# Supplemental Material for "*Testing Many Zero Restrictions in a High Dimensional Linear Regression Setting*"

Jonathan B. Hill*
University of North Carolina, Chapel Hill, U.S.

December 11, 2023

*Dept. of Economics, University of North Carolina, Chapel Hill, North Carolina, U.S.; www.unc.edu/∼jbhill; jbhill@email.unc.edu.

# Contents



# List of Tables

# A Outline

Appendix B presents assumptions. Appendix C contains the omitted proofs of lemmas. In Appendix D we characterize the max-statistic limit law using extreme value theory. Appendix E presents the design set up for robustness checks against the baseline simulation setting in the main paper. Finally, complete simulation results are presented in Appendix F.

# B Assumptions

We assume all random variables exist on a complete measure space. $|x| = \sum_{i,j} |x_{i,j}|$ is the $l_1$-norm, $|x|_2 = (\sum_{i,j} x_{i,j}^2)^{1/2}$ is the Euclidean or $l_2$ norm, and $||A|| = \max_{|\lambda|_2}\{|A\lambda|_2/|\lambda|_2\}$ is the spectral norm for finite dimensional square matrices $A$ (and the Euclidean norm for vectors). $||\cdot||_p$ denotes the $L_p$-norm. $a.s.$ is *almost surely*. $\mathbf{0}_k$ denotes a zero vector with dimension $k \geq 1$. Write $r$-vectors as $x \equiv [x_i]_{i=1}^r$. $[\cdot]$ rounds to the nearest integer. $K > 0$ is non-random and finite, and may take different values in different places. $a_n \propto b_n$ implies $a_n/b_n \to K$. $awp1$ = asymptotically with probability approaching one. We say $z$ has *sub-exponential* distribution tails when $P(|z| > \varepsilon) \leq b\exp\{c\varepsilon\}$ for some $(b,c) > 0$ and all $\varepsilon > 0$ (see, e.g., Vershynin, 2018, Chapt. 2.7).

Throughout $O_p(1)$ and $o_p(1)$ are not functions of model counter $i$. $\{k_{\theta,n}, k_n\}$ are monotonically increasing sequences of positive integers. Define parsimonious parameter spaces $\mathcal{B}_{(i)} \equiv \mathcal{D} \times \Theta_{(i)}$ where $\mathcal{D} \subset \mathbb{R}^{k_\delta}$ and $\Theta_{(i)} \subset \mathbb{R}$ are compact subsets, $\delta_0$ is an interior point of $\mathcal{D}$, and 0 and $\theta_{0,i}$ are interior points of $\Theta_{(i)}$. Recall $\widehat{\mathcal{H}}_{(i)} \equiv 1/n \sum_{t=1}^n x_{(i),t}x_{(i),t}'$ and $\mathcal{H}_{(i)} \equiv E[x_{(i),t}x_{(i),t}']$. Let $\underline{\lambda}_{(i),n}$ and $\underline{\lambda}_{(i)}$ denote the minimum eigenvalues of $\widehat{\mathcal{H}}_{(i)}$ and $\mathcal{H}_{(i)}$

**Assumption 1.**

*a.* $(\epsilon_t, x_t)$ *are iid over* $t$; $E[\epsilon_t] = 0$; $\underline{c} \leq E[\epsilon_t^2] \leq \bar{c}$ *and* $\underline{c} \leq E[x_{j,t}^2] \leq \bar{c}$ *for all* $j$ *and some* $\underline{c},\bar{c} \in (0,\infty)$ *that may differ for different variables;* $E[\epsilon_t^4] < \infty$ *and* $E[x_{j,t}^4] < \infty$ *for all* $j$; $P(E[\epsilon_t^2|x_t] = \sigma^2) = 1$ *for finite* $\sigma^2 > 0$; *and* $\limsup_{n\to\infty} |\theta_0| < \infty$.

*b.* $\beta_0$ *uniquely minimizes* $E[(y_t - \beta' x_t)^2]$ *on* $\mathcal{B}$; $E[(y_t - \beta_{(i)}^{*\prime} x_{(i),t})x_{(i),t}] = \mathbf{0}_{k_\delta+1}$ *for all* $i$ *and unique* $\beta_{(i)}^*$ *in the interior of* $\mathcal{B}_{(i)}$.

*c.* $\liminf_{n\to\infty} \min_{1\leq i\leq k_{\theta,n}} \inf_{\lambda'\lambda=1} E[(\lambda' x_{(i),t})^2] > 0$ *and* $\liminf_{n\to\infty} \min_{1\leq i\leq k_{\theta,n}} \{\underline{\lambda}_{(i)}\} > 0$.

