Supplement for "A nonparametric test for cooperation in discrete games"

Andres Aradillas-Lopez* Lidia Kosenkova†

Abstract

This document begins with an analysis of the 2×2 game case, where we describe the implications of our assumptions, along with the analysis of the power of our test for cooperation (violations to our testable implications of cooperation) when the true underlying behavior is Nash equilibrium. We then proceed with the step-by-step derivations and details pertaining to the proof of our main econometric result, Theorem 2. Next, we include the details of our Monte Carlo experiments and we conclude with results of our empirical illustration using alternative bandwidth choices. Every section in this supplement has the format **SX.X** and every equation has the format **(S-X)**. Any section or equation that we reference here which does not have this format refers to a section or an equation in the main paper or its appendix.

S1 The case of a 2×2 game

Here we focus on the special case of a 2×2 binary choice game. Let us begin by illustrating conditions under which such a game would satisfy the assumptions in the paper. Suppose the normal-form game is given by,

	$Y_2 = 1$	$Y_2 = 0$
$Y_1 = 1$	$\Gamma^1_{11}(X_G, \nu)$, $\Gamma^2_{11}(X_G, \nu)$	$\Gamma^1_{10}(X_G, \nu)$, $\Gamma^2_{10}(X_G, \nu)$
$Y_1 = 0$	$\Gamma^1_{01}(X_G, \nu)$, $\Gamma^2_{01}(X_G, \nu)$	$\Gamma^1_{00}(X_G, \nu)$, $\Gamma^2_{00}(X_G, \nu)$

As we described in Section 2.3.1 of the paper, payoff functions with the following structure satisfy the exchangeability property in Assumption 2 under conditions we described there,

$$u_p(Y_p, Y_{-p}, X_G, \nu) = \phi^0_p(X_G, \nu) + \phi^1_p(X_G, \nu) \cdot \lambda(Y_{-p}) + \phi^2_p(X_G, \nu) \cdot Y_p + \Delta_p(Y_{-p}, X_G, \nu) \cdot Y_p. \quad \text{(S-1)}$$

*Department of Economics, Pennsylvania State University, University Park, PA 16802, United States. Email: aaradill@psu.edu

†Department of Economics, University of Virginia, Charlottesville, VA 22903, United States. Email: lk7cb@virginia.edu

Players' payoff functions in the 2×2 game described above can be represented as in (S-1), with $\lambda(Y_{-p}) \equiv Y_{-p}$, $\phi_p^0(X_G,\nu) \equiv \Gamma_{00}^p(X_G,\nu)$, $\phi_1^1(X_G,\nu) \equiv \Gamma_{01}^1(X_G,\nu) - \Gamma_{00}^1(X_G,\nu)$, $\phi_1^2(X_G,\nu) \equiv \Gamma_{10}^1(X_G,\nu) - \Gamma_{00}^1(X_G,\nu)$, $\phi_2^1(X_G,\nu) \equiv \Gamma_{10}^2(X_G,\nu) - \Gamma_{00}^2(X_G,\nu)$, $\phi_2^2(X_G,\nu) \equiv \Gamma_{01}^2(X_G,\nu) - \Gamma_{00}^2(X_G,\nu)$, and $\Delta_p(Y_{-p},X_G,\nu) \equiv \Big(\Gamma_{11}^p(X_G,\nu) + \Gamma_{00}^p(X_G,\nu) - \Gamma_{01}^p(X_G,\nu) - \Gamma_{10}^p(X_G,\nu)\Big)\cdot Y_{-p}$. From here, the condition in (3) and the exchangeability restriction in Assumption 2 will be satisfied if, $X_1 = X_2$, implies: (a) $\Gamma_{01}^1(X_G,\nu)$ and $\Gamma_{10}^2(X_G,\nu)$ are exchangeable, and (b) $\Gamma_{10}^1(X_G,\nu)$ and $\Gamma_{01}^2(X_G,\nu)$ are exchangeable. In this case, equation (5) reduces to,

$$Q(1,0|x) = Q(0,1|x), \quad \text{a.e } x \in \mathcal{X} \tag{S-2}$$

In other words, we must have $Pr(1,0|X_1 = X_2) = Pr(0,1|X_1 = X_2)$, a.s. A commonly assumed parameterization of payoffs for 2×2 binary choice games in the literature (Bresnahan and Reiss (1990), Bresnahan and Reiss (1991), Tamer (2003)) is,

	$Y_2 = 1$	$Y_2 = 0$
$Y_1 = 1$	$X_1'\beta_1 + \varepsilon_1 + \Delta_1$, $X_2'\beta_2 + \varepsilon_2 + \Delta_2$	$X_1'\beta_1 + \varepsilon_1$, 0
$Y_1 = 0$	0 , $X_2'\beta_2 + \varepsilon_2$	0 , 0

Define $\nu \equiv (\beta_1, \varepsilon_1, \Delta_1, \beta_2, \varepsilon_2, \Delta_2)$ (all allowed to be random). This is a special case of the 2×2 payoffs described above, with $\Gamma_{11}^p(X_G,\nu) \equiv X_p'\beta_p + \Delta_p + \varepsilon_p$, $\Gamma_{00}^p(X_G,\nu) \equiv 0$, $\Gamma_{10}^1(X_G,\nu) \equiv X_1'\beta_1 + \varepsilon_1$, $\Gamma_{10}^2(X_G,\nu) \equiv 0$, $\Gamma_{01}^1(X_G,\nu) \equiv 0$, and $\Gamma_{01}^2(X_G,\nu) \equiv X_2'\beta_2 + \varepsilon_2$. Let us abbreviate $t_1 \equiv X_1'\beta_1 + \varepsilon_1$ and $t_2 \equiv X_2'\beta_2 + \varepsilon_2$. As we described above, the exchangeability restriction in Assumption 2 will be satisfied if, $X_1 = X_2$ implies t_1 and t_2 are exchangeable.

S1.1 Power of our test in the 2×2 case when true behavior is noncooperative

Now let us investigate the power of our test when the true underlying behavior is noncooperative. In a binary choice game where $Y_p \in \{0,1\}$ with two players (i.e, a 2×2 game), equation (5) in the paper reduces to,

$$Q(1,0|x) = Q(0,1|x), \quad \text{a.e } x \in \mathcal{X} \tag{S-3}$$

Let us analyze conditions under which equation (S-3) can be violated in a 2×2 game with the parameterization described above,

	$Y_2 = 1$	$Y_2 = 0$
$Y_1 = 1$	$X_1'\beta_1 + \varepsilon_1 + \Delta_1$, $X_2'\beta_2 + \varepsilon_2 + \Delta_2$	$X_1'\beta_1 + \varepsilon_1$, 0
$Y_1 = 0$	0 , $X_2'\beta_2 + \varepsilon_2$	0 , 0

As described in equation (S-3), cooperation in a 2×2 binary choice game under Assumptions 1, 2 and 3 implies the restriction,

$$Pr(1,0|X_1 = X_2) = Pr(0,1|X_1 = X_2), \quad \text{a.s.}$$

We will focus on Nash equilibrium behavior, either in pure-strategies (PSNE) or mixed-strategies (MSNE), as an alternative to cooperation and we will illustrate how equation S-3 can be violated. Let $t_1 \equiv X_1'\beta_1 + \varepsilon_1$ and $t_2 \equiv X_2'\beta_2 + \varepsilon_2$. The exchangeability restriction in Assumption 2 will be satisfied if, $X_1 = X_2$ implies t_1 and t_2 are exchangeable.

S1.2 Violations of (S-3) when actions are strategic substitutes

The game will be one of strategic substitutes if $\Delta_1 \leq 0$ and $\Delta_2 \leq 0$. We have the following equilibrium regions involving the outcomes $(1,0)$ and $(0,1)$.

- $(1,0)$ will be the unique PSNE if $t_1 > -\Delta_1$ and $t_2 < 0$.
- $(0,1)$ will be the unique PSNE if $t_1 < 0$ and $t_2 > -\Delta_2$.
- If $\Delta_1 \neq 0$ and $\Delta_2 \neq 0$ and $0 \leq t_p \leq -\Delta_p$ for $p = 1,2$, there exist three NE. We have two PSNE: $(1,0)$ and $(0,1)$, and a MSNE where player p chooses $Y_p = 1$ with probability π_p, with $(\pi_1,\pi_2) = \left(-\frac{t_2}{\Delta_2}\,,\,-\frac{t_1}{\Delta_1}\right)$

Let $\mathscr{M}$ denote the underlying equilibrium selection mechanism (which is only relevant in the region where multiple NE exist). Let $\mathscr{M}_{1,0}$ and $\mathscr{M}_{0,1}$ denote the indicator functions for whether PSNE $(1,0)$ or $(0,1)$ are selected and let $\mathscr{M}_\pi$ denote the indicator for whether the MSNE is selected. Let

$$\mathscr{G}_s \equiv \mathbb{1}\Big\{\,0 \leq t_p \leq -\Delta_p \text{ and } \Delta_p \neq 0 \text{ for } p = 1,2\Big\}.$$

That is, the indicator for whether the game is in the multiple NE region. Then,

$$\begin{aligned}
&Pr(1,0|X_1 = X_2) = \\
&Pr\Big(t_1 > -\Delta_1\,,\, t_2 < 0 \,\Big|\, X_1 = X_2\Big) \\
&+Pr\Big(\mathscr{G}_s = 1 \,\Big|\, \mathscr{M}_{1,0} = 1,\, X_1 = X_2\Big)\cdot Pr\Big(\mathscr{M}_{1,0} = 1 \,\Big|\, X_1 = X_2\Big) \\
&+E\left[\left(-\frac{t_2}{\Delta_2}\right)\cdot\left(1+\frac{t_1}{\Delta_1}\right) \,\Bigg|\, \mathscr{G}_s = 1,\, \mathscr{M}_\pi = 1,\, X_1 = X_2\right]\cdot Pr\Big(\mathscr{G}_s = 1 \,\Big|\, \mathscr{M}_\pi = 1,\, X_1 = X_2\Big)\cdot Pr\Big(\mathscr{M}_\pi = 1 \,\Big|\, X_1 = X_2\Big),
\end{aligned}$$

and,

$$
\begin{aligned}
&Pr(0,1|X_1 = X_2) = \\
&Pr\Big(t_1 < 0\ ,\ t_2 > -\Delta_2 \ \Big|\ X_1 = X_2\Big) \\
+&Pr\Big(\mathscr{G}_s = 1 \ \Big|\ \mathscr{M}_{0,1} = 1,\, X_1 = X_2\Big) \cdot Pr\Big(\mathscr{M}_{0,1} = 1 \ \Big|\ X_1 = X_2\Big) \\
+&E\left[\left(1 + \frac{t_2}{\Delta_2}\right) \cdot \left(-\frac{t_1}{\Delta_1}\right) \ \Big|\ \mathscr{G}_s = 1,\, \mathscr{M}_\pi = 1,\, X_1 = X_2\right] \cdot Pr\Big(\mathscr{G}_s = 1 \ \Big|\ \mathscr{M}_\pi = 1,\, X_1 = X_2\Big) \cdot Pr\Big(\mathscr{M}_\pi = 1 \ \Big|\ X_1 = X_2\Big).
\end{aligned}
$$

