On large market asymptotics for spatial price competition models

Taisuke Otsu and Keita Sunada

January 6, 2022

KEO Discussion Paper No. 165

KEO Discussion Paper No. 165

ON LARGE MARKET ASYMPTOTICS FOR SPATIAL PRICE COMPETITION MODELS

TAISUKE OTSU AND KEITA SUNADA

Abstract. We study the problem of weak instruments in a demand estimation of spatial price competition models by Pinkse, Slade, and Brett (2002) (hereafter, PSB). Product characteristics are included in price instruments and have correlation with prices through the markup. We check whether product characteristics hold their identification power as the number of product grows in analogy with random coefficient discrete choice approach investigated by Armstrong (2016). The conventional weak instruments asymptotics do not work in PSB's model because a series estimation is nested in their two-stage least square estimator, and the number of endogenous regressors (and instruments) also grows as the number of products grows. We provide two asymptotic results that indicate the lack of inconsistency of PSB's estimator.

1. Introduction

Economists utilize instruments to solve the simultaneity problem in demand estimation. Since the influential study by Berry, Levinsohn, and Pakes (1995) (hereafter, BLP) for differentiated product markets, many papers adopt product characteristics as price instruments, which correlate with prices through the markup, especially through the market share of each product. However, Armstrong (2016) shows that the market share of each product disappears fast enough as the number of products grows in some demand models of BLP. Since the market share is a function of the product characteristics, these instruments may lose their identifying power and, as a result, lead to inconsistent estimates. Hence the markup formula plays a role of a drifting sequence as in Staiger and Stock (1997).

This paper studies a weak instrument problem of a demand estimation of spatial price competition models by Pinkse, Slade, and Brett (2002) (hereafter, PSB) in which product characteristics are included in price instruments. In PSB's model, consumers' demands are in a product space, not in a product characteristic space, and they can consume more than one good. Since BLP takes a random coefficient discrete choice approach, the demand model of PSB is considerably different from that of BLP. Even under these differences, however, by rewriting the markup formula induced by the Bertrand equilibrium play, one can see that this formula is a function of the demand function of each product instead of the market share in BLP. Since the market size is finite following the existing literature, we expect that the demand function decays to zero as the number of products grows. Hence the instruments in PSB interact with price in a similar way to BLP.

PSB employ a semiparametric approach, and a series estimation is nested in their estimator. In the just identified case, the estimation error of their two-stage least squares estimator $\hat{\theta} - \theta$ is denoted by

$$\left(\frac{1}{\sqrt{n}}\sum_{i=1}^{n} z_i w_i'\right)^{-1}\left(\frac{1}{\sqrt{n}}\sum_{i=1}^{n} z_i v_i\right) =: A_n^{-1} b_n,$$

where z_i, w_i, v_i are vectors of instruments, regressors containing series expansion terms and exogenous variables, and regression and approximation errors, respectively. Notice that we cannot apply conventional weak instruments asymptotics in Staiger and Stock (1997) since A_n and b_n are growing dimensions. Our first result characterizes the stochastic orders of each element of A_n and b_n. We find that these are not degenerate, and b_n may diverge if the number of expansion terms grows at a slower rate. Our second result provides an inconsistency result of $\hat{\theta}$ given a high-level assumption on the maximum eigenvalue of $A_n' A_n$. For further developments of these results, we need to exploit random matrix theory for sample covariance matrix in which elements are not distributed independently and contain a drifting sequence.

2. Model and estimator

Our model follows that of PSB. There are n sellers of a differentiated product. For simplicity, we assume that each firm sells one product. Let q_i, p_i, and y_i be the demand, price, and product characteristic for product i. The demand function for product i is given by

$$q_i(p, y) = a_i + \sum_{j=1}^{n}(b_{ij} p_j + c_{ij} y_j).$$

where $p = (p_1, \ldots, p_n)'$, $y = (y_1, \ldots, y_n)'$, and $(\{a_i\}, \{b_{ij}\}, \{c_{ij}\})$ are parameters to be estimated. Suppose firms play the Bertrand pricing game given rival prices, i.e., firm i chooses p_i to solve

$$\max_{p_i} (p_i - \gamma MC_i) q_i(p, y) - F_i, \tag{1}$$

where MC_i and F_i are firm i's marginal and fixed cost. The best response function of firm i is

$$p_i = -\frac{1}{2\beta_{ii}}\left(a_i - b_{ii}\gamma MC_i + \sum_{j\neq i} b_{ij} p_j + \sum_{j=1}^{n} c_{ij} y_j\right).$$