*d.* $\liminf_{n\to\infty} \inf_{\lambda'\lambda=1} \min_{i\in\mathbb{N}}\{\frac{1}{n}\sum_{t=1}^n (\lambda \mathcal{H}_{(i)}^{-1\prime} x_{(i),t})^2\} > 0$ *a.s.*; $\liminf_{n\to\infty} \min_{i\in\mathbb{N}}\{\underline{\lambda}_{(i),n}\} > 0$ *a.s.*.

# C Omitted Proofs

Recall the high level model

$$y_t = \delta_0' x_{\delta,t} + \theta_0' x_{\theta,t} + \epsilon_t = \beta_0' x_t + \epsilon_t$$

and parsimonious versions:

$$y_t = \delta_{(i)}^{*\prime} x_{\delta,t} + \theta_i^* x_{\theta,i,t} + v_{(i),t} = \beta_{(i)}^{*\prime} x_{(i),t} + v_{(i),t}, \;\; i = 1, ..., k_{\theta,n}, \tag{C.1}$$

Recall least squares first and second order terms for (C.1):

$$\widehat{\mathcal{G}}_{(i)} \equiv -\frac{1}{n}\sum_{t=1}^{n} v_{(i),t} x_{(i),t}, \; \widehat{\mathcal{H}}_{(i)} \equiv \frac{1}{n}\sum_{t=1}^{n} x_{(i),t} x_{(i),t}', \;\; \mathcal{H}_{(i)} \equiv E\left[x_{(i),t} x_{(i),t}'\right],$$

and $\hat{\mathcal{Z}}_{(i)} \equiv -\sqrt{n}\mathcal{H}_{(i)}^{-1}\widehat{\mathcal{G}}_{(i)} = \mathcal{H}_{(i)}^{-1} n^{-1/2} \sum_{t=1}^{n} v_{(i),t} x_{(i),t}$. We have.

$$\sqrt{n}\left(\hat{\beta}_{(i)} - \beta_{(i)}^*\right) = -\sqrt{n}\mathcal{H}_{(i)}^{-1}\widehat{\mathcal{G}}_{(i)} - \left\{\widehat{\mathcal{H}}_{(i)}^{-1} - \mathcal{H}_{(i)}^{-1}\right\}\sqrt{n}\widehat{\mathcal{G}}_{(i)} \equiv \hat{\mathcal{Z}}_{(i)} + \hat{\mathcal{R}}_i.$$

Throughout, asymptotics depend on an upper bound for the $l_\infty$ covariate fourth moment:

$$\mathcal{M}_n \equiv E\left[\max_{1 \leq i \leq k_{\theta,n}} \left|x_{(i),t}\right|^4\right].$$

Recall the four covariate cases, in each case uniformly in $i$:

$$\begin{array}{llll} \textbf{\textit{(i)}} & \text{bounded } x_{(i),t} & \textbf{\textit{(ii)}} & \text{sub-Gaussian } \left|x_{(i),t}\right|^4 \\ \textbf{\textit{(iii)}} & \text{sub-exponential } \left|x_{(i),t}\right|^4 & \textbf{\textit{(iv)}} & \mathcal{L}_p\text{-bounded } x_{(i),t},\ p \geq 4 \end{array} \tag{C.2}$$

Write $l_\infty$ norms compactly as

$$|\cdot|_\infty \equiv \max_{1 \leq i \leq k_{\theta,n}} |\cdot| \;\text{ and }\; \|\cdot\|_\infty \equiv \max_{1 \leq i \leq k_{\theta,n}} \|\cdot\|.$$

Notice $||\cdot||_\infty$ differs from $||\cdot||_p \equiv (E|\cdot|^p)^{1/p}$.

## C.1 Lemma 2.2

**Lemma 2.2.** *Let $H_0$ and Assumption 1 hold. Let the weight sequences $\{\mathcal{W}_{n,i}\}_{n\in\mathbb{N}}$ satisfy $\max_{1\leq i\leq k_{\theta,n}}|\mathcal{W}_{n,i}-\mathcal{W}_i\,|=O_p(\sqrt{\ln(k_{\theta,n})\,\mathcal{M}_n/n})$ for non-stochastic $\mathcal{W}_i\in(0,\infty)$. We have:*

$$\max_{1\leq i\leq k_{\theta,n}}\left|[\mathbf{0}_{k_\delta}',1]\hat{\mathcal{R}}_i\right|=O_p\left(\ln(k_{\theta,n})\,\mathcal{M}_n/\sqrt{n}\right). \tag{C.3}$$

*Then $\max_{1\leq i\leq k_{\theta,n}}|[\mathbf{0}_{k_\delta}',1]\hat{\mathcal{R}}_i|\overset{p}{\to}0$ if under covariate case (i) $\ln(k_{\theta,n})=o(\sqrt{n})$; (ii) $\ln(k_{\theta,n})=o(n^{1/3})$; (iii) $\ln(k_{\theta,n})=o(n^{1/4})$; or (iv) $k_{\theta,n}=o((n/\ln(n)^2)^{p/8})$ where $p\geq 4$.*

**Proof** Let $k_\delta=0$ to reduce notation at no cost. Under $H_0$ recall the parsimonious regression errors $v_{(i),t}=\epsilon_t$ by Theorem 2.1.