We maintain that t_1 and t_2 are exchangeable conditional on $X_1 = X_2$. From the above expressions, can see that we can have $Pr(1,0|X_1 = X_2) \neq Pr(0,1|X_1 = X_2)$ if $Pr\left(\Delta_1 \neq \Delta_2 | X_1 = X_2\right) > 0$. That is, if the strategic effects can be different in magnitude between the two players with nonzero probability even if $X_1 = X_2$. Suppose instead that $Pr(\Delta_1 = \Delta_2 | X_1 = X_2) = 1$, and let $\Delta_1 = \Delta_2 \equiv \Delta$. Suppose $Pr\left(\Delta \neq 0 | X_1 = X_2\right) > 0$. Then,

$$
\begin{aligned}
Pr(1,0|X_1 = X_2) - Pr(0,1|X_1 = X_2) = &\ Pr\left(\mathscr{G}_s = 1 \mid \mathscr{M}_{1,0} = 1,\, X_1 = X_2\right) \cdot Pr\left(\mathscr{M}_{1,0} = 1 \mid X_1 = X_2\right) \\
&- Pr\left(\mathscr{G}_s = 1 \mid \mathscr{M}_{0,1} = 1,\, X_1 = X_2\right) \cdot Pr\left(\mathscr{M}_{0,1} = 1 \mid X_1 = X_2\right).
\end{aligned}
$$

In this case we can see that we can have $Pr(1,0|X_1 = X_2) \neq Pr(0,1|X_1 = X_2)$ if the selection mechanism $\mathscr{M}$ selects $(1,0)$ and $(0,1)$ with different probabilities conditional on $X_1 = X_2$. We conclude that when actions are strategic substitutes, equation (S-3) can be violated with Nash equilibrium behavior if, conditional on $X_1 = X_2$, either the strategic effects can differ in magnitude across players with nonzero probability, or the selection mechanism picks $(1,0)$ and $(0,1)$ in the region of multiple NE with different probabilities.

S1.3 Violations of (S-3) when actions are strategic complements

In this case we have $\Delta_1 \geq 0$ and $\Delta_2 \geq 0$. The equilibrium regions involving the outcomes $(1,0)$ and $(0,1)$ are now the following.

- $(1,0)$ will be the unique PSNE if $t_1 > 0$ and $t_2 < -\Delta_2$.
- $(0,1)$ will be the unique PSNE if $t_1 < -\Delta_1$ and $t_2 > 0$.
- If $\Delta_1 \neq 0$ and $\Delta_2 \neq 0$ and $-\Delta_p \leq t_p \leq 0$ for $p = 1,2$, there exist three NE. We have two PSNE: $(0,0)$ and $(1,1)$, and a MSNE where player p chooses $Y_p = 1$ with probability π_p, with $(\pi_1, \pi_2) = \left(-\frac{t_2}{\Delta_2}\ ,\ -\frac{t_1}{\Delta_1}\right)$

As before, let $\mathscr{M}_\pi$ denote the indicator for whether the MSNE is selected in the multiple equilibrium region. Let

$$\mathscr{G}_c \equiv \mathbb{1}\Big\{\, 0 \leq -\Delta_p \leq t_p \leq 0 \text{ and } \Delta_p \neq 0 \text{ for } p = 1,2 \Big\}.$$

This is the indicator for whether the game is in the multiple NE region. We now have,

$$\begin{aligned}
&Pr(1,0|X_1 = X_2) = \\
&Pr\Big(t_1 > 0 \;,\; t_2 < -\Delta_2 \;\Big|\; X_1 = X_2\Big) \\
&+E\left[\left(-\frac{t_2}{\Delta_2}\right)\cdot\left(1+\frac{t_1}{\Delta_1}\right) \;\Big|\; \mathscr{G}_c = 1,\, \mathscr{M}_\pi = 1,\, X_1 = X_2\right]\cdot Pr\Big(\mathscr{G}_c = 1 \;\Big|\; \mathscr{M}_\pi = 1,\, X_1 = X_2\Big)\cdot Pr\Big(\mathscr{M}_\pi = 1 \;\Big|\; X_1 = X_2\Big),
\end{aligned}$$

and,

$$\begin{aligned}
&Pr(1,0|X_1 = X_2) = \\
&Pr\Big(t_1 < -\Delta_1 \;,\; t_2 > 0 \;\Big|\; X_1 = X_2\Big) \\
&+E\left[\left(1+\frac{t_2}{\Delta_2}\right)\cdot\left(-\frac{t_1}{\Delta_1}\right) \;\Big|\; \mathscr{G}_c = 1,\, \mathscr{M}_\pi = 1,\, X_1 = X_2\right]\cdot Pr\Big(\mathscr{G}_c = 1 \;\Big|\; \mathscr{M}_\pi = 1,\, X_1 = X_2\Big)\cdot Pr\Big(\mathscr{M}_\pi = 1 \;\Big|\; X_1 = X_2\Big),
\end{aligned}$$

Again, we see that even with exchangeability of t_1 and t_2 whenever $X_1 = X_2$, we can still have $Pr(1,0|X_1 = X_2) \neq Pr(0,1|X_1 = X_2)$ if $Pr(\Delta_1 \neq \Delta_2|X_1 = X_2) > 0$; that is, if the magnitude of strategic effects can be different with nonzero probability if $X_1 = X_2$. However, unlike the strategic substitutes case, if $Pr(\Delta_1 = \Delta_2|X_1 = X_2)$ and strategic effects are always identical conditional on $X_1 = X_2$, we will now have $Pr(1,0|X_1 = X_2) = Pr(0,1|X_1 = X_2)$. We conclude that if payoff functions are completely symmetric conditional on $X_1 = X_2$ and the game is one of strategic complements, then Nash equilibrium behavior will be observationally equivalent to cooperation as captured by the restriction in equation (S-3). This condition will be violated with strategic complements if strategic effects are allowed to have different magnitudes across both players with positive probability conditional on $X_1 = X_2$.

S1.4 Violations of (S-3) when Δ_1 and Δ_2 have opposite signs

Suppose $\Delta_1 > 0$ and $\Delta_2 < 0$ (the reverse case is analogous). As we had pointed out before, allowing for the magnitude of strategic effects to differ across players when $X_1 = X_2$ can lead to violations of (S-3) with Nash equilibrium behavior, and we will confirm that in this case. When $\Delta_1 > 0$ and $\Delta_2 < 0$ the game will now have a unique Nash equilibrium for any given realization of payoff covariates. The equilibrium regions involving the outcomes $(1,0)$ and $(0,1)$ are now described as follows.

- $(1,0)$ will be the unique PSNE if $t_1 > 0$ and $t_2 < -\Delta_2$.
- $(0,1)$ will be the unique PSNE if $t_1 < -\Delta_1$ and $t_2 > 0$.

  If $-\Delta_1 \leq t_1 \leq 0$ and $0 \leq t_2 \leq -\Delta_2$, the game has a unique MSNE given by $(\pi_1, \pi_2) = \left(-\frac{t_2}{\Delta_2}\,,\, -\frac{t_1}{\Delta_1}\right)$.

Since equilibria are always unique, the equilibrium selection mechanism is irrelevant. Let

$$\mathcal{G}_m \equiv \mathbb{1}\{-\Delta_1 \leq t_1 \leq 0 \text{ and } 0 \leq t_2 \leq -\Delta_2\}.$$

We now have,

$$\begin{aligned} Pr(1,0|X_1 = X_2) = {} & Pr\Big(t_1 > 0\,,\, t_2 < -\Delta_2 \;\Big|\; X_1 = X_2\Big) \\ & + E\left[\left(-\frac{t_2}{\Delta_2}\right)\cdot\left(1+\frac{t_1}{\Delta_1}\right) \;\middle|\; \mathcal{G}_m = 1,\, X_1 = X_2\right]\cdot Pr\Big(\mathcal{G}_m = 1 \;\Big|\; X_1 = X_2\Big), \end{aligned}$$

and,

$$\begin{aligned} Pr(1,0|X_1 = X_2) = {} & Pr\Big(t_1 < -\Delta_1\,,\, t_2 > 0 \;\Big|\; X_1 = X_2\Big) \\ & + E\left[\left(1+\frac{t_2}{\Delta_2}\right)\cdot\left(-\frac{t_1}{\Delta_1}\right) \;\middle|\; \mathcal{G}_m = 1,\, X_1 = X_2\right]\cdot Pr\Big(\mathcal{G}_m = 1 \;\Big|\; X_1 = X_2\Big), \end{aligned}$$

And so, we see that having $Pr(\Delta_1 \neq \Delta_2 | X_1 = X_2) > 0$ (as it is the case when Δ_1 and Δ_2 have opposite signs) can lead to having $Pr(1,0|X_1 = X_2) \neq Pr(0,1|X_1 = X_2)$, thus violating (S-3) when the true behavior is Nash equilibrium.