PSB estimated this best response function by employing a semiparametric approach. Let x_i be a d_x-vector of MC_i, finite subset of y, and other exogenous demand and cost variables. Also let $\{e_\ell(\cdot)\}_{\ell=1}^{\infty}$ be a sequence of basis functions, $\{d_{ij}\}$ be measures of proximity of firms i and j, and $\tilde{\psi}_{i\ell} = \sum_{j\neq i} e_\ell(d_{ij}) p_j$. Based on this notation, the semiparametric model considered by PSB is written as

$$\begin{aligned} p_i &= \sum_{\ell=1}^{\infty} \tilde{\alpha}_\ell \tilde{\psi}_{i\ell} + x_i'\beta + u_i \qquad (2) \\ &= \psi_i'\alpha + x_i'\beta + v_i, \end{aligned}$$

where $v_i = r_i + u_i$, $r_i = \sum_{\ell=L_n+1}^{\infty} \tilde{\alpha}_\ell \tilde{\psi}_{i\ell}$, $\alpha = (\tilde{\alpha}_1, \ldots, \tilde{\alpha}_{L_n})'$, and

$$\psi_i = \left(\sum_{j \neq i} e_1(d_{ij}) p_j, \sum_{j \neq i} e_2(d_{ij}) p_j, \ldots, \sum_{j \neq i} e_{L_n}(d_{ij}) p_j \right)'.$$

In our setup, the number of endogenous regressors L_n grows with the number of sellers n. Letting $w_i = (\psi_i', x_i')'$ and $\theta = (\alpha', \beta')'$, this model can be concisely written as $p_i = w_i'\theta + v_i$.

For this model, PSB proposed to estimate θ by the (semiparametric) two-stage least squares based on K_n-dimensional vector of instruments z_i. For simplicity, we focus on the just identified case, i.e., $K_n = L_n + d_x$. As in PSB, we adopt transformed variables of x_i as instruments for p_i. Let $z_i = g(x_i)$ be a K_n-dimensional vector-valued function of x_i. Then the semiparametric instrumental variable estimator for θ is written as

$$\hat{\theta} = \left(\sum_{i=1}^{n} z_i w_i' \right)^{-1} \sum_{i=1}^{n} z_i p_i. \tag{3}$$

This paper is concerned with the limiting behaviors of the estimator $\hat{\theta}$ when the number of products n increases to infinity under suitable conditions for the price competition models. To achieve consistency results for $\hat{\theta}$ to θ, it is critical to guarantee sufficiently strong correlations between p_i's contained in the regressors w_i and x_i generating the instruments z_i. To understand the nature of the problem, observe that the first-order condition of (1) can be written as

$$p_i = \gamma MC_i - \frac{q_i(p, y)}{b_{ii}} + u_i. \tag{4}$$

Here MC_i is assumed to be an exogenous regressor included in the regression model (2). Thus we need to guarantee sufficiently strong correlation between the instruments $z_i = g(x_i)$ and markup $q_i(p, y)/b_{ii}$. However, in the current setup, it is common to assume that the market size is finite, i.e., $\lim_{n\to\infty} \sum_{i=1}^{n} q_i(p, y) < \infty$, which implies that $q_i(p, y)$ decays to zero as the number of products n grows. Therefore, the markup $q_i(p, y)/b_{ii}$ may not have enough variations to yield enough correlations with the instruments z_i. This phenomenon is thoroughly studied in Armstrong (2016) for the BLP model on differentiated product demands. Indeed he conjectured may emerge in the current model by PSB (see, p. 1964 of Armstrong, 2016). In the next section, we formally confirm his conjecture.