By multiple uses of the triangle inequality, and $\hat{\mathcal{Z}}_{(i)}\equiv-\sqrt{n}\mathcal{H}_{(i)}^{-1}\widehat{\mathcal{G}}_{(i)}=\mathcal{H}_{(i)}^{-1}n^{-1/2}\sum_{t=1}^n\epsilon_t x_{(i),t}$:

$$\begin{aligned}\left|\left|\sqrt{n}\mathcal{W}_{n,i}\hat{\theta}_i\right|_\infty-\left|\mathcal{W}_i[\mathbf{0}_{k_\delta}',1]\hat{\mathcal{Z}}_{(i)}\right|_\infty\right| &\leq \left|\sqrt{n}\mathcal{W}_{n,i}\hat{\theta}_i-\mathcal{W}_i[\mathbf{0}_{k_\delta}',1]\hat{\mathcal{Z}}_{(i)}\right|_\infty\\ &\leq \max_{i\in\mathbb{N}}|\mathcal{W}_i|\times\left|\sqrt{n}\hat{\theta}_i-[\mathbf{0}_{k_\delta}',1]\hat{\mathcal{Z}}_{(i)}\right|_\infty\\ &\quad+|\mathcal{W}_{n,i}-\mathcal{W}_i|_\infty\times\left|\sqrt{n}\hat{\theta}_i-[\mathbf{0}_{k_\delta}',1]\hat{\mathcal{Z}}_{(i)}\right|_\infty\\ &\quad+\left|[\mathbf{0}_{k_\delta}',1]\hat{\mathcal{Z}}_{(i)}\right|_\infty\times|\mathcal{W}_{n,i}-\mathcal{W}_i|_\infty.\end{aligned}$$

Lemma A.1, arguments in the proof of Lemma A.1.c, and the fact that $|a|=||a||$ for scalar $a$, yield:

$$\begin{aligned}\left|\sqrt{n}\hat{\theta}_i-[\mathbf{0}_{k_\delta}',1]\hat{\mathcal{Z}}_{(i)}\right|_\infty &= \left|[\mathbf{0}_{k_\delta}',1]\left\{\widehat{\mathcal{H}}_{(i)}^{-1}-\mathcal{H}_{(i)}^{-1}\right\}\sqrt{n}\widehat{\mathcal{G}}_{(i)}\right|_\infty\\ &\leq \left\|\widehat{\mathcal{H}}_{(i)}^{-1}-\mathcal{H}_{(i)}^{-1}\right\|_\infty\times\left\|\sqrt{n}\widehat{\mathcal{G}}_{(i)}\right\|_\infty=O_p\left(\ln(k_{\theta,n})\mathcal{M}_n/\sqrt{n}\right).\end{aligned}$$

Similarly

$$\left|[\mathbf{0}_{k_\delta}',1]\hat{\mathcal{Z}}_{(i)}\right|_\infty\leq K\left\|\sqrt{n}\widehat{\mathcal{G}}_{(i)}\right\|_\infty=O_p\left(\sqrt{\ln(k_{\theta,n})\mathcal{M}_n}\right).$$

Further, by assumption $|\mathcal{W}_{n,i}-\mathcal{W}_i|_\infty=O_p(\sqrt{\ln(k_{\theta,n})\mathcal{M}_n/n})$. Combine results to yield the desired result:

$$\begin{aligned}&\left|\left|\sqrt{n}\mathcal{W}_{n,i}\hat{\theta}_i\right|_\infty-\left|\mathcal{W}_i[\mathbf{0}_{k_\delta}',1]\hat{\mathcal{Z}}_{(i)}\right|_\infty\right|\\ &\qquad=O\left(\frac{\ln(k_{\theta,n})\mathcal{M}_n}{\sqrt{n}}\right)+O_p\left(\frac{[\ln(k_{\theta,n})\mathcal{M}_n]^{3/2}}{n}\right)=O_p\left(\frac{\ln(k_{\theta,n})\mathcal{M}_n}{\sqrt{n}}\right).\end{aligned} \tag{C.4}$$

It remains to bound $\mathcal{M}_n$ under (C.2) in order to achieve $\ln(k_{\theta,n})\mathcal{M}_n/\sqrt{n}\to 0$.

$(i)$ If $x_{(i),t}$ is (uniformly in $i$) bounded $a.s.$, then $\mathcal{M}_n\leq\mathcal{C}$. Hence $\ln(k_{\theta,n})=o(\sqrt{n})$.

