On Sharpness of Bounds for Dynamic Nonseparable Panel Models

V. Chernozhukov, I. Fernandez-Val, J. Hahn, W. Newey

MIT, BU, UCLA, MIT

December 2011

1 Proof of Sharpness

This is a not very elegant proof that the bounds in the dynamic model are sharp. The argument is based on partitioning α in different components and choosing suitable conditional distributions for the components of α that are not restricted given the "bad" sequences $\bar{\mathcal{X}}(x)$. It requires an additional assumption on the structural function g_0.

Since $X_{it} \in \{X^1, ..., X^K\}$, there exists a representation for g_0 such that

$$g_0(x, \alpha_i, \varepsilon_{it}) = \sum_{k=1}^{K} g_k(\alpha_i^k, \varepsilon_{it}^k) 1(x = X^k),$$

for $\alpha_i = (\alpha_i^1, ..., \alpha_i^K)$ and $\varepsilon_{it} = (\varepsilon_{it}^1, ..., \varepsilon_{it}^K)$.

We are interested in the ASF evaluated at X^k. Using the same partition of X_i as in the paper,

$$\mu(X^k) = \sum_{t=1}^{T} E[1\{X_i \in \mathcal{X}_t(X^k)\} g_k(\alpha_i^k, \varepsilon_{it}^k)] + E[g_k(\alpha_i^k, \varepsilon_{iT}^k) \mid X_i \in \bar{\mathcal{X}}(X^k)] \bar{\mathcal{P}}(X^k).$$

To find the upper bound of the ASF, we solve the program

$$\begin{aligned} \max \quad & E[g_k(\alpha_i^k, \varepsilon_{iT}^k) \mid X_i \in \bar{\mathcal{X}}(X^k)] \\ \text{s.t.} \quad & g_{k'}(\alpha_i^{k'}, \varepsilon_{it}^{k'}) \stackrel{d}{=} Y_{it} \mid X_{it} = X^{k'}, X_{i1}, \quad t \in \{1, ..., T\}, k' \in \{1, .., K\}, \end{aligned}$$

where we maximize over the conditional distributions of $(\alpha_i^k, \varepsilon_{iT}^k)$ given $X_i \in \bar{\mathcal{X}}(X^k)$ that satisfy Assumption 3, and the constraints of the program impose Assumption 1. By Assumption 3, the objective function is

$$E[g_k(\alpha_i^k, \varepsilon_{iT}^k) \mid X_i \in \bar{\mathcal{X}}(X^k)] = \int m_k(\alpha_i^k, X_{i1}) dF(\alpha_i^k \mid X_i \in \bar{\mathcal{X}}(X^k)),$$

where $m_k(\alpha_i^k, X_{i1}) = E[g_k(\alpha_i^k, \varepsilon_{iT}^k) \mid \alpha_i^k, X_{i1}]$ is restricted by the constraints since

$$E[m_k(\alpha_i^k, X_{i1}) \mid X_{it} = X^k, X_{i1}] = E[Y_{it} \mid X_{it} = X^k, X_{i1}].$$

However, we can unrectrictively maximize over the conditional distribution of α_i^k given $X_i \in \bar{\mathcal{X}}(X^k)$.

Assume that the following condition holds.

Assumption S *There exists* $\bar{\alpha}_u^k(X_{i1})$ *such that* $m_k(\bar{\alpha}_u^k(X_{i1}), X_{i1}) = B_u$ *for* $X_{i1} \neq X^k$.

The following condition on g_k is sufficient for Assumption S: there exists $\bar{\alpha}_u^k(X_{i1})$ such that $g_k(\bar{\alpha}_u^k(X_{i1}), \varepsilon^k) = B_u$ for all ε^k in the support of the conditional distribution of ε_{i1}^k given (α_i^k, X_{i1}). If g_k does not satisfy this condition, we can define $\tilde{g}_k(\alpha_i^k, \varepsilon_{it}^k) = g_k(\alpha_i^k, \varepsilon_{it}^k)1\{\alpha_i^k \neq \bar{\alpha}_u^k(X_{i1})\} + B_u 1\{\alpha_i^k = \bar{\alpha}_u^k(X_{i1})\}$. Under Assumption S, the solution to the program for the upper bound is $\alpha_i^k = \bar{\alpha}_u^k(X_{i1})$ conditional on $X_i \in \bar{\mathcal{X}}(X^k)$ and X_{i1}.

Note that knowing that $\alpha_i^k = \bar{\alpha}_u^k(X_{i1})$ and therefore $X_i \in \bar{\mathcal{X}}(X^k)$ does not convey any information about ε_{it}^k, $(t = 1, ..., T)$, under Assumption S.