3. Large market asymptotics

We now study asymptotic properties of the instrumental variables estimator $\hat{\theta}$ under the large market asymptotics, the number of products n diverges to infinity. Based on the existing literature, we impose the following assumptions of the demand function $q_i(p, y)$ and market size.

Assumption Q. *(i)* $\operatorname{p\,lim}_{n\to\infty} \sum_{i=1}^{n} q_i(y, p) < \infty$. *(ii)* $\sqrt{n} \max_{1 \leq i \leq n} q_i(y, p)/b_{ii}^1 \xrightarrow{p} 0$.

Assumption Q (i) says that the market size $\sum_{i=1}^{n} q_i(y, p)$ remains finite as the number of products n diverges to infinity. This assumption implies that the demand $q_i(y, p)$ for each product i decays to 0. Assumption Q (ii) requires that the decay rate of $q_i(y, p)$ normalized by

b_{ii}^1 should be faster than $n^{-1/2}$ uniformly over i. An analogous is employed by Armstrong (2016, Theorem 1) for the BLP model.

We also impose some regularity conditions on the series expansion in (2).

Assumption S. *(i)* $\sup_{1\le i\le n,\ell\in\mathbb{N}}\sum_{j\neq i}|e_\ell(d_{ij})|=O(1)$. *(ii)* $\max_{1\le i\le n}\sum_{j\neq i}e_\ell(d_{ij})^2=O(1)$ *for each* $\ell\in\mathbb{N}$. *(iii)* $\sup_{\ell\in\mathbb{N}}|\tilde{\alpha}_\ell\ell^\lambda|<\infty$ *for some* $\lambda>1$.

Assumptions S (i) and (iii) are also employed by PSB (their assumptions (vi) and (vii), respectively). Assumptions S (i) and (ii) are on the basis functions $\{e_\ell(d)\}_{\ell\in\mathbb{N}}$. When the supports of $\{e_\ell(d)\}_{\ell\in\mathbb{N}}$ are finite, these assumptions require that the number of non-zero elements of $e_\ell(d_{ij})$ for $i,j=1,\dots n$ should be finite. If the supports of $\{e_\ell(d)\}_{\ell\in\mathbb{N}}$ are infinite, Assumptions S (i) and (ii) require that $e_\ell(d)$ should decay fast enough as $d\to\infty$. Assumption S (iii) can be understood as a smoothness condition for the function to be approximated by the series expansion. Intuitively, larger λ is associated with a smoother function.

Based on the above assumptions, we now study the asymptotic properties of the semiparametric instrumental variable estimator $\hat{\theta}$. From (2) and (3), the estimation error of $\hat{\theta}$ can be written as

$$\hat{\theta}-\theta=\left(\frac{1}{\sqrt{n}}\sum_{i=1}^n z_i w_i'\right)^{-1}\left(\frac{1}{\sqrt{n}}\sum_{i=1}^n z_i v_i\right)=:A_n^{-1}b_n. \tag{5}$$

There are two notable features in this expression. First, the matrix A_n is normalized by $n^{-1/2}$, instead of n^{-1} for the case of the conventional instrumental variable regression with strong instruments. This normalization by $n^{-1/2}$ for A_n is employed by Staiger and Stock (1997) for the weak instruments asymptotics. As indicated in the last section, in our setup, the markup $q_i(p,y)/b_{ii}$ (and thus w_i) may not have enough correlations with the instruments z_i, and hence we adopt the analogous normalization. Second, in contrast to the conventional or weak instruments asymptotic analysis in Staiger and Stock (1997), A_n is a $K_n\times K_n$ matrix and b_n is a $K_n\times 1$ vector so that both components have growing dimensions. In other words, we need to deal with the problem of weak instruments for semiparametric or series estimators, where not only the number of instruments K_n but also the number of endogenous regressors L_n grow with the sample size n. Such an analysis is a substantial challenge in the econometrics literature.[1]

Although full development of the asymptotic theory for (5) by extending the random matrix theory is beyond the scope of this paper, we can present two theoretical results to indicate lack of consistency of the estimator $\hat{\theta}$. The first proposition characterizes the stochastic orders of the elements of A_n and b_n.