(*ii*) If $x^4_{(i),t}$ is uniformly sub-Gaussian then for any $\gamma > 0$ and some finite $\mathcal{K} > 0$, by Jensen's inequality and a standard log-exp argument:

$$\begin{aligned}\mathcal{M}_n &\leq \gamma^{-1} \ln \left( k_{\theta,n} \max_{1 \leq i \leq k_{\theta,n}} E\left[\exp\{\gamma |x_{(i),t}|^4\}\right]\right) \\ &\leq \gamma^{-1} \ln \left( k_{\theta,n} \exp\{\gamma^2 \mathcal{K}\}\right) = \gamma^{-1} \ln(k_{\theta,n}) + \gamma \mathcal{K}\end{aligned}$$

Minimize the upper bound over $\gamma$ to yield $\mathcal{M}_n = O(\sqrt{\ln(k_{\theta,n})})$. Now use $\ln(k_{\theta,n})\mathcal{M}_n/\sqrt{n} \to 0$ to yield $\ln(k_{\theta,n}) = o(n^{1/3})$.

(*iii*) If $x^4_{(i),t}$ is uniformly sub-exponential then by the same type of argument $\mathcal{M}_n = O(\ln(k_{\theta,n}))$ (cf. Vershynin, 2018, Proposition 2.7.1). Thus $\ln(k_{\theta,n}) = o(n^{1/4})$.

(*iv*) Finally, if $x_{(i),t}$ is uniformly $\mathcal{L}_p$-bounded, $p \geq 4$, then Liapunov's inequality and $|x_{(i),t}|_\infty \leq \sum_{i=1}^{k_{\theta,n}} |x_{(i),t}|$ yield:

$$\mathcal{M}_n \leq k^{4/p}_{\theta,n} \left( \max_{1 \leq i \leq k_{\theta,n}} E\left|x_{(i),t}\right|^p \right)^{4/p} = O(k^{4/p}_{\theta,n}).$$

We therefore need $\ln(k_{\theta,n})\, k^{4/p}_{\theta,n}/\sqrt{n} \to 0$, where $k_{\theta,n} = o(n^{p/8}/\ln(n)^{\varpi})$ for any $\varpi \geq p/4$ suffices. Hence it suffices to have $k_{\theta,n} = o((n/\ln(n)^2)^{p/8})$. $\mathcal{QED}$.

**Remark 1.** For future reference, the above proof yields by case:

$$\begin{aligned}&i.\ \mathcal{M}_n \leq \mathcal{C}; \text{ and } ii.\ \mathcal{M}_n = O(\sqrt{\ln(k_{\theta,n})}) \\ &iii.\ \mathcal{M}_n = O(\ln(k_{\theta,n})); \text{ and } iv.\ \mathcal{M}_n = O(k^{4/p}_{\theta,n}) \text{ for } p \geq 4.\end{aligned} \tag{C.5}$$

## C.2 Lemma 2.3

Define parsimonious regression residuals

$$\hat{v}_{(i),t} = y_t - \hat{\beta}'_{(i)} x_{(i),t}.$$

Squared standard errors for $\sqrt{n}\hat{\theta}_i$ are

$$\widehat{\mathcal{S}}^2_{(i)} \equiv \left[\widehat{\mathcal{H}}^{-1}_{(i)}\right]_{1,1} \times \hat{\mathcal{V}}^2_{(i),n} \text{ where } \hat{\mathcal{V}}^2_{(i),n} = \frac{1}{n}\sum_{t=1}^{n} \hat{v}^2_{(i),t},$$

and the asymptotic value is $\mathcal{S}^2_{(i)} \equiv [\mathcal{H}^{-1}_{(i)}]_{1,1} \times E[v^2_{(i),t}]$.

**Lemma 2.3.** *Let Assumption 1 hold. Then* $|\widehat{\mathcal{S}}^2_{(i)} - \mathcal{S}^2_{(i)}|_\infty = O_p(\sqrt{\ln(k_{\theta,n})\,\mathcal{M}_n/n})$.

The proof requires a high dimensional max-law of large numbers. We refer to the following result in subsequent proofs as well.