S2 Details of the proof of Theorem 2

The steps of the proof were described in the appendix of the paper. Here we present the step-by-step details and derivations. As we described in the main body of the paper (equation (9)), the population statistic we employ to test for cooperation is given by

$$\mathcal{T} \equiv \sum_{y\in\mathcal{S}} \sum_{y'\in\sigma(y)} \sum_{x^d\in\mathcal{X}^d} \int_{x^c\in\mathcal{X}^c} \tau(y,y'|x)^2 \omega(x) dx^c,$$

where

$$\tau(y,y'|x) \equiv (Q(y|x) - Q(y'|x))\cdot f_{X_G}(\boldsymbol{e}\otimes x).$$

Next, recall that $X_{Gi} \equiv (X_{1i},\ldots,X_{Pi})$ denotes the collection of all players' observable characteristics, and that $X_G^c \equiv (X_1^c,\ldots,X_P^c)$ and $X_G^d \equiv (X_1^d,\ldots,X_P^d)$ denote the collection of all players' continuous and discrete observable characteristics, respectively. Also, recall that, for a given $x \equiv (x^c, x^d)$, we defined,

$$\mathcal{K}\left(\frac{X^c_{Gi} - \boldsymbol{e} \otimes x^c}{h_n}\right) \equiv \prod_{p=1}^{P} \prod_{\ell=1}^{r} \kappa\left(\frac{X^{c_\ell}_{pi} - x^{c_\ell}}{h_n}\right), \quad \mathbb{1}\left\{X^d_{Gi} = \boldsymbol{e} \otimes x^d\right\} \equiv \prod_{p=1}^{P} \prod_{s=1}^{m} \mathbb{1}\left\{X^{d_s}_{pi} = x^{d_s}\right\},$$

$$\Gamma(X_{Gi}, x, h_n) \equiv \frac{1}{h_n^L} \mathcal{K}\left(\frac{X^c_{Gi} - \boldsymbol{e} \otimes x^c}{h_n}\right) \cdot \mathbb{1}\left\{X^d_{Gi} = \boldsymbol{e} \otimes x^d\right\},$$

where $\boldsymbol{e} \equiv (1, \ldots, 1) \in \mathbb{R}^P$ denotes the vector of all-ones in $\mathbb{R}^P$, and where $r \equiv \dim(x^c)$ and $L \equiv P \cdot r$, so r is the number of continuously distributed covariates in X^c_{pi}, and L is the total number of continuously distributed covariates in X_{Gi}. As described in Assumption E3 , the support of the kernel $\kappa(\cdot)$ is given by the interval $[-S, S]$ (i.e, $\kappa(z) = 0$ for all $z \notin [-S, S]$).

As in the paper, group $U_i \equiv (Y_i, X_{Gi})$, let $S(Y_i, y, y') \equiv \mathbb{1}\{Y_i = y\} - \mathbb{1}\{Y_i = y'\}$, denote $S(Y_i, Y_j, y, y') \equiv S(Y_i, y, y') \cdot S(Y_j, y, y')$, and

$$\overline{S}(Y_i, Y_j) \equiv \sum_{y \in \mathcal{S}} \sum_{y' \in \sigma(y)} S(Y_i, Y_j, y, y'),$$

$$\varphi(X_{Gi}, X_{Gj}, h_n) \equiv \sum_{x^d \in \mathcal{X}^d} \int_{x^c \in \mathcal{X}^c} \Gamma(X_{Gi}, x, h_n) \cdot \Gamma(X_{Gj}, x, h_n) \omega(x) dx^c,$$

$$H_n(U_i, U_j) \equiv \overline{S}(Y_i, Y_j) \cdot \varphi(X_{Gi}, X_{Gj}, h_n).$$

Note that H_n is symmetric, since $H_n(U_i, U_j) = H_n(U_j, U_j)$. Our estimator for $\mathcal{T}$ can be expressed as,

$$\widehat{\mathcal{T}} = \binom{n}{2}^{-1} \sum_{i<j} H_n(U_i, U_j)$$

We will establish the asymptotic properties of $\widehat{\mathcal{T}}$ by verifying the conditions in Theorem 1 of Hall (1986), which result in asymptotic normality of degenerate U-statistics of order two. To verify these conditions we need to analyze various functionals, which we will describe next. In what follows, let (i, j, k) denote three distinct observations from our iid sample $(U_i)_{i=1}^n$. Let

$$\widetilde{H}_n(U_j, U_k) \equiv E\left[H_n(U_i, U_j) \cdot H_n(U_i, U_k) \middle| U_k, U_k\right]. \tag{S-4}$$

We will begin by analyzing $E[\widetilde{H}_n(U_j,U_k)]$ and $E[\widetilde{H}_n(U_j,U_k)^2]$. For given y_1, y_2, let

$$\eta_a(y_1,y_2|X_G) \equiv E[\overline{S}(Y,y_1)\cdot\overline{S}(Y,y_2)|X_G],$$
$$\Xi_a\left(X^c_{Gj},X^c_{Gk},x^d\middle|u\right) \equiv E\left[\eta_a\left(Y_j,Y_k|u,\boldsymbol{e}\otimes x^d\right)\middle|X^c_{Gj},X^c_{Gk},X^d_{Gj}=X^d_{Gk}=\boldsymbol{e}\otimes x^d\right]$$

We have,

$$E\left[\widetilde{H}_n(U_j,U_k)\right]=\frac{1}{h_n^{4L}}\sum_{x^d}\int_{x_1^c}\int_{x_2^c}\int_{u_1}\int_{u_2}\int_{u_3}\left(\Xi_a(u_2,u_3,x^d|u_1)\mathcal{K}\left(\frac{u_1-\boldsymbol{e}\otimes x_1^c}{h_n}\right)\mathcal{K}\left(\frac{u_1-\boldsymbol{e}\otimes x_2^c}{h_n}\right)\mathcal{K}\left(\frac{u_2-\boldsymbol{e}\otimes x_1^c}{h_n}\right)\right.$$
$$\left.\cdot\mathcal{K}\left(\frac{u_3-\boldsymbol{e}\otimes x_2^c}{h_n}\right)f_{X_G^c|X_G^d}(u_1|\boldsymbol{e}\otimes x^d)f_{X_G^c|X_G^d}(u_2|\boldsymbol{e}\otimes x^d)f_{X_G^c|X_G^d}(u_3|\boldsymbol{e}\otimes x^d)\omega(x_1^c,x^d)\omega(x_2^c,x^d)\right)$$
$$du_3du_2du_1dx_2^cdx_1^c\cdot f_{X_G^d}(\boldsymbol{e}\otimes x^d)^3 \tag{S-5}$$

Let

$$\xi_{a,n}\equiv\sum_{x^d}\int_{x^c}\left(\Xi_a\left(\boldsymbol{e}\otimes x^c,\boldsymbol{e}\otimes x^c,x^d\middle|\boldsymbol{e}\otimes x^c\right)f_{X_G}(\boldsymbol{e}\otimes x)^3\omega(x)^2\int_{\Delta}\left\{\int_{\psi_1}\mathcal{K}(\psi_1)\mathcal{K}(\psi_1-\boldsymbol{e}\otimes\Delta)d\psi_1\int_{\psi_2}\mathcal{K}(\psi_2)d\psi_2\right.\right.$$
$$\left.\left.\cdot\int_{\psi_3}\mathcal{K}(\psi_3-\boldsymbol{e}\otimes\Delta)d\psi_3\right\}d\Delta\right)dx^c,$$

The sum for x^d and the integral for x^c are taken over $\mathcal{X}^d$ and $\mathcal{X}^c$, respectively. We have $\psi_j\in\mathbb{R}^L$ and $\Delta\in\mathbb{R}^r$, and the limits of integration are $[-\boldsymbol{e}\otimes S,\boldsymbol{e}\otimes S]$ for each ψ_j, and $[\ell b_n(x^c),ub_n(x^c)]$ for Δ, where

$$\ell b_n(x^c)\equiv\frac{\underline{x}^c-x^c}{h_n}\vee -2S,\quad\text{and}\quad ub_n(x^c)\equiv\frac{\underline{x}^c-x^c}{h_n}\wedge 2S, \tag{S-6}$$

where $\cdot\vee\cdot$ denotes the element-wise maximum and $\cdot\wedge\cdot$ denotes the element-wise minimum, and where (as defined in Assumption E1), $\underline{x}^c$ and $\underline{x}^c$ denote the (element-wise) minimum and maximum values of x^c for which $\omega(x^c,x^d)>0$ for some x^d. Relabeling $x_1^c\equiv x^c$ and using the change of variables $u_j=\boldsymbol{e}\otimes x^c+h_n\cdot\psi_j$, for $j=1,2,3$, and $x_2^c=x^c+h_n\cdot\Delta$, under Assumptions E1-E3, the expectation in (S-5) becomes,

$$E\left[\widetilde{H}_n(U_j,U_k)\right]=\frac{1}{h_n^{L-r}}\cdot\left(\xi_{a,n}+o(1)\right)=O\left(\frac{1}{h_n^{L-r}}\right). \tag{S-7}$$

Next we focus on $E\left[\widetilde{H}_n(U_j,U_k)^2\right]$. Let

$$\Xi_b\left(X^c_{Gj},X^c_{Gk},x^d\middle|u_1,u_2\right)\equiv E\left[\eta_a\left(Y_j,Y_k\middle|u_1,\boldsymbol{e}\otimes x^d\right)\cdot\eta_a\left(Y_j,Y_k\middle|u_2,\boldsymbol{e}\otimes x^d\right)\middle|X^c_{Gj},X^c_{Gk},X^d_{Gj}=X^d_{Gk}=\boldsymbol{e}\otimes x^d\right].$$

We have,

$$\begin{aligned}E\left[\widetilde{H}_n(U_j,U_k)^2\right]&=\frac{1}{h_n^{8L}}\sum_{x^d}\int_{x_1^c}\int_{x_2^c}\int_{x_3^c}\int_{u_1}\int_{u_2}\int_{u_3}\int_{u_4}\Bigg(\Xi_b(u_3,u_4,x^d|u_1,u_2)\mathcal{K}\left(\frac{u_1-\boldsymbol{e}\otimes x_1^c}{h_n}\right)\mathcal{K}\left(\frac{u_1-\boldsymbol{e}\otimes x_2^c}{h_n}\right)\\&\cdot\mathcal{K}\left(\frac{u_2-\boldsymbol{e}\otimes x_3^c}{h_n}\right)\mathcal{K}\left(\frac{u_2-\boldsymbol{e}\otimes x_4^c}{h_n}\right)\mathcal{K}\left(\frac{u_3-\boldsymbol{e}\otimes x_1^c}{h_n}\right)\mathcal{K}\left(\frac{u_3-\boldsymbol{e}\otimes x_3^c}{h_n}\right)\mathcal{K}\left(\frac{u_4-\boldsymbol{e}\otimes x_2^c}{h_n}\right)\mathcal{K}\left(\frac{u_4-\boldsymbol{e}\otimes x_4^c}{h_n}\right)\\&\cdot f_{X_G^c|X_G^d}(u_1|\boldsymbol{e}\otimes x^d)f_{X_G^c|X_G^d}(u_2|\boldsymbol{e}\otimes x^d)f_{X_G^c|X_G^d}(u_3|\boldsymbol{e}\otimes x^d)f_{X_G^c|X_G^d}(u_4|\boldsymbol{e}\otimes x^d)\omega(x_1^c,x^d)\omega(x_2^c,x^d)\omega(x_3^c,x^d)\\&\cdot\omega(x_4^c,x^d)\Bigg)du_4du_3du_2du_1dx_4^cdx_3^cdx_2^cdx_1^c\cdot f_{X_G^d}(\boldsymbol{e}\otimes x^d)^4.\end{aligned}\tag{S-8}$$