Proposition 1. *Suppose* $\{p_i,x_i,z_i\}_{i=1}^n$ *is an i.i.d. triangular array, where each element has the finite fourth moments, and Assumptions Q and S hold true. Then each element of* A_n *is of order* $O_p(1)$, *and each element of* b_n *is of order* $O_p(\max\{1,\sqrt{n}L_n^{1-\lambda}\})$.

[1] There are few papers tackle weak instruments in a nonparametric framework despite the problem's importance. Han (2020) analyzes this in a nonparametric estimation model of a triangular system. Freyberger (2017) provides positive testability results for the key identification condition in a nonparametric framework, completeness, through the diameter of an identified set.

This proposition says that the elements in A_n and b_n do not degenerate, and b_n may even diverge when the L_n (and thus K_n) grows at a slower rate. Although this result is not enough to characterize the stochastic order of the whole vector $\hat{\theta} - \theta = A_n^{-1} b_n$, we can observe analogous behaviors of the corresponding terms of A_n and b_n for the case of the weak instruments asymptotics in Staiger and Stock (1997).

Additionally we provide a lack of consistency result in terms of the Euclidean norm $||\hat{\theta} - \theta||$ under some high level assumption on the matrix A_n. Let $\lambda_{\max}(A)$ be the maximum eigenvalue of a matrix A.

Proposition 2. *Suppose $\{p_i, x_i, z_i\}_{i=1}^n$ is an i.i.d. triangular array, where each element has the finite fourth moments, and Assumptions Q and S hold true. If $\lambda_{\max}(A_n A_n') \le C_n$ with probability approaching one (w.p.a.1) and $nL_n^{2-2\lambda}/C_n \to 0$ for some C_n, then $||\hat{\theta} - \theta|| \xrightarrow{p} +\infty$.*

This proposition provides sufficient conditions to induce inconsistency of the estimator $\hat{\theta}$. The additional condition $nL_n^{2-2\lambda}/C_n \to 0$ is analogous to Assumption (viii) in PSB (which requires $nL_n^{2-2\lambda}/\zeta_n \to 0$ for a sequence $\{\zeta_n\}$ associated with the minimum eigenvalue of $\sum_{i=1}^n z_i w_i'$). In our setup, it is beyond the scope of this paper to characterize the upper bound C_n for the maximum eigenvalue of the product matrix $A_n A_n'$ with growing dimension, which requires further developments of the random matrix theory.

To illustrate this point, suppose that A_n is a $K_n \times K_n$ matrix of independent standard normal random variables. Then Johnstone (2001, Theorem 1.1) showed that

$$\frac{\lambda_{\max}(A_n A_n') - \mu_n}{\sigma_n} \xrightarrow{d} \text{Tracy-Widom law of order 1},$$

where $\mu_n = k_n^2$ and $\sigma_n = k_n\{(K_n - 1)^{-1/2} + K_n^{-1/2}\}^{1/3}$ for $k_n = (K_n - 1)^{1/2} + K_n^{1/2}$. Thus, in this case, the upper bound C_n can be set as K_n. By $K_n = L_n + d_x$, the additional condition in Proposition 2 will be $nL_n^{1-2\lambda} \to 0$, which is satisfied when L_n grows fast enough and/or λ is large enough.

Finally, we mention how to test the null hypothesis $H_0 : \theta = \theta_0$. In the conventional weak identification framework, some asymptotically valid test statistics are proposed. For example, the S-statistic in Stock and Wright (2000) takes the form

$$S_{K_n}(\theta) := \frac{1}{\sqrt{n}} \sum_{i=1}^n g_i(\theta)' \left[\frac{1}{n} \sum_{i=1}^n g_i(\theta) g_i(\theta)' \right]^{-1} \frac{1}{\sqrt{n}} \sum_{i=1}^n g_i(\theta),$$

where $g_i(\theta) = z_i(p_i - w_i'\theta)$. If $K_n = K$ is fixed, Stock and Wright (2000, Theorem 2) implies that $S_K(\theta)$ converges in distribution to χ_K^2. Then, applying the central limit theorem yields

$$T_n = \frac{S_{K_n}(\theta) - K_n}{\sqrt{2K_n}} \xrightarrow{d} N(0, 1).$$

Hence normalized $S_{L_n}(\theta)$ can be used for constructing an asymptotically valid hypothesis test. However, we cannot use this statistic if we are interested in a subset of parameters like finite dimensional parameter β in (2), not in the whole parameter θ. Developing a framework for this situation is beyond the scope our paper and left for future research.