**Lemma C.1** (max-LLN). *Let* $\{y_{(i),t}\}_{t=1}^{n}$ *be zero mean, scalar random variables on a common probability space,* $1 \leq i \leq k$*, iid over* $t$*..Then*

$$\max_{1\leq i\leq k}\left|\frac{1}{n}\sum_{t=1}^{n}y_{(i),t}\right| = O_p\left(\sqrt{\frac{1}{n}8\ln\left(k\right)E\left[\max_{1\leq i\leq k}y_{(i),t}^{2}\right]}\right).$$

**Proof.** Nemirovski's $\mathcal{L}_q$-moment bound, $q \geq 1$, under an iid assumption can be represented as (see, e.g., Bühlmann and Van De Geer, 2011, Lemma 14.24):

$$E\left[\max_{1\leq i\leq k}\left|\frac{1}{n}\sum_{t=1}^{n}y_{(i),t}\right|^{q}\right] \leq \left\{\frac{1}{n}8\ln\left(2k\right)E\left[\max_{1\leq i\leq k}y_{(i),t}^{2}\right]\right\}^{q/2}.$$

Put $q = 1$ to deduce:

$$E\left[\max_{1\leq i\leq k}\left|\frac{1}{n}\sum_{t=1}^{n}y_{(i),t}\right|\right] \leq \sqrt{\frac{1}{n}8\ln\left(k\right)E\left[\max_{1\leq i\leq k}y_{(i),t}^{2}\right]}.$$

The claim now follows from Chebyschev's inequality. $\mathcal{QED}$.

**Proof of Lemma 2.3.** It suffices to prove

$$\Delta_n \equiv \left\|\widehat{\mathcal{H}}_{(i)}^{-1}\hat{\mathcal{V}}_{(i),n}^{2} - \mathcal{H}_{(i)}^{-1}E[v_{(i),t}^{2}]\right\|_{\infty} = O_p\left(\sqrt{\ln\left(k_{\theta,n}\right)\mathcal{M}_n/n}\right).$$

After adding and subtracting like terms, we have:

$$\begin{aligned}\Delta_n \;\leq\; & \left\|\widehat{\mathcal{H}}_{(i)}^{-1} - \mathcal{H}_{(i)}^{-1}\right\|_{\infty} \times E\left[v_{(i),t}^{2}\right] \\ & + \left(\left\|\widehat{\mathcal{H}}_{(i)}^{-1} - \mathcal{H}_{(i)}^{-1}\right\|_{\infty} + \left\|\mathcal{H}_{(i)}^{-1}\right\|_{\infty}\right) \times \left\|\hat{\mathcal{V}}_{(i),n}^{2} - E\left[v_{(i),t}^{2}\right]\right\|_{\infty}.\end{aligned}$$

By Assumption 1.c $||\mathcal{H}_{(i)}^{-1}||_{\infty} = (\min_{1\leq i\leq k_{\theta,n}}\{\underline{\lambda}_{(i)}\})^{-1} = O(1)$. Further, apply Lemma A.1.b to yield:

$$\left\|\widehat{\mathcal{H}}_{(i)}^{-1} - \mathcal{H}_{(i)}^{-1}\right\|_{\infty} = O_p\left(\sqrt{\ln\left(k_{\theta,n}\right)\mathcal{M}_n/n}\right),$$

Next, observe that:

$$\begin{aligned}\left|\hat{\mathcal{V}}_{(i),n}^{2} - E\left[v_{(i),t}^{2}\right]\right| \;\leq\; & \left|\frac{1}{n}\sum_{t=1}^{n}v_{(i),t}^{2} - E\left[v_{(i),t}^{2}\right]\right| + \left\|\hat{\beta}_{(i)} - \beta_{(i)}^{*}\right\|^{2}\left\|\widehat{\mathcal{H}}_{(i)} - \mathcal{H}_{(i)}\right\| \\ & + \left\|\hat{\beta}_{(i)} - \beta_{(i)}^{*}\right\|^{2}\left\|\mathcal{H}_{(i)}\right\| + 2\left\|\hat{\beta}_{(i)} - \beta_{(i)}^{*}\right\|\left\|\frac{1}{n}\sum_{t=1}^{n}x_{(i),t}v_{(i),t}\right\|.\end{aligned} \tag{C.6}$$

We now bound each term uniformly over $i$. Lemma A.1 and its proof imply

$$\left\|\hat{\beta}_{(i)} - \beta^*_{(i)}\right\|_\infty = O_p\left(\sqrt{\ln\left(k_{\theta,n}\right)\mathcal{M}_n/n}\right) \text{ and } \left\|\widehat{\mathcal{H}}_{(i)} - \mathcal{H}_{(i)}\right\|_\infty = O_p\left(\sqrt{\ln\left(k_{\theta,n}\right)\mathcal{M}_n/n}\right).$$

Invoke Lemma A.1.a to deduce

$$\left\|\frac{1}{n}\sum_{t=1}^{n} x_{(i),t}v_{(i),t}\right\|_\infty = O_p\left(\sqrt{\ln\left(k_{\theta,n}\right)\mathcal{M}_n/n}\right).$$

Finally, mimicking the proof of Lemma A.1.a, apply max-LLN Lemma C.1 to $v_{(i),t}$ to yield

$$\left|\frac{1}{n}\sum_{t=1}^{n} v^2_{(i),t} - E[v^2_{(i),t}]\right|_\infty = O_p\left(\sqrt{\ln\left(k_{\theta,n}\right)\mathcal{M}_n/n}\right).$$