Let

$$\begin{aligned}\xi_{b,n}\equiv\sum_{x^d}\int_{x^c}\Bigg(&\Xi_b\left(\boldsymbol{e}\otimes x^c,\boldsymbol{e}\otimes x^c,x^d\middle|\boldsymbol{e}\otimes x^c,\boldsymbol{e}\otimes x^c\right)f_{X_G}(\boldsymbol{e}\otimes x)^4\omega(x)^4\int_{\Delta_2}\int_{\Delta_3}\int_{\Delta_4}\Bigg\{\int_{\psi}\mathcal{K}(\psi_1)\mathcal{K}(\psi_1-\boldsymbol{e}\otimes\Delta_2)d\psi_1\\&\int_{\psi_2}\mathcal{K}(\psi_2-\boldsymbol{e}\otimes\Delta_3)\mathcal{K}(\psi_2-\boldsymbol{e}\otimes\Delta_4)d\psi_2\mathcal{K}(\psi_3)\mathcal{K}(\psi_3-\boldsymbol{e}\otimes\Delta_3)d\psi_3\int_{\psi_4}\mathcal{K}(\psi_4-\boldsymbol{e}\otimes\Delta_2)\mathcal{K}(\psi_4-\boldsymbol{e}\otimes\Delta_4)d\psi_4\Bigg\}\\&d\Delta_4d\Delta_3d\Delta_2\Bigg)dx^c,\end{aligned}$$

where $\psi_j\in\mathbb{R}^L$ and $\Delta_\ell\in\mathbb{R}^r$, and the limits of integration are $[-\boldsymbol{e}\otimes S,\boldsymbol{e}\otimes S]$ for each ψ_j, and $[\ell b_n(x^c),ub_n(x^c)]$ (as defined in (S-6)) for each Δ_ℓ. As before, the sum for x^d and the integral for x^c are taken over $\mathcal{X}^d$ and $\mathcal{X}^c$, respectively. Relabeling $x_1^c\equiv x^c$ and using the change of variables $u_j=\boldsymbol{e}\otimes x^c+h_n\cdot\psi_j$, for $j=1,2,3,4$, and $x_\ell^c=x^c+h_n\cdot\Delta_\ell$, for $\ell=2,3,4$, under Assumptions E1-E3, the expectation in (S-8) becomes,

$$E\left[\widetilde{H}_n(U_j,U_k)^2\right]=\frac{1}{h_n^{4L-3r}}\cdot\left(\xi_{b,n}+o(1)\right)=O\left(\frac{1}{h_n^{4L-3r}}\right).\tag{S-9}$$

Next we focus our attention on $E\left[H_n(U_i,U_j)^2\right]$. Let

$$\mu_2\left(X^c_{Gi},X^c_{Gj},x^d\right)\equiv E\left[\overline{S}(Y_i,Y_j)^2\middle|X^c_{Gi},X^c_{Gj},X^d_{Gi}=X^d_{Gj}=\boldsymbol{e}\otimes x^d\right].$$

We have,

$$E\Big[H_n(U_i,U_j)^2\Big] = \frac{1}{h_n^{4L}}\sum_{x^d}\int_{x_1^c}\int_{x_2^c}\int_{u_1}\int_{u_2}\Bigg(\mu_2(u_1,u_2,x_d)\mathcal{K}\left(\frac{u_1-\boldsymbol{e}\otimes x_1^c}{h_n}\right)\mathcal{K}\left(\frac{u_1-\boldsymbol{e}\otimes x_1^c}{h_n}\right)\mathcal{K}\left(\frac{u_2-\boldsymbol{e}\otimes x_1^c}{h_n}\right)$$
$$\cdot\mathcal{K}\left(\frac{u_2-\boldsymbol{e}\otimes x_1^c}{h_n}\right)f_{X_G^c|X_G^d}(u_1|\boldsymbol{e}\otimes x^d)f_{X_G^c|X_G^d}(u_2|\boldsymbol{e}\otimes x^d)\omega(x_1^c,x^d)\omega(x_2^c,x^d)\Bigg)du_2du_1dx_2^cdx_1^c\cdot f_{X_G^d}(\boldsymbol{e}\otimes x^d)^2 \tag{S-10}$$

Let

$$\sigma_n^2 \equiv \sum_{x^d}\int_{x^c}\Bigg(\mu_2\Big(\boldsymbol{e}\otimes x^c,\boldsymbol{e}\otimes x^c,x^d\Big)f_{X_G}(\boldsymbol{e}\otimes x)^2\omega(x)^2\int_{\Delta}\left[\int_{\psi}\mathcal{K}(\psi)\mathcal{K}(\psi-\boldsymbol{e}\otimes\Delta)d\psi\right]^2 d\Delta\Bigg)dx^c \tag{S-11}$$

Once again, the sum for x^d and the integral for x^c are taken over $\mathcal{X}^d$ and $\mathcal{X}^c$, respectively, while $\psi\in\mathbb{R}^L$ is integrated over the range $[-\boldsymbol{e}\otimes S,\boldsymbol{e}\otimes S]$ and $\Delta\in\mathbb{R}^r$ is integrated over $[\ell b_n(x^c),ub_n(x^c)]$, as described in (S-6). Relabeling $x_1^c\equiv x^c$ and using the change of variables $u_j=\boldsymbol{e}\otimes x^c+h_n\cdot\psi_j$, for $j=1,2$, and $x_2^c=x^c+h_n\cdot\Delta$, under Assumptions E1-E3, the expectation in (S-10) becomes,

$$E\Big[H_n(U_i,U_j)^2\Big] = \frac{1}{h_n^{2L-r}}\cdot\Big(\sigma_n^2+o(1)\Big). \tag{S-12}$$

Let

$$\mu_3\Big(X_{Gi}^c,X_{Gj}^c,x^d\Big)\equiv E\Big[\overline{S}(Y_i,Y_j)^3\Big|X_{Gi}^c,X_{Gj}^c,X_{Gi}^d=X_{Gj}^d=\boldsymbol{e}\otimes x^d\Big],$$
$$\mu_4\Big(X_{Gi}^c,X_{Gj}^c,x^d\Big)\equiv E\Big[\overline{S}(Y_i,Y_j)^4\Big|X_{Gi}^c,X_{Gj}^c,X_{Gi}^d=X_{Gj}^d=\boldsymbol{e}\otimes x^d\Big].$$

And,

$$\lambda_{3,n}\equiv\sum_{x^d}\int_{x^c}\Bigg(\mu_3\Big(\boldsymbol{e}\otimes x^c,\boldsymbol{e}\otimes x^c,x^d\Big)f_{X_G}(\boldsymbol{e}\otimes x)^2\omega(x)^3\int_{\Delta_2}\int_{\Delta_3}\left[\int_{\psi}\mathcal{K}(\psi)\mathcal{K}(\psi-\boldsymbol{e}\otimes\Delta_2)\mathcal{K}(\psi-\boldsymbol{e}\otimes\Delta_3)d\psi\right]^2$$
$$d\Delta_3d\Delta_2\Bigg)dx^c,$$

$$\lambda_{4,n}\equiv\sum_{x^d}\int_{x^c}\Bigg(\mu_4\Big(\boldsymbol{e}\otimes x^c,\boldsymbol{e}\otimes x^c,x^d\Big)f_{X_G}(\boldsymbol{e}\otimes x)^2\omega(x)^4\int_{\Delta_2}\int_{\Delta_3}\int_{\Delta_4}\Bigg[\int_{\psi}\mathcal{K}(\psi)\mathcal{K}(\psi-\boldsymbol{e}\otimes\Delta_2)\mathcal{K}(\psi-\boldsymbol{e}\otimes\Delta_3)$$
$$\cdot\mathcal{K}(\psi-\boldsymbol{e}\otimes\Delta_4)d\psi\Bigg]^2 d\Delta_4d\Delta_3d\Delta_2\Bigg)dx^c,$$

As before, the sum for x^d and the integral for x^c are taken over $\mathcal{X}^d$ and $\mathcal{X}^c$, respectively, while each $\psi_j\in\mathbb{R}^L$ is integrated over the range $[-\boldsymbol{e}\otimes S,\boldsymbol{e}\otimes S]$ and each $\Delta_\ell\in\mathbb{R}^r$ is integrated over $[\ell b_n(x^c), ub_n(x^c)]$, as described in (S-6). Using the same change of variables as in the previous results, under Assumptions E1-E3, we obtain,

$$
\begin{aligned}
E\left[H_n(U_i,U_j)^3\right] &= \frac{1}{h_n^{4L-2r}}\cdot\left(\lambda_{3,n}+o(1)\right),\\
E\left[H_n(U_i,U_j)^4\right] &= \frac{1}{h_n^{6L-3r}}\cdot\left(\lambda_{4,n}+o(1)\right).
\end{aligned}
\tag{S-13}
$$

Equipped with the results in equations (S-7), (S-9), (S-12) and (S-13), we proceed to analyze the Hoeffding decomposition of $\widehat{\mathcal{T}}$. Recall that $\widehat{\mathcal{T}}$ is given by,

$$
\widehat{\mathcal{T}} = \binom{n}{2}^{-1} H_n(U_i,U_j),
$$

where, as we pointed out previously, $H_n(U_i,U_j)$ is symmetric. Denote

$$
m_n(U_i) \equiv E\left[H_n(U_i,U_j)|U_i\right].
$$

And note that $E\left[m_n(U_i)\right] = E\left[H_n(U_i,U_j)\right]$ by iterated expectations. Let

$$
\overline{\varphi}(X_{Gi}|y,y',h_n) \equiv \sum_{x^d}\int_{x^c} \tau(y,y'|x)\Gamma(X_{Gi},x,h_n)\omega(x)dx^c.
$$