2 Example: Dynamic Binary Choice

Consider the model

$$Y_{it} \mid Y_{i,t-1}, ..., Y_{i0}, \alpha_i \sim i.i.d.\ Be(q_i + (p_i - q_i)Y_{i,t-1}), \quad t = \{1, 2\}, \quad Y_{i0} \mid \alpha_i \sim Be(r_i),$$

where all the random variables are independent across i. The support of $X_i = (Y_{i0}, Y_{i1})$ is $\{0, 1\}^2$. The observed probabilities are:

$$P(Y_{i0} = 1) = E[r_i], \quad P(Y_{i1} = 1 \mid Y_{i0}) = E[q_i \mid Y_{i0}] + E[p_i - q_i \mid Y_{i0}]Y_{i0},$$
$$P(Y_{i2} = 1 \mid Y_{i1}, Y_{i0}) = E[q_i \mid Y_{i1}, Y_{i0}] + E[p_i - q_i \mid Y_{i1}, Y_{i0}]Y_{i1}.$$

Moreover, for $x \in \{0, 1\}$

$$P(X_i \in \mathcal{X}_1(x)) = E[r_i^x (1-r_i)^{1-x}], \; P(X_i \in \mathcal{X}_2(x)) = E[r_i^{1-x}(1-r_i)^x] E[q_i^x (1-p_i)^{1-x} \mid Y_{i0} = 1-x],$$
$$P(X_i \in \bar{\mathcal{X}}(x)) = E[r_i^{1-x}(1-r_i)^x] E[(1-q_i)^x p_i^{1-x} \mid Y_{i0} = 1-x].$$

The ASF of Y is

$$\mu(x) = E[q_i] + E[p_i - q_i]x.$$

For $\tilde{\alpha}_i = (\tilde{\alpha}_i^0, \tilde{\alpha}_i^1)$ and $\varepsilon_{it} = (\varepsilon_{it}^0, \varepsilon_{it}^1)$, construct the nonseparable model

$$g(x, \tilde{\alpha}_i, \varepsilon_{it}) = 1(\varepsilon_{it}^0 < \tilde{\alpha}_i^0)1(x = 0) + 1(\varepsilon_{it}^1 < \tilde{\alpha}_i^1)1(x = 1),$$

where ε_{it}^0 and ε_{it}^1 are i.i.d. $U(0, 1)$ random variables independent of each other, and

$$\tilde{\alpha}_i^0 = \begin{cases} C(0) & \text{if } X_i \in \bar{\mathcal{X}}(0), \\ q_i & \text{otherwise;} \end{cases} \qquad \tilde{\alpha}_i^1 = \begin{cases} C(1) & \text{if } X_i \in \bar{\mathcal{X}}(1), \\ p_i & \text{otherwise.} \end{cases}$$

Let $\tilde{Y}_{it} = g(\tilde{Y}_{i,t-1}, \tilde{\alpha}_i, \varepsilon_{it})$ for $t = \{1, 2\}$, $\tilde{Y}_{i0} = Y_{i0}$, and $\tilde{X}_i = (\tilde{Y}_{i0}, \tilde{Y}_{i1})$. Then, $\tilde{Y}$ is observationally equivalent to Y because

$$P(\tilde{Y}_{i1} = 1 \mid \tilde{Y}_{i0}) = E[1(\varepsilon_{i1}^0 < \alpha_i^0)1(\tilde{Y}_{i0} = 0) + 1(\varepsilon_{i1}^1 < \alpha_i^1)1(\tilde{Y}_{i0} = 1) \mid \tilde{Y}_{i0}] = E[q_i \mid \tilde{Y}_{i0}] + E[p_i - q_i \mid \tilde{Y}_{i0}]\tilde{Y}_{i0},$$

and

$$P(\tilde{Y}_{i2} = 1 \mid \tilde{Y}_{i1}, \tilde{Y}_{i0}) = E[1(\varepsilon_{i2}^0 < \alpha_i^0)1(\tilde{Y}_{i1} = 0) + 1(\varepsilon_{i2}^1 < \alpha_i^1)1(\tilde{Y}_{i1} = 1) \mid \tilde{Y}_{i1}, \tilde{Y}_{i0}]$$
$$= E[q_i \mid \tilde{Y}_{i1}, \tilde{Y}_{i0}] + E[p_i - q_i \mid \tilde{Y}_{i1}, \tilde{Y}_{i0}]\tilde{Y}_{i1}.$$

Moreover,

$$\begin{aligned}
P(\tilde{X}_i \in \mathcal{X}_1(x)) &= P(\tilde{Y}_{i0} = x) = E[r_i^x(1 - r_i)^{1-x}], \\
P(\tilde{X}_i \in \mathcal{X}_2(x)) &= P(\tilde{Y}_{i0} = 1 - x, \tilde{Y}_{i1} = x) = E[r_i^{1-x}(1 - r_i)^x]E[q_i^x(1 - p_i)^{1-x} \mid \tilde{Y}_{i0} = 1 - x], \\
P(\tilde{X}_i \in \bar{\mathcal{X}}(x)) &= P(\tilde{Y}_{i0} = 1 - x, \tilde{Y}_{i1} = 1 - x) = E[r_i^{1-x}(1 - r_i)^x]E[(1 - q_i)^x p_i^{1-x} \mid \tilde{Y}_{i0} = 1 - x],
\end{aligned}$$

The ASF of $\tilde{Y}$ is

$$\tilde{\mu}(x) = \sum_{t=1}^{T} E[g(x, \tilde{\alpha}_i, \varepsilon_{it})1(\tilde{X}_i \in \mathcal{X}_t(x))] + E[g(x, \tilde{\alpha}_i, \varepsilon_{iT})1(\tilde{X}_i \in \bar{\mathcal{X}}(x))]$$
$$= \sum_{t=1}^{T} E[\{q_i + (p_i - q_i)x\}1\{\tilde{X}_i \in \mathcal{X}_t(x)\}] + C(x)P(\tilde{X}_i \in \bar{\mathcal{X}}(x)),$$

and reaches the bounds by setting $C(x) \in \{0, 1\}$