The dominant term in (C.6) is therefore $|1/n\sum_{t=1}^{n} v^2_{(i),t} - E[v^2_{(i),t}]|_\infty$. Hence:

$$\left|\hat{\mathcal{V}}^2_{(i),n} - E\left[v^2_{(i),t}\right]\right|_\infty = O_p\left(\sqrt{\ln\left(k_{\theta,n}\right)\mathcal{M}_n/n}\right).$$

Combine the above results with $||\mathcal{H}^{-1}_{(i)}||_\infty < \infty$ to yield $\Delta_n = O_p(\sqrt{\ln\left(k_{\theta,n}\right)\mathcal{M}_n/n})$,proving the claim. $\mathcal{QED}$.

## C.3 Lemma 3.1

Recall $\hat{\beta}^{(0)} \equiv [\hat{\delta}^{(0)\prime}, \mathbf{0}'_{k_\theta}]'$ where $\hat{\delta}^{(0)}$ minimizes $\sum_{t=1}^{n}(y_t - \delta' x_{\delta,t})^2$. Write $\hat{\beta}^{(0)}_{(i)} \equiv [\hat{\delta}^{(0)\prime}, 0]'$. Let $\delta^{(0)}$ minimize $E[(y_t - \delta' x_{\delta,t})^2]$ on compact $\mathcal{D}$, and write $\beta^{(0)} = [\delta^{(0)\prime}, \mathbf{0}'_{k_\theta}]'$ and $\beta^{(0)}_{(i)} = [\delta^{(0)\prime}, 0]'$. Recall

$$y^*_{n,t} \equiv \hat{\delta}^{(0)\prime}x_{\delta,t} + \epsilon^{(0)}_{n,t}\eta_t \text{ and } \epsilon^{(0)}_{n,t} \equiv y_t - \hat{\delta}^{(0)\prime}x_{\delta,t},$$

and note by construction $\hat{\beta}^{(0)\prime}_{(i)}x_{(i),t} = \hat{\delta}^{(0)\prime}x_{\delta,t}$. Recall

$$\tilde{\mathcal{Z}}^{(0)}_{(i)} \equiv \bar{\mathcal{H}}^{(0)-1}_{(i)}\frac{1}{\sqrt{n}}\sum_{t=1}^{n}\eta_t G^{(0)}_{i,t} \tag{C.7}$$

where $G^{(0)}_{i,t} \equiv (y_t - \delta^{(0)\prime}x_{\delta,t})x_{\delta,t}$ and $\bar{\mathcal{H}}^{(0)}_{(i)} \equiv E\left[x_{\delta,t}x'_{\delta,t}\right]$

**Lemma 3.1.** *Let Assumption 1 hold, and let the weights satisfy* $\max_{1\leq i\leq k_{\theta,n}} |\mathcal{W}_{n,i} - \mathcal{W}_i| = O_p(\sqrt{\ln\left(k_{\theta,n}\right)\mathcal{M}_n/n})$ *for non-stochastic* $\mathcal{W}_i \in (0,\infty)$. *Then:*

$$\left|\max_{1\leq i\leq k_{\theta,n}}\left|\sqrt{n}\mathcal{W}_{n,i}\hat{\tilde{\theta}}_i\right| - \max_{1\leq i\leq k_{\theta,n}}\left|\mathcal{W}_i[\mathbf{0}'^{he}_{k_\delta}, 1]\tilde{\mathcal{Z}}^{(0)}_{(i)}\right|\right| = O_p\left(\ln\left(k_{\theta,n}\right)\mathcal{M}^{3/4}_n/\sqrt{n}\right).$$

*Hence the $O_p(\cdot)$ term is $o_p(1)$ if by the ([C.2]) covariate case: (i) $\ln(k_{\theta,n}) = o(n^{1/2})$; (ii) $\ln(k_{\theta,n}) = o(n^{4/11})$; (iii) $\ln(k_{\theta,n}) = o(n^{2/7})$; (iv) $k_{\theta,n} = o((n/[\ln(n)]^2)^{p/6})$, $p \geq 4$.*

The proof of Lemma 3.1 requires the following result.