Under the conditions of Assumptions E1-E3, there exists a finite constant $\overline{B}$ such that, using an M^{th}–order approximation,

$$
\begin{aligned}
m_n(U_i) &= \sum_{y\in\mathcal{S}}\sum_{y'\in\sigma(y)} S(Y_i,y,y')\overline{\varphi}(X_{Gi}|y,y',h_n) + B_{2n}(U_i),\\
E\left[m_n(U_i)\right] &= \mathcal{T} + B_{1n},
\end{aligned}
\tag{S-14}
$$

where $|B_{2n}(U_i)| \le \overline{B}\cdot h_n^{M+r-L}$, and $|B_{1n}| \le \overline{B}\cdot h_n^{M+r-L}$. Note that under the null hypothesis of cooperation, we have $\tau(y,y'|x) = 0$ for all $y\in\mathcal{S}$, $y'\in\sigma(y)$ and all $x\in\mathcal{X}$. Therefore, under the null hypothesis of cooperation, we have

$$
\left.\begin{aligned}
m_n(U_i) &= B_{2n}(U_i),\\
E\left[m_n(U_i)\right] &= B_{1n}
\end{aligned}\right\}\quad \text{where}\quad |B_{2n}(U_i)| \le \overline{B}\cdot h_n^{M+r-L}, \text{ and } |B_{1n}| \le \overline{B}\cdot h_n^{M}.
\tag{S-14'}
$$

Next, denote

$$\begin{aligned} b_n(U_i) &\equiv m_n(U_i) - E[m_n(U_i)], \\ G_n(U_i,U_j) &\equiv H_n(U_i,U_j) - E\Big[H_n(U_i,U_j)\Big] - b_n(U_i) - b_n(U_j). \end{aligned} \tag{S-15}$$

The Hoeffding decomposition of the U-statistic $\binom{n}{2}^{-1}\sum_{i<j} H_n(U_i,U_j)$ (see (see Serfling (1980, pages 177-178)) is given by,

$$\begin{aligned} \widehat{\mathcal{T}} &= E\Big[H_n(U_i,U_j)\Big] + \frac{2}{n}\sum_{i=1}^{n} b_n(U_i) + \binom{n}{2}^{-1}\sum_{i<j} G_n(U_i,U_j) \\ &= \mathcal{T} + \frac{2}{n}\sum_{i=1}^{n} b_n(U_i) + \binom{n}{2}^{-1}\sum_{i<j} G_n(U_i,U_j) + O\Big(h_n^{M}\Big), \end{aligned} \tag{S-16}$$

where the last result follows from the result in (S-14). We begin by analyzing the asymptotic properties of the decomposition in (S-16) under the null hypothesis of cooperation. In this case, using the result in (S-14'), we have $\mathcal{T} = 0$ and $\left|\frac{1}{n}\sum_{i=1}^{n} b_n(U_i)\right| \leq \overline{B}\cdot h_n^{M+r-L}$, and (S-16) becomes,

$$\widehat{\mathcal{T}} = \binom{n}{2}^{-1}\sum_{i<j} G_n(U_i,U_j) + O_p\Big(h_n^{M+r-L}\Big) \tag{S-16'}$$

Thus, under the null hypothesis of cooperation, our statistic $\widehat{\mathcal{T}}$ is the sum of a degenerate U-statistic of order two, plus a remainder that vanishes to zero and has order of magnitude $O\Big(h_n^{M+r-L}\Big)$. We apply the results in Hall (1986) (specifically, Theorem 1 in that paper) to analyze the asymptotic properties of this U-statistic. As Hall (1986) shows, degenerate U-statistics of order two can be asymptotically normal when their kernel function depends on n. This is precisely the case of $G_n(U_i,U_j)$, which depends on n through the presence of the bandwidth sequence h_n. Let (i,j,k) denote three distinct observations from our iid sample $(U_i)_{i=1}^n$. Let

$$\widetilde{G}_n(U_j,U_k) \equiv E\Big[G_n(U_i,U_j)\cdot G_n(U_i,U_k)\Big|U_j,U_k\Big].$$

Let $G_n(U_i,U_j)$ be a symmetric function satisfying $E[G_n(U_i,U_j)|U_i] = 0$ almost surely, and $E[G_n(U_i,U_j)^2] < \infty$ for each n. Suppose,

$$\frac{E\Big[\widetilde{G}_n(U_j,U_k)^2\Big] + n^{-1}E\Big[G_n(U_i,U_j)^4\Big]}{\Big(E\Big[G_n(U_i,U_j)^2\Big]\Big)^2} \longrightarrow 0, \tag{S-17}$$

as $n \to \infty$. Theorem 1 in Hall (1986) shows that, if these conditions are satisfied, then $\sum_{i<j} G_n(U_i, U_j)$ is asymptotically normally distributed with zero mean and variance given by $\frac{1}{2}n^2 \cdot E\left[G_n(U_i, U_j)^2\right]$. Therefore, $\binom{n}{2}^{-1} \sum_{i<j} G_n(U_i, U_j)$ will be asymptotically normally distributed with zero mean and variance given by $\frac{2}{(n-1)^2} \cdot E\left[G_n(U_i, U_j)^2\right]$. Since $n/(n-1) \to 1$ as $n \to \infty$, it follows that if the condition in (S-17) is satisfied, we have

$$n \cdot \left(\frac{\binom{n}{2}^{-1} \sum_{i<j} G_n(U_i, U_j)}{\sqrt{2 \cdot E\left[G_n(U_i, U_j)^2\right]}} \right) \xrightarrow{d} \mathcal{N}(0, 1). \tag{S-18}$$

Plugging in (S-14') into (S-15), we have that, under the null hypothesis of cooperation,

$$G_n(U_i, U_j) = H_n(U_i, U_j) - \left(B_{2n}(U_i) + B_{2n}(U_j) - B_{1n}\right), \tag{S-19}$$

where $\left|B_{2n}(U_i) + B_{2n}(U_j)\right| \leq 2\overline{B} \cdot h_n^{M+r-L}$, and $|B_{1n}| \leq \overline{B} \cdot h_n^M$, so under the null hypothesis of cooperation,

$$\begin{aligned} \left|B_{2n}(U_i) + B_{2n}(U_j) - B_{1n}\right| &= O\left(h_n^{M+r-L}\right) + O\left(h_n^M\right) = O\left(h_n^{M+r-L}\right), \\ G_n(U_i, U_j) &= H_n(U_i, U_j) + O\left(h_n^{M+r-L}\right). \end{aligned} \tag{S-20}$$

Let $\mu_1(y, X_G) \equiv E\left[\left|\overline{S}(Y, y)\right| \,\middle|\, X_G\right]$. We have,

$$\begin{aligned} E\left[\left|H_n(U_i, U_j)\right| \,\middle|\, U_j\right] = \frac{1}{h_n^{2L}} \cdot \sum_{x^d} \int_{x^c} \Bigg(\int_u \mu_1\left(Y_j, u, \boldsymbol{e} \otimes x^d\right) \cdot \left|\mathcal{K}\left(\frac{u - \boldsymbol{e} \otimes x^c}{h_n}\right)\right| f_{X_G^c | X_G^d}(u | \boldsymbol{e} \otimes x^d) du \\ \cdot \left|\mathcal{K}\left(\frac{X_{Gj}^c - \boldsymbol{e} \otimes x^c}{h_n}\right)\right| \cdot \omega(x) \Bigg) dx^c \cdot f_{X_G^d}(\boldsymbol{e} \otimes x^d) \cdot \mathbb{1}\left\{X_{Gj}^d = \boldsymbol{e} \otimes x^d\right\} \end{aligned}$$

Using the change of variables $u = \boldsymbol{e} \otimes x^c + h_n \cdot \psi$, with $\psi \in \mathbb{R}^L$, and $x^c = X_{1j}^c + h_n \cdot \Delta$, with $\Delta \in \mathbb{R}^r$ (note that X_{1j}^c are the continuous payoff covariates of player 1, we could use those of any other player in our change of variables), we have,

$$\begin{aligned} E\left[\left|H_n(U_i, U_j)\right| \,\middle|\, U_j\right] = \frac{1}{h_n^{L-r}} \cdot \sum_{x^d} \int_{\Delta} \Bigg(\int_{\psi} \mu_1\left(Y_j, \boldsymbol{e} \otimes (X_{1j}^c + h_n \cdot \Delta) + h_n \cdot \psi, \boldsymbol{e} \otimes x^d\right) \cdot \left|\mathcal{K}(\psi)\right| \\ \cdot f_{X_G^c | X_G^d}(\boldsymbol{e} \otimes (X_{1j}^c + h_n \cdot \Delta) + h_n \cdot \psi | \boldsymbol{e} \otimes x^d) d\psi \cdot \left|\mathcal{K}\left(\frac{X_{Gj}^c - \boldsymbol{e} \otimes X_{1j}^c}{h_n} - \Delta\right)\right| \cdot \omega(X_{1j}^c + h_n \cdot \Delta, x^d) \Bigg) d\Delta \\ \cdot f_{X_G^d}(\boldsymbol{e} \otimes x^d) \cdot \mathbb{1}\left\{X_{Gj}^d = \boldsymbol{e} \otimes x^d\right\} = O\left(\frac{1}{h_n^{L-r}}\right). \end{aligned}$$

