Simulation Error and Root-n Inference in Dynamic Inequality Estimation: A Comment on Bajari, Benkard, and Levin (2007)

Chae-Yeon Xon
Yonsei University

September 2026

Abstract

Bajari, Benkard, and Levin (2007) derive a root-n distribution for a simulated minimum-distance estimator based on equilibrium inequalities. Their published growth condition permits the number of sampled inequalities to increase too slowly to remove centered forward-simulation error. A bounded one-player dynamic model satisfies S1 and every displayed part of S2. Along $n = m^4$, $n_I = m$, and $n_s = m^3$, its simulation score contributes an $N(0, 2)$ term while the published covariance is zero. This Comment derives a replacement limit indexed by $\kappa = \lim n/(n_I n_s)$. The published covariance applies at $\kappa = 0$; finite positive κ adds simulation covariance; divergent κ changes the convergence rate. Binding inequalities can also generate a root-n mean shift governed by n/n_s. The stronger rates in the paper's working-paper version cover both channels. Under the stated uniformity conditions, the correction leaves the equilibrium inequalities, forward simulator, and point target unchanged.

JEL Codes: C13, C15, C57, L13

Keywords: dynamic games, simulation error, moment inequalities, minimum distance, asymptotic distribution

Yonsei University. eunixon@yonsei.ac.kr. Early stage draft; please do not cite or circulate without author's written permission. All errors are my own.

1. Introduction

Bajari, Benkard, and Levin (2007) propose a two-step estimator for dynamic models of imperfect competition. A first step estimates state transitions and policy rules. A second step simulates continuation values under observed and alternative policies, forms equilibrium inequalities, and minimizes their squared violations. Proposition 1 gives a root-n distribution whose covariance contains first-stage sampling variation. Its simulation condition requires $n_I, n_s \to \infty$ and $n/n_s^2 \to 0$, where n_I is the number of sampled inequalities and n_s is the number of simulated paths per inequality.

The second-stage score averages n_I simulated inequalities. Its centered simulation component has variance of order $(n_I n_s)^{-1}$. At the root-n scale, the governing ratio is $n/(n_I n_s)$. The printed condition on n_s leaves this ratio unrestricted when n_I grows slowly. Section 3 gives a bounded Markov decision problem satisfying the displayed assumptions and the printed rates for which the omitted component converges to $N(0, 2)$.

The same score has a separate centering channel at binding inequalities. The criterion uses the squared negative part, whose derivative has a kink at zero. With a binding atom, simulation noise can produce score bias of order $n_s^{-1/2}$; with a regular slack density, aggregation around the kink yields the smoother order n_s^{-1}. The two channels therefore call for separate rate conditions.

The analysis first locates the two simulation terms in the published estimator. A dynamic counterexample then shows how the omitted term enters, and a replacement expansion separates the resulting three variance regimes. Binding inequalities receive their own treatment before the discussion turns to implementation and to the conclusions that remain in place. The working-paper version of Bajari, Benkard, and Levin (2007) imposed $n/n_I \to r < \infty$ and $n/n_s \to 0$. Those rates imply $n/(n_I n_s) \to 0$ and remove the leading binding-atom mean shift.

2. The simulated inequality estimator

Let X index a state, player, observed policy, and alternative policy. Write the population equilibrium slack as

$$g(X; \theta, \alpha) = V(X; \theta, \alpha) - V'(X; \theta, \alpha), \tag{1}$$

where α collects first-stage objects. Equilibrium gives $g(X; \theta_0, \alpha_0) \geq 0$ for the distribution of inequalities used by the estimator. Let

$$h(u) = \{\min(u, 0)\}^2. \tag{2}$$

For n_I sampled inequalities and n_s forward paths per inequality, the second-stage criterion is

$$Q_n(\theta) = \frac{1}{n_I} \sum_{k=1}^{n_I} h\{\widehat{g}_k(\theta, \widehat{\alpha})\}. \tag{3}$$

The population criterion is identified at θ_0, and the local Hessian in Proposition 1 converges to a nonsingular matrix B_0.

At θ_0, decompose the score into a first-stage term, a mean simulation term, and centered simulation variation. For one sampled inequality define

$$s_{n_s}(X) = \partial_\theta h\{\widehat{g}(X; \theta_0, \alpha_0)\}, \qquad b_{n_s} = \mathbb{E} s_{n_s}(X), \tag{4}$$

and suppose

$$n_s \mathbb{V}\{s_{n_s}(X) - b_{n_s}\} \longrightarrow \Omega_s. \tag{5}$$

The centered average in (3) then has root-n variance of order $n/(n_I n_s)$. This ratio is absent from S2(iii) in the version of record.