**Lemma C.2.** *Under Assumption 1 $||\sqrt{n}(\hat{\tilde{\beta}}_{(i)} - \hat{\beta}^{(0)}_{(i)})||_\infty = O_p(\sqrt{\ln(k_{\theta,n})\mathcal{M}_n})$.*

**Proof.** By construction

$$\begin{aligned}\sqrt{n}\left(\hat{\tilde{\beta}}_{(i)} - \hat{\beta}^{(0)}_{(i)}\right) &= \widehat{\mathcal{H}}^{-1}_{(i)}\frac{1}{\sqrt{n}}\sum_{t=1}^{n} x_{(i),t}\epsilon^{(0)}_{n,t}\eta_t \\ &= -\widehat{\mathcal{H}}^{-1}_{(i)}\frac{1}{\sqrt{n}}\sum_{t=1}^{n}\eta_t x_{(i),t}x'_{(i),t}\left(\hat{\beta}^{(0)} - \beta^*_{(i)}\right) + \widehat{\mathcal{H}}^{-1}_{(i)}\frac{1}{\sqrt{n}}\sum_{t=1}^{n}\eta_t x_{(i),t}v_{(i),t}.\end{aligned}$$

Therefore:

$$\begin{aligned}\left\|\sqrt{n}\left(\hat{\tilde{\beta}}_{(i)} - \hat{\beta}^{(0)}_{(i)}\right)\right\|_\infty \leq\ & \left\|\mathcal{H}^{-1}_{(i)}\right\|_\infty\left\|\frac{1}{\sqrt{n}}\sum_{t=1}^{n}\eta_t x_{(i),t}x'_{(i),t}\right\|_\infty\left\|\hat{\beta}^{(0)} - \beta^*_{(i)}\right\|_\infty \\ &+\left\|\widehat{\mathcal{H}}^{-1}_{(i)} - \mathcal{H}^{-1}_{(i)}\right\|_\infty\left\|\frac{1}{\sqrt{n}}\sum_{t=1}^{n}\eta_t x_{(i),t}x'_{(i),t}\right\|_\infty\left\|\hat{\beta}^{(0)} - \beta^*_{(i)}\right\|_\infty \\ &+\left\|\mathcal{H}^{-1}_{(i)}\right\|_\infty\left\|\frac{1}{\sqrt{n}}\sum_{t=1}^{n}\eta_t x_{(i),t}v_{(i),t}\right\|_\infty \\ &+\left\|\widehat{\mathcal{H}}^{-1}_{(i)} - \mathcal{H}^{-1}_{(i)}\right\|_\infty\left\|\frac{1}{\sqrt{n}}\sum_{t=1}^{n}\eta_t x_{(i),t}v_{(i),t}\right\|_\infty.\end{aligned} \tag{C.8}$$

By Lemma A.1.b,

$$\left\|\widehat{\mathcal{H}}^{-1}_{(i)} - \mathcal{H}^{-1}_{(i)}\right\|_\infty = O_p\left(\sqrt{\ln(k_{\theta,n})\mathcal{M}_n/n}\right).$$

By construction of $\hat{\beta}^{(0)}$, Lemma A.1.c, and parameter space compactness:

$$\left\|\hat{\beta}^{(0)} - \beta^*_{(i)}\right\|_\infty \leq \left\|\hat{\delta}^{(0)}\right\| + \left\|\beta^*_{(i)}\right\|_\infty = O_p(1).$$

Note that it is easily checked that $||\hat{\delta}^{(0)}|| \xrightarrow{p} K \in [0,\infty)$; see also the proof of Lemma A.1.c.

It remains to bound $||1/\sqrt{n}\sum_{t=1}^{n}\eta_t x_{(i),t}x'_{(i),t}||_\infty$ in ([C.8]). In order to reduce notion at no cost, assume $k_\delta = 0$ such that $x_{(i),t}$ is a scalar. Recall $\eta_t$ are iid $N(0,1)$ random variables. Apply Nemirovski's inequality, conditional on $x_{(i),t}$ (e.g. Bühlmann and Van De Geer, 2011, Lemma 14.24), and $E[\eta_t^2] = 1$:

$$E\left[\max_{1\le i\le k_{\theta,n}}\left(\frac{1}{\sqrt{n}}\sum_{t=1}^{n}\eta_t x_{(i),t}^2\right)^2 |x_{(i),t}\right] \le 8\ln\left(2k_{\theta,n}\right)E\left[\max_{1\le i\le k_{\theta,n}}\frac{1}{n}\sum_{t=1}^{n}\eta_t^2 x_{(i),t}^4|x_{(i),t}\right] \quad \text{(C.9)}$$

$$= 8\ln\left(2k_{\theta,n}\right)\frac{1}{n}\sum_{t=1}^{n}\max_{1\le i\le k_{\theta,n}} x_{(i),t}^4.$$

Hence unconditionally

$$E\left[\max_{1\le i\le k_{\theta,n}}\left(\frac{1}{\sqrt{n}}\sum_{t=1}^{n}\eta_t x_{(i),t}^2\right)^2\right] \le 8\ln\left(2k_{\theta,n}\right)\mathcal{M}_n.$$

Therefore

$$\max_{1\le i\le k_{\theta,n}}\left|\frac{1}{\sqrt{n}}\sum_{t=1}^{n}\eta_t x_{(i),t}^2\right| = O_p\left(\sqrt{\ln\left(k_{\theta,n}\right)\mathcal{M}_n}\right).$$