The same steps show that $E\left[|H_n(U_i,U_k)| \,\middle|U_k\right] = O\left(\frac{1}{h_n^{L-r}}\right)$. Combining these results with (S-20), we obtain,

$$\begin{aligned}
&\widetilde{G}_n(U_j,U_k) \equiv E\left[G_n(U_i,U_j)\cdot G_n(U_i,U_k)\middle|U_j,U_k\right] \\
&= E\left[\left(H_n(U_i,U_j)-\left(B_{2n}(U_i)+B_{2n}(U_j)-B_{1n}\right)\right)\cdot\left(H_n(U_i,U_k)-(B_{2n}(U_i)+B_{2n}(U_k)-B_{1n})\right)\middle|U_j,U_k\right] \\
&= E\left[H_n(U_i,U_j)\cdot H_n(U_i,U_k)\middle|U_j,U_k\right]+O\left(h_n^{M+r-L}\right)\cdot O\left(h_n^{r-L}\right)+O\left(h_n^{2\cdot(M+r-L)}\right) \\
&\equiv \widetilde{H}_n(U_j,U_k)+O\left(h_n^{M+2\cdot(r-L)}\right),
\end{aligned} \tag{S-21}$$

where $\widetilde{H}_n(U_j,U_k) \equiv E\left[H_n(U_i,U_j)\cdot H_n(U_i,U_k)\middle|U_k,U_k\right]$, as we defined in equation (S-4). Combining (S-21) with the results we obtained for $E\left[\widetilde{H}_n(U_j,U_k)\right]$ and $E\left[\widetilde{H}_n(U_j,U_k)^2\right]$ in equations (S-7) and (S-9), we have

$$\begin{aligned}
E\left[\widetilde{G}_n(U_j,U_k)^2\right] &= E\left[\widetilde{H}_n(U_j,U_k)^2\right]+O\left(h_n^{M+3\cdot(r-L)}\right) \\
&= O\left(\frac{1}{h_n^{4L-3r}}\right)+O\left(\frac{h_n^M}{h_n^{3L-3r}}\right)=\frac{1}{h_n^{4L-3r}}\cdot\left(O(1)+\underbrace{O\left(h_n^{M+L}\right)}_{=o(1)}\right) \\
&= O\left(\frac{1}{h_n^{4L-3r}}\right).
\end{aligned} \tag{S-22}$$

Next, combining (S-20) with the results we obtained for $E\left[H_n(U_i,U_j)^2\right]$, $E\left[H_n(U_i,U_j)^3\right]$ and $E\left[H_n(U_i,U_j)^4\right]$ in equations (S-12) and (S-13), and the result in (S-14), where we have $E\left[H_n(U_i,U_j)\right] = E\left[m_n(U_i)\right] = \mathcal{T}+B_{1n} = O(1)$, we obtain,

$$\begin{aligned}
E\left[G_n(U_i,U_j)^4\right] &= E\left[H_n(U_i,U_j)^4\right]+O\left(h_n^{M+3r-5L}\right) \\
&= O\left(\frac{1}{h_n^{6L-3r}}\right)+O\left(h_n^{M+3r-5L}\right)=\frac{1}{h_n^{6L-3r}}\cdot\left(O(1)+\underbrace{O\left(h_n^{M+L}\right)}_{=o(1)}\right) \\
&= O\left(\frac{1}{h_n^{6L-3r}}\right),
\end{aligned} \tag{S-23}$$

and,

$$
\begin{aligned}
E\left[G_n(U_i,U_j)^2\right] &= E\left[H_n(U_i,U_j)^2\right] + O\left(h_n^{M+r-L}\right) \\
&= \frac{1}{h_n^{2L-r}} \cdot \left(\sigma_n^2 + o(1)\right) + O\left(h_n^{M+r-L}\right) = \frac{1}{h_n^{2L-r}} \cdot \left(\left(\sigma_n^2 + o(1)\right) + \underbrace{O\left(h_n^{M+L}\right)}_{=o(1)}\right) \\
&= \frac{1}{h_n^{2L-r}} \cdot \left(\sigma_n^2 + o(1)\right),
\end{aligned}
\tag{S-24}
$$

where $\sigma_n^2 > 0$ is as described in (S-11). Plugging in the results in equations (S-22), (S-23) and (S-24) into the Hall (1986, Theorem 1) criterion described in (S-17), we have

$$
\begin{aligned}
\frac{E\left[\widetilde{G}_n(U_j,U_k)^2\right] + n^{-1}E\left[G_n(U_i,U_j)^4\right]}{\left(E\left[G_n(U_i,U_j)^2\right]\right)^2} &= \frac{O\left(\frac{1}{h_n^{4L-3r}}\right) + O\left(\frac{1}{n \cdot h_n^{6L-3r}}\right)}{\frac{1}{h_n^{4L-2r}} \cdot \left(\sigma_n^2 + o(1)\right)^2} \\
&= \frac{O(h_n^r) + O\left(\frac{1}{n \cdot h_n^{2L-r}}\right)}{\left(\sigma_n^2 + o(1)\right)^2} \longrightarrow 0,
\end{aligned}
\tag{S-25}
$$

where the last result follows since, as described in Assumption E3, we have $n \cdot h_n^{2L-r} \longrightarrow \infty$. From the result in (S-25), Theorem 1 in Hall (1986) implies that $n \cdot \left(\frac{\binom{n}{2}^{-1}\sum_{i<j} G_n(U_i,U_j)}{\sqrt{2 \cdot E\left[G_n(U_i,U_j)^2\right]}}\right) \xrightarrow{d} \mathcal{N}(0,1)$. Using the fact that $E\left[G_n(U_i,U_j)^2\right] = E\left[H_n(U_i,U_j)^2\right] + o(1)$, going back to the Hoeffding decomposition in (S-16′) and using the result in (S-24), we have that under the null hypothesis of cooperation,

$$
\begin{aligned}
\frac{n \cdot \widehat{\mathcal{T}}}{\sqrt{2 \cdot E\left[H_n(U_i,U_j)^2\right]}} &= n \cdot \left(\frac{\binom{n}{2}^{-1}\sum_{i<j} G_n(U_i,U_j)}{\sqrt{2 \cdot E\left[G_n(U_i,U_j)^2\right] + o(1)}}\right) + \frac{n \cdot h_n^{\frac{2L-r}{2}}}{\sqrt{2 \cdot \left(\sigma_n^2 + o(1)\right)}} \cdot O_p\left(h_n^{M+r-L}\right) \\
&= n \cdot \left(\frac{\binom{n}{2}^{-1}\sum_{i<j} G_n(U_i,U_j)}{\sqrt{2 \cdot E\left[G_n(U_i,U_j)^2\right] + o(1)}}\right) + \underbrace{O_p\left(n \cdot h_n^{M+\frac{r}{2}}\right)}_{=o_p(1)} \\
&\xrightarrow{d} \mathcal{N}(0,1).
\end{aligned}
\tag{S-26}
$$

where the last result follows from the bandwidth convergence restriction $n \cdot h_n^{M+\frac{r}{2}} \longrightarrow 0$, described in Assumption E3.

Now let us describe the asymptotic properties o $\widehat{\mathcal{T}}$ under the alternative hypothesis of no cooperation where (8) is violated. In this case we have $\mathcal{T} > 0$. Let us go back to the general Hoeffding decomposition result in (S-16),

$$\widehat{\mathcal{T}} = \mathcal{T} + \frac{2}{n}\sum_{i=1}^{n} b_n(U_i) + \binom{n}{2}^{-1}\sum_{i<j} G_n(U_i, U_j) + O\left(h_n^M\right),$$

We proceed by characterizing $Var[b_n(U_i)]$. Let

$$\Lambda(X_{Gi}, y, y', v, v') \equiv E\Big[S(Y_i, y, y')\cdot S(Y_i, v, v')\Big|X_{Gi}\Big].$$

And let

$$\lambda_{2,n} \equiv \sum_{y\in\mathcal{S}}\sum_{y'\in\sigma(y)}\sum_{v\in\mathcal{S}}\sum_{v'\in\sigma(v)}\sum_{x^d}\int_{x^c}\Bigg(\Lambda(\boldsymbol{e}\otimes x, y, y', v, v')\cdot f_{X_G}(\boldsymbol{e}\otimes x)\cdot\tau(y, y'|x)\cdot\tau(v, v'|x)\cdot\omega(x)^2$$
$$\cdot\int_{\Delta}\int_{\psi}\mathcal{K}(\psi)\mathcal{K}(\psi-\boldsymbol{e}\otimes\Delta)d\psi d\Delta\Bigg)dx^c.$$

The sum for x^d and the integral for x^c are taken over $\mathcal{X}^d$ and $\mathcal{X}^c$, respectively, while $\psi\in\mathbb{R}^L$ is integrated over the range $[-\boldsymbol{e}\otimes S, \boldsymbol{e}\otimes S]$ and $\Delta\in\mathbb{R}^r$ is integrated over $[\ell b_n(x^c), ub_n(x^c)]$, as described in (S-6). Under the null hypothesis of cooperation, we have $\lambda_{2,n} = 0$ (since $\tau(y, y'|x) = 0$ for all $y\in\mathcal{S}$, $y'\in\sigma(y)$), but under the alternative hypothesis of no cooperation where (8) is violated and $\mathcal{T} > 0$, we have $\lambda_{2,n}\neq 0$, and

$$Var[b_n(U_i)] = \frac{1}{h_n^{L-r}}\cdot(\lambda_{2,n} + o(1)), \quad \text{and} \quad \frac{2}{n}\sum_{i=1}^{n} b_n(U_i) = O_p\left(\frac{1}{n^{\frac{1}{2}}\cdot h_n^{\frac{L-r}{2}}}\right). \tag{S-27}$$

Next, recall that $\binom{n}{2}^{-1}\sum_{i<j} G_n(U_i, U_j)$ is a degenerate U-statistic of order 2. Thus, using Lemma 5.2.1.A in Serfling (1980) and the result in (S-24), we have

$$Var\left[\binom{n}{2}^{-1}\sum_{i<j} G_n(U_i, U_j)\right] = \frac{2}{n\cdot(n-1)}\cdot E\left[G_n(U_i, U_j)^2\right] = O\left(\frac{E\left[G_n(U_i, U_j)^2\right]}{n^2}\right) = O\left(\frac{1}{n^2\cdot h_n^{2L-r}}\right)$$

Therefore,

$$\binom{n}{2}^{-1}\sum_{i<j} G_n(U_i, U_j) = O_p\left(\frac{1}{n\cdot h_n^{\frac{2L-r}{2}}}\right) \tag{S-28}$$

Therefore, under the alternative hypothesis of no cooperation where (8) is violated and

$\mathcal{T} > 0$,

$$\widehat{\mathcal{T}} = \mathcal{T} + O_p\left(\frac{1}{n^{\frac{1}{2}} \cdot h_n^{\frac{L-r}{2}}}\right) + O_p\left(\frac{1}{n \cdot h_n^{\frac{2L-r}{2}}}\right) + O\left(h_n^M\right) = \mathcal{T} + O_p\left(\frac{1}{n^{\frac{1}{2}} \cdot h_n^{\frac{L-r}{2}}}\right) + O\left(h_n^M\right).$$