Appendix A. Mathematical appendix

A.1. **Proof of Proposition 1.** We first consider the matrix A_n. Without loss of generality, we consider the $(1,1)$-element of A_n, say $A_n^{(1,1)}$. Also we assume $E[z_{1i}] = 0$ to simplify the presentation. By inserting the markup formula in (4), we can decompose

$$\begin{aligned} A_n^{(1,1)} &= \frac{1}{\sqrt{n}} \sum_{i=1}^n z_{1i} \sum_{j \neq i} e_1(d_{ij}) p_j \\ &= \frac{1}{\sqrt{n}} \sum_{i=1}^n \sum_{j \neq i} z_{1i} e_1(d_{ij}) \frac{q_j}{b_{jj}^1} + \frac{1}{\sqrt{n}} \sum_{i=1}^n \sum_{j \neq i} z_{1i} e_1(d_{ij}) MC_j \gamma + \frac{1}{\sqrt{n}} \sum_{i=1}^n \sum_{j \neq i} z_{1i} e_1(d_{ij}) u_j \\ &=: T_1 + T_2 + T_3. \end{aligned}$$

For T_1, Assumptions Q (ii) and S (i) and the law of large numbers imply

$$|T_1| \leq \left\{ \sqrt{n} \max_{1 \leq j \leq n} \frac{q_j}{b_{jj}^1} \right\} \left(\max_{1 \leq i \leq n} \sum_{j \neq i} |e_1(d_{ij})| \right) \frac{1}{n} \sum_{i=1}^n |z_{1i}| = o_p(1).$$

For T_2, observe that

$$\begin{aligned} E[T_2^2] &= \frac{\gamma^2}{n} \sum_{i=1}^n \sum_{i_1=1}^n \sum_{j \neq i} \sum_{j_1 \neq i_1} E[z_{1i} z_{1i_1} MC_j MC_{j_1}] e_1(d_{ij}) e_1(d_{i_1 j_1}) \\ &= \frac{\gamma^2}{n} \sum_{i=1}^n \sum_{j \neq i} \{E[z_{1i} MC_i] E[z_{1j} MC_j] + E[z_{1i}^2] E[MC_j^2]\} e_1(d_{ij})^2 \\ &\leq \frac{C_1}{n} \sum_{i=1}^n \sum_{j \neq i} e_1(d_{ij})^2 \leq C_1 \max_{1 \leq i \leq n} \sum_{j \neq i} e_1(d_{ij})^2 = O(1), \end{aligned}$$

for some $C_1 > 0$, where the inequality follows the assumption that z_{1i} and MC_i have the finite fourth moments, and the last equality follows from Assumption S (ii). Thus, Chebyshev's inequality implies $T_2 = O_p(1)$. For T_3, let $R_i = z_{1i} \sum_{j \neq i} e_1(d_{ij}) u_j$ so that $T_3 = n^{-1/2} \sum_{i=1}^n R_i$. Note that $E[R_i] = 0$,

$$E[R_i^2] = E[z_{1i}^2] \sum_{j \neq i} E[u_j^2] e_1(d_{1j})^2 = E[z_{1i}^2] E[u_i^2] \sum_{j \neq i} e_1(d_{1j})^2,$$

