \equiv \underset{\sim}{\Sigma}_\theta^*\end{aligned}$$

And therefore (A8) simplifies to $\langle\tau^*, \dot{\tau}\rangle_F = \theta^{*\prime}\underset{\sim}{\Sigma}_\theta^*\dot{\theta}\ \ \forall\,\dot{\theta}\in\dot{\mathscr{T}}$. The R-F condition becomes $c'\dot{\theta} = \theta^{*\prime}\underset{\sim}{\Sigma}_\theta^*\dot{\theta}\ \ \forall\,\dot{\theta}\in\dot{\mathscr{T}}$. Assuming that $\underset{\sim}{\Sigma}_\theta^*$ is invertible, the representer that satisfies the R-F condition is therefore $\theta^* = \underset{\sim}{\Sigma}_\theta^{*-1}\cdot c$. From here, the R-F Theorem described in Section 2 yields the efficiency bound for $\sqrt{n}-$consistent, regular estimators of $\rho(\theta_0) = c'\theta_0$,

$$l.b = \|\tau^*\|_F^2 = \langle\tau^*, \tau^*\rangle_F = \theta^{*\prime}\underset{\sim}{\Sigma}_\theta^*\theta^* = c'\underset{\sim}{\Sigma}_\theta^{*-1}\underset{\sim}{\Sigma}_\theta^*\underset{\sim}{\Sigma}_\theta^{*-1}\cdot c = c'\underset{\sim}{\Sigma}_\theta^{*-1}c.$$

Since c is arbitrary, the lower bound for the asymptotic variance of $\sqrt{n}-$consistent, regular estimators of θ_0 is given by $\underset{\sim}{\Sigma}_{\theta}^{*-1}$, which concludes the proof of Proposition 2. ■

References

Ahn, H. (1995). Nonparametric two-stage estimation of conditional choice probabilities in a binary choice model under uncertainty. *Journal of Econometrics 67*, 337–378.

Ahn, H., H. Ichimura, J. L. Powell, and P. A. Ruud (2018). Simple estimators for invertible index models. *Journal of Business & Economic Statistics 36*(1), 1–10.

Ahn, H. and C. Manski (1993). Distribution theory for the analysis of binary choice under uncertainty with nonparametric estimation of expectations. *Journal of Econometrics 56*, 291–321.

Aradillas-López, A. (2012). Pairwise difference estimation of incomplete information games. *Journal of Econometrics 168*, 120–140.

Aradillas-López, A. (2019). Computing semiparametric efficiency bounds in linear models with nonparametric regressors. *Economics Letters 185*, 1–5.

Bickel, P. (1982). On adaptive estimation. *Annals of Statistics 10*, 647–671.

Bickel, P., C. Klaasen, Y. Ritov, and J. Wellner (1998). *Efficient and Adaptive Estimation for Semiparametric Models*. Springer.

Brock, W. and S. Durlauf (2001). Interactions-based models. In J. Heckman and E. Leamer (Eds.), *The Handbook of Econometrics*, Volume 5, pp. 3297–3380. North-Holland.

Donkers, B. and M. Schafgans (2008). Specification and estimation of semiparametric multiple-index models. *Econometric Theory 24*, 1584–1606.

Ichimura, H. (1993). Semiparametric least squares (sls) and weighted sls estimation of single-index models. *Journal of Econometrics 58*, 71 – 120.

Ichimura, H. and L.-F. Lee (1991). Semiparametric least squares estimation of multiple index models: Single equation estimation. In W. Barnett, J. Powell, and G. Tauchen (Eds.), *Nonparametric and Semiparametric Estimation Methods in Econometrics and Statistics*, pp. 3–49. Cambridge University Press.

Kim, J. and D. Pollard (1990). Cube root asymptotics. *Annals of Statistics 18*, 191–219.

Kitamura, Y., G. Tripathi, and H. Ahn (2004). Empirical likelihood-based inference in conditional moment restriction models. *Econometrica 72*, 1667–1714.

Klaasen, C. (1987). Consistent estimation of the influence function of locally asymptotically linear estimators. *Annals of Statistics 15*, 1548–1562.

Klein, R. and R. Spady (1993). An efficient semiparametric estimator for binary response models. *Econometrica 61*(2), 387–421.

Lewbel, A. and X. Tang (2015). Identification and estimation of games with incomplete information using excluded regressors. *Journal of Econometrics 189*, 229–244.

Li, Q. and J. Wooldridge (2002). Semiparametric estimation of partially linear models for dependent data with generated regressors. *Econometric Theory 18*, 625–645.

Luenberger, D. (1969). *Optimization by vector space methods.* Wiley Inter-Science.

Manski, C. (1975). Maximum score estimation of the stochastic utility model of choice. *Journal of Econometrics 3*, 205 – 228.

Manski, C. F. (1993). Identification of endogenous social effects: the reflection problem. *Review of Economic Studies 60*, 531–542.

Manski, C. F. (1995). *Identification Problems in the Social Sciences.* Cambridge, MA: Harvard University Press.

Mas-Colell, A., M. Whinston, and J. Green (1995). *Microeconomic Theory.* Oxford University Press.

Newey, W. (1990). Semiparametric efficiency bounds. *Journal of Applied Econometrics 5*, 99–135.

Newey, W. (2004). Efficient semiparametric estimation via moment restrictions. *Econometrica 72*, 1877–1897.

Pesendorfer, M. and P. Schmidt-Dengler (2008). Asymptotic least squares estimators for dynamic games. *Review of Economic Studies 75*, 901–928.

Powell, J. L., J. Stock, and T. M. Stoker (1989). Semiparametric estimation of index coefficients. *Econometrica 57*(6), 1403–30.

Rilstone, P. (1993). Calculating the (local) semiparametric efficiency bounds for the generated regressors problem. *Journal of Econometrics 56*, 357–370.

Schick, A. (1986). On asymptotically efficient estimation in semiparametric models. *Annals of Statistics 14*, 1139–1151.

Severini, T. and G. Tripathi (2001). A simplified approach to computing efficiency bounds in semiparametric models. *Journal of Econometrics 102*, 23–66.

Stein, C. (1956). Efficient nonparametric testing and estimation. In M. Armstrong and P. Rob (Eds.), *Berkeley symposium on mathematical statistics and probability*, Volume 1, pp. 187–195. University of California Press.

Young, N. (1988). *An introduction to Hilbert space.* Cambridge.