Where the last result follows since the restriction $n \cdot h_n^{2L-r} \longrightarrow \infty$ implies $n \cdot h_n^L \longrightarrow \infty$. Therefore, under the alternative hypothesis of no cooperation where (8) is violated and $\mathcal{T} > 0$,

$$\begin{aligned}
&\frac{n \cdot \widehat{\mathcal{T}}}{\sqrt{2 \cdot E\left[H_n(U_i, U_j)^2\right]}} \\
&= \frac{n \cdot h_n^{\frac{2L-r}{2}} \cdot \mathcal{T}}{\sqrt{2 \cdot \left(\sigma_n^2 + o(1)\right)}} + \frac{n \cdot h_n^{\frac{2L-r}{2}}}{\sqrt{2 \cdot \left(\sigma_n^2 + o(1)\right)}} \cdot O_p\left(\frac{1}{n^{\frac{1}{2}} \cdot h_n^{\frac{L-r}{2}}}\right) + \frac{n \cdot h_n^{\frac{2L-r}{2}}}{\sqrt{2 \cdot \left(\sigma_n^2 + o(1)\right)}} \cdot O\left(h_n^M\right) \\
&= \frac{n \cdot h_n^{\frac{2L-r}{2}} \cdot \mathcal{T}}{\sqrt{2 \cdot \left(\sigma_n^2 + o(1)\right)}} + O_p\left(n^{\frac{1}{2}} \cdot h_n^{\frac{L}{2}}\right) + O\left(n \cdot h_n^{M + \frac{2L-r}{2}}\right) \\
&= \frac{n \cdot h_n^{\frac{2L-r}{2}} \cdot \mathcal{T}}{\sqrt{2 \cdot \left(\sigma_n^2 + o(1)\right)}} + n^{\frac{1}{2}} \cdot h_n^{\frac{L}{2}} \cdot \left(O_p(1) + \underbrace{O\left(n^{\frac{1}{2}} \cdot h_n^{M + \frac{L-r}{2}}\right)}_{=o(1)}\right) \\
&= \frac{n \cdot h_n^{\frac{2L-r}{2}} \cdot \mathcal{T}}{\sqrt{2 \cdot \left(\sigma_n^2 + o(1)\right)}} + n^{\frac{1}{2}} \cdot h_n^{\frac{L}{2}} \cdot \vartheta_n, \quad \text{where} \quad |\vartheta_n| = O_p(1),
\end{aligned} \tag{S-29}$$

In the above result we used the fact that $n^{\frac{1}{2}} \cdot h_n^{M + \frac{L-r}{2}} \longrightarrow 0$ since $n^{\frac{1}{2}} \cdot h_n^{M + \frac{L-r}{2}} = n \cdot h_n^{M + \frac{r}{2}} \cdot \frac{h_n^{\frac{L}{2}-r}}{n^{\frac{1}{2}}}$, and $n \cdot h_n^{M + \frac{r}{2}} \longrightarrow 0$ by Assumption E3 , while $\frac{L}{2} - r = \frac{P \cdot r}{2} - r = r \cdot \left(\frac{P}{2} - 1\right) \geq 0$, since $P \geq 2$ (the game has at least two players). Therefore, $\frac{L}{2} - r \geq 0$ and we either have $h_n^{\frac{L}{2}-r} = 1$ (if $P = 2$), or $h_n^{\frac{L}{2}-r} \longrightarrow 0$ (if $P \geq 3$). Therefore, $\frac{h_n^{\frac{L}{2}-r}}{n^{\frac{1}{2}}} \longrightarrow 0$ and thus, $n^{\frac{1}{2}} \cdot h_n^{M + \frac{L-r}{2}} \longrightarrow 0$.

Under the bandwidth convergence restrictions in Assumption E3 we have[1] $n^{\frac{1}{2}} \cdot h_n^{\frac{L}{2}} \longrightarrow$

[1] Note that, $n^{\frac{1}{2}} \cdot h_n^{\frac{L}{2}} = \left(n \cdot h_n^{2L-r} \cdot \frac{1}{h_n^{L-r}}\right)^{\frac{1}{2}}$ and $n^{\frac{1}{2}} \cdot h_n^{\frac{L-r}{2}} = \left(n \cdot h_n^{2L-r} \cdot \frac{1}{h_n^L}\right)^{\frac{1}{2}}$. Thus, we have $n^{\frac{1}{2}} \cdot h_n^{\frac{L-r}{2}} \longrightarrow \infty$ since $n \cdot h_n^{2L-r} \longrightarrow \infty$ by Assumption E3. This condition and the fact that $L - r = r \cdot (P - 1) > 0$ yields $n^{\frac{1}{2}} \cdot h_n^{\frac{L}{2}} \longrightarrow \infty$.

∞ and $n^{\frac{1}{2}} \cdot h_n^{\frac{L-r}{2}} \longrightarrow \infty$. Take any sequence $c_n > 0$ such that,

$$c_n \longrightarrow +\infty, \quad \text{and} \quad \frac{c_n}{n^{\frac{1}{2}} \cdot h_n^{\frac{L}{2}}} \longrightarrow 0.$$

Going back to (S-29), under the alternative hypothesis of no cooperation where (8) is violated and $\mathcal{T} > 0$,

$$Pr\left(\frac{n \cdot \widehat{\mathcal{T}}}{\sqrt{2 \cdot E\left[H_n(U_i, U_j)^2\right]}} > c_n\right) = Pr\left(\vartheta_n > \frac{c_n}{n^{\frac{1}{2}} \cdot h_n^{\frac{L}{2}}} - \frac{n^{\frac{1}{2}} \cdot h_n^{\frac{L-r}{2}} \cdot \mathcal{T}}{\sqrt{2 \cdot \left(\sigma_n^2 + o(1)\right)}}\right) \longrightarrow 1,$$

where the last result follows from the fact that $\vartheta_n = O_p(1)$ and

$$\frac{c_n}{n^{\frac{1}{2}} \cdot h_n^{\frac{L}{2}}} - \frac{n^{\frac{1}{2}} \cdot h_n^{\frac{L-r}{2}} \cdot \mathcal{T}}{\sqrt{2 \cdot \left(\sigma_n^2 + o(1)\right)}} \longrightarrow -\infty.$$

Thus, under the alternative hypothesis of no cooperation where (8) is violated and $\mathcal{T} > 0$,

$$\frac{n \cdot \widehat{\mathcal{T}}}{\sqrt{2 \cdot E\left[H_n(U_i, U_j)^2\right]}} \longrightarrow +\infty \quad \text{with probability approaching one (w.p.a.1)} \tag{S-30}$$

The results in (S-26) and (S-30) prove Theorem 2. ■

S3 Monte Carlo experiments

Section 5 in the paper included a summary of the results of our Monte Carlo experiments. We present the full results and the details here. We applied our test to data generated for a 2×2 game with normal-form payoffs given as follows,

	$Y_2 = 1$	$Y_2 = 0$
$Y_1 = 1$	$X_1^1 + X_1^2 + \Delta_1 + \varepsilon_1$, $X_2^1 + X_2^2 + \Delta_2 + \varepsilon_2$	$X_1^1 + X_1^2 + \varepsilon_1$, 0
$Y_1 = 0$	0 , $X_2^1 + X_2^2 + \varepsilon_2$	0 , 0

In all our experiments, $(X_p^1, X_p^2, \varepsilon_p)$ are iid $\mathcal{N}(0,1)$, so we have $X_p \equiv (X_p^1, X_p^2)$, which includes two continuously distributed covariates (i.e, $r = 2$). The strategic interaction parameters (Δ_1, Δ_2) are constant, with $\Delta_p \leq 0$ (strategic substitutes). We will apply our test to data generated by cooperation (assuming three different joint objective functions that satisfy our assumptions) and to data generated by pure-strategy Nash equilibrium (PSNE) noncooperative behavior. As we show in Section S1.1 of this supplement, in this setting our test will have power when true behavior is Nash equilibrium if either $\Delta_1 \neq \Delta_2$ or if the selection mechanism chooses both coexisting Nash equilibria –(0,1) and (1,0)– with unequal probability in the multiple NE region.

We will use two DGPs. In the first DGP, we will set $\Delta_1 = -2$ and $\Delta_2 = -1$ and we will assume that the NE selection mechanism chooses both coexisting Nash equilibria in the multiple NE region with equal probability. Our goal here is to evaluate the power of our test when the strategic effects are different across players. In the second DGP, we will set $\Delta_1 = \Delta_2 = -2$ but we will assume that the selection mechanism chooses the coexisting NE with different probabilities. We will first assume that (1,0) is selected with probability 25% and then we will reduce this probability to 10%. Our goal is to evaluate the power of our test when the strategic effects are equal but the selection mechanism does not choose the coexisting NE with uniform probability. For both DGPs we will also generate data assuming cooperation and looking at three joint objective functions: $V(u_1,u_2) = u_1 + u_2$, $V(u_1,u_2) = \max\{u_1,u_2\}$ and $V(u_1,u_2) = \min\{u_1,u_2\}$. These three examples satisfy the symmetry conditions in Assumption 3 (see our examples in Section 2.4.1).

S3.1 Choice of testing range, tuning parameters

Let $Z_{(\tau)}$ denote the τ^{th} quantile of the r.v Z. For $\ell = 1,2$, let $\underline{x}^{\ell} \equiv X^{\ell}_{1,(0.01)} \vee X^{\ell}_{2,(0.01)}$ and $\overline{x}^{\ell} \equiv X^{\ell}_{1,(0.99)} \wedge X^{\ell}_{2,(0.99)}$. Our testing range is $\mathcal{X} \equiv [\underline{x}^1, \overline{x}^1] \times [\underline{x}^2, \overline{x}^2]$. Our weight function $\omega(\cdot)$ is the uniform distribution over $\mathcal{X}$. Regarding our tuning parameters, we choose a bandwidth of the form $h_n = c_h \cdot \widehat{\sigma}(X) \cdot n^{-\alpha_h}$, where $\widehat{\sigma}(X)$ is the sample standard deviation of X (we use covariate-specific bandwidths). As we discussed in the paragraph following Assumption E3, we must have $\frac{1}{M+r/2} < \alpha_h < \frac{1}{r(2P-1)}$, and the smallest value of M is $M = \lceil \frac{5r}{2} \rceil$. Since $P = 2$ and $r = 2$, we set $M = 6$ and $\alpha_h = 0.16$. We will present results for $c_h = \{1, 1.25, 1.5, 1.75\}$. We employed a bias-reducing kernel of order $M = 6$ of the form, $\kappa(\psi) = \Big(c_1 \cdot (S^2 - \psi^2)^2 + c_2 \cdot (S^2 - \psi^2)^4 + c_3 \cdot (S2 - \psi^2)^6\Big) \cdot \mathbb{1}\{|\psi| \leq S\}$. The kernel has support $[-S, S]$, with $S = 10$. The coefficients c_1, c_2, c_3 were chosen to satisfy the conditions of a bias-reducing kernel of order $M = 6$.