and

$$\begin{aligned} \mathrm{Cov}(R_1, R_2) &= \sum_{j_1 \neq 1} \sum_{j_2 \neq 2} E[z_{11} z_{12} u_{j_1} u_{j_2}] e_1(d_{1j_1}) e_1(d_{2j_2}) \\ &= \sum_{j_2 \neq 2} E[z_{11} z_{12} u_2 u_{j_2}] e_1(d_{12}) e_1(d_{2j_2}) + \sum_{j_1 \neq 1,2} \sum_{j_2 \neq 1,2} E[z_{11} z_{12} u_{j_1} u_{j_2}] e_1(d_{1j_1}) e_1(d_{2j_2}) \\ &= \sum_{j_2 \neq 2} E[z_{11} z_{12} u_2 u_{j_2}] e_1(d_{12}) e_1(d_{2j_2}) + \sum_{j_1 \neq 1,2} \sum_{j_2 \neq 1,2,j_1} E[z_{11} z_{12} u_{j_1} u_{j_2}] e_1(d_{1j_1}) e_1(d_{2j_2}) \\ &\quad + \sum_{j_1 \neq 1,2} E[z_{11} z_{12} u_{j_1}^2] e_1(d_{1j_1}) e_1(d_{1j_2}) \\ &= 0. \end{aligned}$$

Thus, we have

$$\begin{aligned} Var(T_3) &= \frac{1}{n}\sum_{i=1}^{n} Var(R_i) = \frac{1}{n}\sum_{i=1}^{n} E[R_i^2] = \frac{1}{n}\sum_{i=1}^{n} E[z_{1i}^2]E[u_i^2]\sum_{j\neq i} e_1(d_{1j})^2 \\ &\leq C_2 \max_{1\leq i\leq n}\sum_{j\neq i} e_1(d_{1j})^2 = O(1), \end{aligned}$$

for some $C_2 > 0$, where the last equality follows from Assumption S (ii). Now Chebyshev's inequality implies $T_3 = O_p(1)$. Combining these results, we obtain $A_n^{(1,1)} = O_p(1)$.

We next consider the vector b_n. Without loss of generality, we consider the first element of b_n, say

$$b_n^{(1)} = \frac{1}{\sqrt{n}}\sum_{i=1}^{n} z_{1i}u_i + \frac{1}{\sqrt{n}}\sum_{i=1}^{n} z_{1i}r_i =: T_4 + T_5.$$

For T_4, the i.i.d. and finite fourth moments assumptions guarantees

$$E[T_4^2] = \frac{1}{n}\sum_{i=1}^{n}\sum_{j=1}^{n} E[z_{1i}z_{1j}u_iu_j] = \frac{1}{n}\sum_{i=1}^{n} E[z_{1i}^2u_i^2] = O(1).$$

Thus, T_4 is $O_p(1)$. For T_5, note that $E[T_5^2] = \frac{1}{n}\sum_{i=1}^{n}\sum_{j=1}^{n} E[z_{1i}z_{1j}r_ir_j]$ by the i.i.d. assumption, and thus we have

$$\begin{aligned} E[T_5^2] &= \frac{1}{n}\sum_{i=1}^{n}\sum_{j=1}^{n}\sum_{\ell=L_n+1}^{\infty}\sum_{\ell'=L_n+1}^{\infty}\sum_{k\neq i}\sum_{k'\neq j} E[z_{1i}z_{1j}p_kp_{k'}]\tilde{\alpha}_\ell\tilde{\alpha}_{\ell'}e_\ell(d_{ik})e_{\ell'}(d_{jk'}) \\ &\leq \frac{C_3}{n}\sum_{i=1}^{n}\sum_{j=1}^{n}\sum_{\ell=L_n+1}^{\infty}\sum_{\ell'=L_n+1}^{\infty}\sum_{k\neq i}\sum_{k'\neq j} |\tilde{\alpha}_\ell||\tilde{\alpha}_{\ell'}||e_\ell(d_{ik})||e_{\ell'}(d_{jk'})| \\ &\leq C_3 n\left(\sum_{\ell=L_n+1}^{\infty}|\tilde{\alpha}_\ell|\right)^2\left(\sup_{1\leq i\leq n,\ell\in\mathbb{N}}\sum_{k\neq i}|e_\ell(d_{ik})|\right)^2 \\ &\leq C_4 nL_n^{2-2\lambda}\left(\sup_{1\leq i\leq n,\ell\in\mathbb{N}}\sum_{k\neq i}|e_\ell(d_{ik})|\right)^2 = O(nL_n^{2-2\lambda}), \end{aligned} \tag{6}$$