3. A bounded dynamic counterexample

Consider one player, discount factor $\beta = 1/2$, parameter space $\Theta = [-1, 1]$, and two decision states s_a indexed by $a \in \{-1, 1\}$. Action zero at s_a pays $a\theta$ and moves to a terminal-payoff state. The terminal payoff Y is uniform on $[-2\sqrt{3}, 2\sqrt{3}]$, is received once, and is followed by absorption. Action one pays zero and leads directly to absorption. The observed policy selects action zero at both decision states; the comparison policy selects action one.

At $\theta_0 = 0$, both actions have value zero, so the observed policy is a selected Markov perfect equilibrium. The exact slacks are

$$g_a(\theta) = a\theta. \tag{6}$$

If the inequality sampler gives each state probability one half, the population objective is

$$Q(\theta) = \frac{1}{2}h(\theta) + \frac{1}{2}h(-\theta) = \frac{\theta^2}{2}, \tag{7}$$

with unique minimizer zero and $B_0 = 1$.

Simulate the value of action zero with n_s independent paths. Its continuation error is

$$e = \frac{1}{n_s} \sum_{r=1}^{n_s} \frac{Y_r}{2}, \tag{8}$$

an average of iid uniform variables on $[-\sqrt{3}, \sqrt{3}]$. Hence $\mathbb{E}e = 0$ and $\mathbb{E}e^2 = 1/n_s$. At the true parameter, one simulated score contribution equals

$$\psi = 2a \min(e, 0). \tag{9}$$

The sampled sign a is independent of e, so

$$\mathbb{E}\psi = 0, \qquad \mathbb{E}\psi^2 = \frac{2}{n_s}. \tag{10}$$

Consequently,

$$\mathbb{V}\left\{ \frac{\sqrt{n}}{n_I} \sum_{k=1}^{n_I} \psi_k \right\} = \frac{2n}{n_I n_s}. \tag{11}$$

Choose $n = m^4$, $n_I = m$, and $n_s = m^3$. Every simulation index diverges and

$$\frac{n}{n_s^2} = \frac{1}{m^2} \longrightarrow 0, \qquad \frac{n}{n_I n_s} = 1. \tag{12}$$

The innovations are bounded, and a triangular-array Lyapunov condition follows because the fourth-moment ratio is of order n_I^{-1}. Uniform convergence of the sample Hessian is given by the two threshold classes in θ. The local first-order condition therefore yields

$$\sqrt{n}(\widehat{\theta} - \theta_0) = -\{1 + o_p(1)\} \frac{\sqrt{n}}{n_I} \sum_{k=1}^{n_I} \psi_k \Longrightarrow N(0, 2). \tag{13}$$

Treating the transition and policy objects as known gives $V_\alpha = 0$. A nondegenerate first stage can be appended through an independent transition component with $\Lambda_0 = 0$. In either version, the covariance displayed in Proposition 1 is zero. Equation (13) gives a different limit under every displayed part of S1 and S2.

4. A replacement root-n expansion

The following proposition separates first-stage, mean-simulation, and centered-simulation terms. It states the high-level conditions used in the expansion so that its role is distinct from a primitive theorem for every dynamic model.

PROPOSITION 1 (Simulation-aware expansion). *Suppose $\widehat{\theta}$ has a uniform local score expansion with nonsingular limit Hessian B_0, $\sqrt{n}(\widehat{\alpha} - \alpha_0) \Rightarrow N(0, V_\alpha)$, and the first-stage derivative converges to Λ_0. Suppose* (5) *holds and the centered simulated scores obey a triangular-array* *central limit theorem, independently of the first-stage sample. If*

$$\frac{n}{n_I n_s} \longrightarrow \kappa \in [0, \infty), \qquad \sqrt{n} b_{n_s} \longrightarrow b, \tag{14}$$

then

$$\sqrt{n}(\widehat{\theta} - \theta_0) \Rightarrow N\left(-B_0^{-1} b,\ B_0^{-1}\{\Lambda_0 V_\alpha \Lambda_0' + \kappa \Omega_s\} B_0^{-1\prime}\right). \tag{15}$$

If $n/(n_I n_s) \to \infty$ and Ω_s is positive in some direction, the root-n sequence is unbounded in probability along that direction under the corresponding central limit approximation.