S3.2 Results

Rejection rates for our test are included in Tables S1 and S2. Each case shown corresponds to 1,000 simulations. We found that, for all the cases of cooperation included, as the sample size increases, the results become increasingly closer to the asymptotic size predictions in Theorem 2. We also find that, consistent with our analysis in Section S1.1

of this supplement, our test has power to reject cooperation when the true behavior is noncooperative (PSNE in this case) and in our experiments we show that this power can be derived from an asymmetry in players' strategic-interaction effects or from the properties of the equilibrium selection mechanism. For the tuning parameters used in our analysis, we found that values of c_h around 1.25 produced the best results for size and power, but our results are robust to a reasonably wide range of values for c_h.

S4 Results for our empirical illustration with alternative bandwidth choices

Table 5 in Section 6 of the paper presents the results of our test for our empirical application, using the bandwidth $h_n = c_h \cdot \widehat{\sigma}(X) \cdot n^{-0.11}$, where $\widehat{\sigma}(X)$ is the sample standard deviation of X (we employed covariate-specific bandwidths). Following our Monte Carlo experiment findings, the results presented in the paper correspond to $c_h = 1.25$. Here we present the results for alternative values of c_h, ranging from $c_h = 1$ to $c_h = 1.75$. As Table S3 shows, our findings are robust, suggesting that, while cooperation in expansion/entry decisions can be rejected in larger markets, this type of behavior cannot be rejected in smaller markets. To be precise, we fail to reject cooperation in markets below the 70th percentile in population size, as well as in markets that did not have any stores in 2008. On the other hand, when we consider all markets or, for example, markets whose size are above the 85th percentile, we reject cooperation at a significance level $< 1\%$.

Table S1: Monte Carlo experiment results. Rejection rates for the null hypothesis of cooperation when $\Delta_1 = -2$, $\Delta_2 = -1$ and PSNE selection mechanism is uniform

	Rejection rates for $c_h = 1$			
	Cooperative behavior			**PSNE behavior**
Sample size	$V(u_1,u_2) = u_1 + u_2$	$V(u_1,u_2) = \max\{u_1,u_2\}$	$V(u_1,u_2) = \min\{u_1,u_2\}$	$P_{\mathcal{M}}(1,0) = 0.50$
$n = 250$	8.9%	8.8%	6.6%	29.4%
$n = 500$	5.8%	6.1%	7.0%	45.0%
$n = 1000$	5.7%	5.7%	5.7%	65.2%

	Rejection rates for $c_h = 1.25$			
	Cooperative behavior			**PSNE behavior**
Sample size	$V(u_1,u_2) = u_1 + u_2$	$V(u_1,u_2) = \max\{u_1,u_2\}$	$V(u_1,u_2) = \min\{u_1,u_2\}$	$P_{\mathcal{M}}(1,0) = 0.50$
$n = 250$	8.2%	8.1%	6.7%	36.8%
$n = 500$	6.2%	6.2%	7.3%	58.3%
$n = 1000$	5.6%	5.5%	4.6%	82.4%

	Rejection rates for $c_h = 1.5$			
	Cooperative behavior			**PSNE behavior**
Sample size	$V(u_1,u_2) = u_1 + u_2$	$V(u_1,u_2) = \max\{u_1,u_2\}$	$V(u_1,u_2) = \min\{u_1,u_2\}$	$P_{\mathcal{M}}(1,0) = 0.50$
$n = 250$	7.6%	7.5%	7.0%	38.8%
$n = 500$	6.6%	6.7%	6.6%	65.3%
$n = 1000$	6.3%	6.2%	5.5%	89.8%

	Rejection rates for $c_h = 1.75$			
	Cooperative behavior			**PSNE behavior**
Sample size	$V(u_1,u_2) = u_1 + u_2$	$V(u_1,u_2) = \max\{u_1,u_2\}$	$V(u_1,u_2) = \min\{u_1,u_2\}$	$P_{\mathcal{M}}(1,0) = 0.50$
$n = 250$	7.2%	7.2%	7.1%	40.2%
$n = 500$	6.9%	6.9%	7.6%	67.4%
$n = 1000$	6.6%	6.4%	5.3%	92.8%

- $P_{\mathcal{M}}(y) \equiv Pr$(mechanism $\mathcal{M}$ will choose PSNE y), with $P_{\mathcal{M}}(1,0) + P_{\mathcal{M}}(0,1) = 1$.
- 1000 simulations in each case.

Table S2: Monte Carlo experiment results. Rejection rates for the null hypothesis of cooperation when $\Delta_1 = -2$, $\Delta_2 = -2$ and PSNE selection mechanism is non-uniform

Sample size	Rejection rates for $c_h = 1$				
	Cooperative behavior			**PSNE behavior**	
	$V(u_1,u_2) = u_1 + u_2$	$V(u_1,u_2) = \max\{u_1,u_2\}$	$V(u_1,u_2) = \min\{u_1,u_2\}$	$P_{\mathcal{M}}(1,0) = 0.25$	$P_{\mathcal{M}}(1,0) = 0.10$
$n = 250$	8.9%	8.9%	6.6%	27.6%	52.1%
$n = 500$	6.1%	6.1%	7.0%	40.3%	74.5%
$n = 1000$	5.7%	5.7%	5.7%	58.8%	93.6%
Sample size	**Rejection rates for $c_h = 1.25$**				
	Cooperative behavior			**PSNE behavior**	
	$V(u_1,u_2) = u_1 + u_2$	$V(u_1,u_2) = \max\{u_1,u_2\}$	$V(u_1,u_2) = \min\{u_1,u_2\}$	$P_{\mathcal{M}}(1,0) = 0.25$	$P_{\mathcal{M}}(1,0) = 0.10$
$n = 250$	8.1%	8.1%	6.7%	30.8%	62.3%
$n = 500$	6.2%	6.2%	7.3%	51.9%	87.0%
$n = 1000$	5.5%	5.5%	4.6%	74.4%	98.9%
Sample size	**Rejection rates for $c_h = 1.5$**				
	Cooperative behavior			**PSNE behavior**	
	$V(u_1,u_2) = u_1 + u_2$	$V(u_1,u_2) = \max\{u_1,u_2\}$	$V(u_1,u_2) = \min\{u_1,u_2\}$	$P_{\mathcal{M}}(1,0) = 0.25$	$P_{\mathcal{M}}(1,0) = 0.10$
$n = 250$	7.5%	7.5%	7.0%	32.9%	65.7%
$n = 500$	6.6%	6.7%	6.6%	55.8%	91.1%
$n = 1000$	6.2%	6.2%	5.5%	82.4%	99.6%
Sample size	**Rejection rates for $c_h = 1.75$**				
	Cooperative behavior			**PSNE behavior**	
	$V(u_1,u_2) = u_1 + u_2$	$V(u_1,u_2) = \max\{u_1,u_2\}$	$V(u_1,u_2) = \min\{u_1,u_2\}$	$P_{\mathcal{M}}(1,0) = 0.25$	$P_{\mathcal{M}}(1,0) = 0.10$
$n = 250$	7.1%	7.2%	7.1%	33.4%	66.9%
$n = 500$	6.9%	6.9%	7.6%	58.1%	92.8%
$n = 1000$	6.4%	6.4%	5.3%	86.0%	99.9%

- $P_{\mathcal{M}}(y) \equiv Pr(\text{mechanism } \mathcal{M} \text{ will choose PSNE } y)$, with $P_{\mathcal{M}}(1,0) + P_{\mathcal{M}}(0,1) = 1$.
- 1000 simulations in each case.

Table S3: Test results for cooperation in expansion decisions.

Results for $c_h = 1$				
All markets	**Markets below the 85^{th} percentile in size**	**Markets below the 70^{th} percentile in size**	**Markets below the 50^{th} percentile in size**	**Markets with no stores in 2008**
$21.650^{\dagger}$ (0.000)	$4.336^{\dagger}$ (0.000)	−1.016 (0.845)	−0.700 (0.758)	−0.620 (0.732)

Results for $c_h = 1.25$				
All markets	**Markets below the 85^{th} percentile in size**	**Markets below the 70^{th} percentile in size**	**Markets below the 50^{th} percentile in size**	**Markets with no stores in 2008**
$21.642^{\dagger}$ (0.000)	$4.294^{\dagger}$ (0.000)	−0.912 (0.819)	−0.773 (0.780)	−0.543 (0.706)

Results for $c_h = 1.5$				
All markets	**Markets below the 85^{th} percentile in size**	**Markets below the 70^{th} percentile in size**	**Markets below the 50^{th} percentile in size**	**Markets with no stores in 2008**
$21.472^{\dagger}$ (0.000)	$4.328^{\dagger}$ (0.000)	−0.845 (0.801)	−0.818 (0.793)	−0.478 (0.684)

Results for $c_h = 1.75$				
All markets	**Markets below the 85^{th} percentile in size**	**Markets below the 70^{th} percentile in size**	**Markets below the 50^{th} percentile in size**	**Markets with no stores in 2008**
$21.348^{\dagger}$ (0.000)	$4.438^{\dagger}$ (0.000)	−0.784 (0.783)	−0.835 (0.798)	−0.449 (0.673)

• Results show the value of our test-statistic, with p-value in parenthesis.

(†) Cooperation rejected at < 1% significance level.

References

Bresnahan, T. F. and P. J. Reiss (1990). Entry in monopoly markets. *Review of Economic Studies 57*, 531–553.

Bresnahan, T. F. and P. J. Reiss (1991). Empirical models of discrete games. *Journal of Econometrics 48*(1-2), 57–81.

Hall, P. (1986). Central limit theorem for integrated square error of multivariate nonparametric density estimators. *Journal of Multivariate Analysis 14*, 1–16.

Serfling, R. J. (1980). *Approximation Theorems of Mathematical Statistics*. Wiley. New York, NY.

Tamer, E. (2003, January). Incomplete simultaneous discrete response model with multiple equilibria. *Review of Economic Studies 70*(1), 147–167.