for some $C_3, C_4 > 0$, where the first inequality follows from the Cauchy-Schwarz inequality and finite fourth moments assumption, the third inequality follows from $\sum_{\ell=L_n+1}^{\infty}|\tilde{\alpha}_\ell| \leq \sum_{\ell=L_n+1}^{\infty} C_5\ell^{-\lambda} \leq C_5L_n^{2-2\lambda}$ for some $C_5 > 0$ by using Assumption S (iii), and the last equality follows from Assumption S (i). Thus, Chebyshev's inequality implies $T_5 = O_p(\sqrt{n}L_n^{1-\lambda})$. Combining these results, we obtain $b_n^{(1)} = O_p(\max\{1, \sqrt{n}L_n^{1-\lambda}\})$.

A.2. **Proof of Proposition 2.** Let $\lambda_{\max}(A)$ and $\lambda_{\min}(A)$ be the maximum and minimum eigenvalues of a matrix A, respectively. It is sufficient to show that $\Pr\{(\hat{\theta}-\theta)'(\hat{\theta}-\theta) \leq M\} \to 0$ for each $M > 0$. Take any $M > 0$. Note that

$$(\hat{\theta}-\theta)'(\hat{\theta}-\theta) = b_n'(A_n^{-1})'A_n^{-1}b_n \geq \lambda_{\min}((A_n^{-1})'A_n^{-1})b_n'b_n = \frac{b_n'b_n}{\lambda_{\max}(A_nA_n')},$$

where the last equality follows from

$$\lambda_{\min}((A^{-1})'A^{-1}) = \lambda_{\min}((AA')^{-1}) = \frac{1}{\lambda_{\max}(AA')},$$

for any invertible matrix A. Thus, we have

$$
\begin{aligned}
&\Pr\{(\hat{\theta}-\theta)'(\hat{\theta}-\theta)\le M\} \le \Pr\left\{\frac{b_n' b_n}{\lambda_{\max}(A_n A_n')}\le M\right\} \\
\le\ & \Pr\left\{\frac{b_n' b_n}{\lambda_{\max}(A_n A_n')}\le M,\ \lambda_{\max}(A_n A_n')\le C_n\right\} + \Pr\{\lambda_{\max}(A_n A_n')>C_n\} \\
\le\ & \Pr\{b_n' b_n \le C_n M\} + o(1) \le \frac{E[b_n' b_n]}{C_n M} + o(1),
\end{aligned}
$$

where the third inequality follows from the assumption $\lambda_{\max}(A_n A_n') \le C_n$ w.p.a.1, and the last inequality follows from Markov's inequality. By using the definition $b_n = \frac{1}{\sqrt{n}}\sum_{i=1}^n z_i v_i = \frac{1}{\sqrt{n}}\sum_{i=1}^n z_i(r_i+u_i)$, we can decompose

$$
\begin{aligned}
E[b_n' b_n] &= \frac{1}{n}\sum_{i=1}^n\sum_{j=1}^n E[z_i' z_j r_i r_j] + \frac{1}{n}\sum_{i=1}^n\sum_{j=1}^n E[z_i' z_j u_i u_j] + \frac{2}{n}\sum_{i=1}^n\sum_{j=1}^n E[z_i' z_j r_i u_j] \\
&=: T_1 + T_2 + 2T_3.
\end{aligned}
$$

For T_1, similar arguments to (6) in the proof of Proposition 1 yield

$$
\begin{aligned}
T_1 &= \frac{1}{n}\sum_{i=1}^n\sum_{j=1}^n E\left[z_i' z_j \left(\sum_{\ell=L_n+1}^{\infty}\tilde{\alpha}_\ell \sum_{h\neq i} e_\ell(d_{ih})p_h\right)\left(\sum_{\ell=L_n+1}^{\infty}\tilde{\alpha}_\ell \sum_{k\neq j} e_\ell(d_{jk})p_k\right)\right] \\
&\le O(n)\left(\sum_{\ell=L_n+1}^{\infty}|\tilde{\alpha}_\ell|\right)^2\left(\sup_{1\le i\le n,\ell\in\mathbb{N}}\sum_{j\neq i}|e_\ell(d_{ij})|\right)^2 = O(nL_n^{2-2\lambda}),
\end{aligned}
$$

where the inequality follows from the Cauchy-Schwarz inequality and finite fourth moments assumption, and the second equality follows from Assumptions S (i) and (iii).