PROOF. The local score expansion gives

$$\sqrt{n}(\widehat{\theta} - \theta_0) = -B_0^{-1}\left[\Lambda_0 \sqrt{n}(\widehat{\alpha} - \alpha_0) + \sqrt{n} b_{n_s} + \sqrt{\frac{n}{n_I n_s}} Z_{s,n}\right] + o_p(1), \tag{16}$$

where $Z_{s,n} \Rightarrow N(0, \Omega_s)$. Joint convergence and Slutsky's theorem give (15). When the multiplier of a nondegenerate $Z_{s,n}$ diverges, root-n tightness fails in its positive-variance directions. □

With shared randomness between the first-stage estimate and forward simulator, (15) includes the two cross-covariance terms from their joint central limit theorem. The centered simulation variance vanishes under

$$\frac{n}{n_I n_s} \longrightarrow 0. \tag{17}$$

This condition allows n_I and n_s to trade off in controlling variance. Bias has its own rate, addressed next.

5. Binding inequalities

Let $G = g(X; \theta_0, \alpha_0) \geq 0$ be the exact slack and $d = \partial_\theta g(X; \theta_0, \alpha_0)$. At a binding inequality suppose

$$\left(\sqrt{n_s}(\widehat{g} - G), \partial_\theta \widehat{g}\right) \Rightarrow (\zeta, d) \tag{18}$$

conditionally on X, with uniform integrability of the scaled score. Since $h'(u) = 2\min(u, 0)$, the leading score bias from an atom at $G = 0$ is

$$\sqrt{n_s} b_{n_s} \longrightarrow b_A = 2\mathbb{E}\left[\mathbf{1}\{G = 0\} d\, \mathbb{E}\{\min(\zeta, 0) \mid X\}\right]. \tag{19}$$

For conditionally Gaussian noise with standard deviation $\sigma(X)$,

$$b_A = -\sqrt{\frac{2}{\pi}}\mathbb{E}\{\mathbf{1}\{G = 0\}d\sigma(X)\}. \tag{20}$$

Zero-centered root-n inference with $b_A \neq 0$ requires $n/n_s \to 0$. If $n/n_s \to \tau \in (0, \infty)$, Proposition 1 has mean shift $-B_0^{-1}\sqrt{\tau}b_A$.

When the slack distribution has no atom at zero and has a bounded density there, the probability mass crossing the kink is of order $n_s^{-1/2}$ and its conditional score size is of the same order. Their product gives $b_{n_s} = O(n_s^{-1})$ when the simulated derivative is exact or has a root-n_s second-moment bound. The printed condition $n/n_s^2 \to 0$ then removes this smooth-crossing bias. The same order can hold with binding components when their b_A terms cancel. A primitive limit theorem can state the slack density, simulator moments, and cancellation conditions that select between these cases.

LEMMA 1 (A sufficient density condition). *Write $\delta_{n_s} = \widehat{g} - G$ and $\widehat{d}_{n_s} = \partial_\theta \widehat{g}$. Suppose $|d| \le \bar{d}$. Conditional on δ_{n_s}, let G have a density on $[0, \infty)$ bounded by $\bar{f}$, and suppose $\mathbb{E}\delta_{n_s}^2 \le \bar{c}/n_s$ and $\mathbb{E}(\widehat{d}_{n_s} - d)^2 \le \bar{c}_d/n_s$. Then*

$$\left|\mathbb{E}\left[2\widehat{d}_{n_s}\min\{G + \delta_{n_s}, 0\}\right]\right| \le \frac{\bar{d}\bar{f}\bar{c} + 2\sqrt{\bar{c}_d\bar{c}}}{n_s}. \tag{21}$$

PROOF. Decompose $\widehat{d}_{n_s} = d + (\widehat{d}_{n_s} - d)$. Fix $\delta_{n_s} = -a < 0$. Only slacks $G \in [0, a]$ cross zero, and conditional integration for the first term gives

$$\mathbb{E}\left[2|d|\{a - G\}\mathbf{1}\{0 \le G \le a\} \mid \delta_{n_s} = -a\right] \le 2\bar{d}\bar{f}\int_0^a (a - g)\, dg = \bar{d}\bar{f}a^2. \tag{22}$$

The contribution is zero when $\delta_{n_s} \ge 0$. On the crossing event, $|\min\{G + \delta_{n_s}, 0\}| \le |\delta_{n_s}|$. Cauchy–Schwarz therefore bounds the derivative-error term by $2\sqrt{\bar{c}_d\bar{c}}/n_s$. Taking expectations gives (21). □

Lemma 1 gives one set of primitive conditions for the n_s^{-1} row. The density bound is conditional on the simulator disturbance. Other implementations can establish the same crossing bound directly from their transition and payoff simulators.