For T_2, the i.i.d. assumption and Cauchy-Schwarz inequality imply

$$
T_2 = \frac{1}{n}\sum_{i=1}^n E[z_i' z_i u_i^2] \le \sqrt{E[||z_i||^4]}\sqrt{E[u_i^4]} = O(1).
$$

For T_3, observe that

$$
\begin{aligned}
T_3 &= \frac{1}{n}\sum_{i=1}^n\sum_{j=1}^n\sum_{h\neq i} E[z_i' z_j u_j p_h]\sum_{\ell=L_n+1}^{\infty}\tilde{\alpha}_\ell e_\ell(d_{ih}) \\
&= \frac{1}{n}\sum_{i=1}^n\sum_{j\neq i} E[z_i]' E[z_j u_j p_j]\sum_{\ell=L_n+1}^{\infty}\tilde{\alpha}_\ell e_\ell(d_{ij}) \\
&\le \frac{C_1}{n}\sum_{i=1}^n\sum_{j\neq i}\sum_{\ell=L_n+1}^{\infty}|\tilde{\alpha}_\ell||e_\ell(d_{ij})| \\
&\le C_1\left(\sup_{1\le i\le n,\ell\in\mathbb{N}}\sum_{j\neq i}|e_\ell(d_{ij})|\right)\left(\sum_{\ell=L_n+1}^{\infty}|\tilde{\alpha}_\ell|\right) = O(L_n^{1-\lambda}),
\end{aligned}
$$

for some $C_1 > 0$, where the second equality follows from $E[z_j' z_j u_j] = E[z_j u_j] = 0$, the first inequality follows from the Cauchy-Schwarz inequality and finite fourth moments assumption, and the last equality follows from Assumptions S (i) and (iii).

Combining these results, $E[b_n' b_n] = O(nL_n^{2-2\lambda})$, and thus

$$\Pr\{(\hat{\theta}-\theta)'(\hat{\theta}-\theta) \le M\} \le O(nL_n^{2-2\lambda}/C_n).$$

Therefore, the conclusion follows by the assumption $nL_n^{2-2\lambda}/C_n \to 0$.

References

[1] Armstrong, T. (2016) Large market asymptotics for differentiated product demand estimators with economic models of supply, *Econometrica*, 84, 1961-1980.

[2] Berry, S., Levinsohn, J. and A. Pakes (1995) Automobile prices in market equilibrium, *Econometrica*, 63, 841-890.

[3] Freyberger, J. (2017) On completeness and consistency in nonparametric instrumental variable models, *Econometrica*, 85, 1629-1644.

[4] Han, S. (2020) Nonparametric estimation of triangular simultaneous equations models under weak identification, *Quantitative Economics*, 11, 161-202.

[5] Johnstone, I. M. (2001) On the distribution of the largest eigenvalue in principal components analysis, *Annals of Statistics*, 29, 295-327.

[6] Pinkse, J., Slade, M. E. and C. Brett (2002) Spatial price competition: a semiparametric approach, *Econometrica*, 70, 1111-1153.

[7] Staiger, D. and J. H. Stock (1997) Instrumental variables regression with weak instruments, *Econometrica*, 65, 557-586.

[8] Stock, J. H. and J. H. Wright (2000) GMM with weak identification, *Econometrica*, 68, 1055-1096.

Department of Economics, London School of Economics, Houghton Street, London, WC2A 2AE, UK, and Keio Economic Observatory (KEO), 2-15-45 Mita, Minato-ku, Tokyo 108-8345, Japan.

Email address: t.otsu@lse.ac.uk

Graduate School of Economics, Keio University, 2-15-45 Mita, Minato-ku, Tokyo 108-8345, Japan.

Email address: sunada@keio.jp