6. Implementation and scope

The correction changes analytic inference whenever forward paths are reused as part of the second-stage criterion. A simulation budget can be chosen to make both (17) and the applicable bias condition hold. Alternatively, the standard-error calculation can retain $\kappa\Omega_S$ and any finite mean shift. A bootstrap or subsampling procedure can carry simulation variation when it regenerates forward paths within every resample; a procedure that holds those paths fixed generally conditions on their realized error.

TABLE 1. Simulation treatment and repeated-sample implication

Implementation	Implication
Printed analytic covariance under S2(iii)	Omits the $\kappa\Omega_s$ component and any score-bias limit allowed by the displayed rates.
Fresh first stage and fresh forward paths in every resample	Can carry both simulation channels; validity depends on the joint data–simulation rates and the behavior of binding inequalities.
Forward paths held fixed across resamples	Conditions on the realized path error and omits its repeated-sample variation.
Simulation counts chosen with $n/n_s \to 0$ and $n/(n_I n_s) \to 0$	Removes the binding-atom bias and centered simulation component at the root-n scale.

The author-distributed Monte Carlo program redraws value simulations inside its subsampling loop. Its reported finite experiment therefore provides its own resampling calculation. Checked later applications of the BBL architecture use subsampling or bootstrap procedures. Article descriptions frequently leave path regeneration undocumented, so their exposure cannot be assigned from the published prose alone. Their point estimates and counterfactuals remain tied to their respective estimated models.

The equilibrium inequalities in equation (4) of Bajari, Benkard, and Levin (2007), the population penalty criterion, and forward simulation as a method for computing value differences remain in place. The counterexample concerns Proposition 1's root-n law under S2(iii). Proposition 2 uses a set-estimation argument with its own fixed- or growing-inequality target and uniformity requirements.

The stronger rates in Bajari, Benkard, and Levin (2004) give a source-based sufficient repair. Its $n/n_I \to r < \infty$ and $n/n_s \to 0$ imply

$$\frac{n}{n_I n_s} = \frac{n}{n_I}\frac{1}{n_s} \longrightarrow 0 \tag{23}$$

and remove the $n_s^{-1/2}$ binding-atom term at the root-n scale. The version history thus separates the rates needed for the two simulation channels.

7. Conclusion

The simulated inequality score has centered variance of order $(n_I n_s)^{-1}$ and can have bias of order $n_s^{-1/2}$ at binding atoms. Proposition 1's published growth condition controls the smoother n_s^{-1} bias and leaves the centered component unrestricted. The bounded dynamic model makes that component the entire root-n limit. A simulation-aware expansion adds $\kappa\Omega_s$ to the covariance and carries the applicable score-bias limit into the mean. These two quantities provide the rate accounting for analytic and resampling implementations of the estimator.

References

Bajari, Patrick, C. Lanier Benkard, and Jonathan Levin. 2004. "Estimating Dynamic Models of Imperfect Competition." Working Paper 10450, National Bureau of Economic Research. 10.3386/w10450.

Bajari, Patrick, C. Lanier Benkard, and Jonathan Levin. 2007. "Estimating Dynamic Models of Imperfect Competition." *Econometrica* 75 (5): 1331–1370. 10.1111/j.1468-0262.2007.00796.x.

Appendix A. Verification of the counterexample assumptions

The model uses a finite action set, a Markov transition kernel, bounded payoffs, and a pure Markov policy. Degenerate private shocks satisfy the independence condition. The selected equilibrium premise permits the observed action at the tie $\theta_0 = 0$. The inequality sampler is iid on two states. Each simulated value is an average of bounded iid paths, is linear in θ, and is constant in the first-stage parameter. Equation (7) verifies identification. For the Hessian, away from probability-zero knots,

$$B_n(\theta) = \frac{2}{n_I} \sum_{k=1}^{n_I} \mathbf{1}\{a_k\theta + e_k < 0\}. \tag{A1}$$

Symmetry gives $\mathbb{E}B_n(\theta) = F_{n_s}(-\theta) + F_{n_s}(\theta) = 1$. The two threshold subclasses are VC classes, giving $\sup_{\theta\in\Theta} |B_n(\theta) - 1| = o_p(1)$ as $n_I \to \infty$. Integrating this Hessian yields the local expansion used in (13).

Appendix B. Rate conditions across versions

The Econometrica version states S2 and Proposition 1 on pages 1348–1349 and proves the result on pages 1365–1368. Its S2(iii) requires $n_s, n_I \to \infty$ and $n/n_s^2 \to 0$. NBER Working Paper 10450 states the earlier rates $n/n_I \to r < \infty$ and $n/n_s \to 0$. The counterexample uses the version-of-record condition; the working-paper rates appear only as a sufficient comparison in Section 6.