Combine bounds to yield as claimed:

$$\begin{aligned}&\left\|\sqrt{n}\left(\hat{\tilde{\beta}}_{(i)} - \hat{\beta}_{(i)}^{(0)}\right)\right\|_{\infty} \\ &\quad = O_p\left(\sqrt{\ln\left(k_{\theta,n}\right)\mathcal{M}_n}\right) + O_p\left(\sqrt{\frac{\ln\left(k_{\theta,n}\right)\mathcal{M}_n}{n}}\right) \times O_p\left(\sqrt{\ln\left(k_{\theta,n}\right)\mathcal{M}_n}\right) \\ &\quad = O_p\left(\sqrt{\ln\left(k_{\theta,n}\right)\mathcal{M}_n}\right). \ \mathcal{QED}.\end{aligned}$$

**Proof of Lemma 3.1.** By the triangle inequality:

$$\begin{aligned}\left|\left|\sqrt{n}\mathcal{W}_{n,i}\hat{\tilde{\theta}}_i\right|_{\infty} - \left|\mathcal{W}_i[\mathbf{0}_{k_\delta}', 1]\tilde{\mathcal{Z}}_{(i)}^{(0)}\right|_{\infty}\right| \le\ & |\mathcal{W}_i|_{\infty}\left|\sqrt{n}\hat{\tilde{\theta}}_i - [\mathbf{0}_{k_\delta}', 1]\tilde{\mathcal{Z}}_{(i)}^{(0)}\right|_{\infty} \\ &+ |\mathcal{W}_{n,i} - \mathcal{W}_i|_{\infty}\left|\sqrt{n}\hat{\tilde{\theta}}_i - [\mathbf{0}_{k_\delta}', 1]\tilde{\mathcal{Z}}_{(i)}^{(0)}\right|_{\infty} \\ &+ |\mathcal{W}_{n,i} - \mathcal{W}_i|_{\infty} \times \left|[\mathbf{0}_{k_\delta}', 1]\tilde{\mathcal{Z}}_{(i)}^{(0)}\right|_{\infty}.\end{aligned} \quad \text{(C.10)}$$

By assumption $|\mathcal{W}_i|_{\infty} \le K$ and $|\mathcal{W}_{n,i} - \mathcal{W}_i|_{\infty} = O_p(\sqrt{\ln\left(k_{\theta,n}\right)\mathcal{M}_n/n})$.

Consider $|[\mathbf{0}_{k_\delta}', 1]\tilde{\mathcal{Z}}_{(i)}^{(0)}|_{\infty}$. Since $|\bar{\mathcal{H}}_{(i)}^{(0)-1}|_{\infty} \le K$ under Assumption 1.c:

$$\left|\tilde{\mathcal{Z}}_{(i)}^{(0)}\right|_{\infty} \le K \times \left|\frac{1}{\sqrt{n}}\sum_{t=1}^{n}\eta_t\left(y_t - \delta^{(0)\prime}x_{\delta,t}\right)x_{(i),t}\right|_{\infty}. \quad \text{(C.11)}$$

By the same conditional Nemirovski's inequality leading to (C.9), and noting $\delta^{(0)\prime}x_{\delta,t} =$

$\beta_{(i)}^{(0)\prime} x_{(i),t}$:

$$\sup_{\zeta'\zeta=1} E\left[\max_{1\leq i\leq k_{\theta,n}}\left(\frac{1}{\sqrt{n}}\sum_{t=1}^{n}\eta_t\left(y_t-\delta^{(0)\prime}x_{\delta,t}\right)\zeta' x_{(i),t}\right)^2\right]$$
$$\leq 8\ln\left(2k_{\theta,n}\right)\times\sup_{\zeta'\zeta=1} E\left[\left\{y_t-\delta^{(0)\prime}x_{\delta,t}\right\}^2\max_{1\leq i\leq k_{\theta,n}}\left(\zeta' x_{(i),t}\right)^2\right]$$
$$\leq 8\ln\left(2k_{\theta,n}\right)\times\sup_{\zeta'\zeta=1}\left(E\left[\left(y_t-\delta^{(0)\prime}x_{\delta,t}\right)^4\right]\right)^{1/2}\left(E\left[\max_{1\leq i\leq k_{\theta,n}}\left(\zeta' x_{(i),t}\right)^4\right]\right)^{1/2}$$
$$\leq K\ln\left(2k_{\theta,n}\right)\times\left(E\left[\left(y_t-\delta^{(0)\prime}x_{\delta,t}\right)^4\right]\right)^{1/2}\left(E\left[\max_{1\leq i\leq k_{\theta,n}}\left|x_{(i),t}\right|^4\right]\right)^{1/2}$$

Adding and subtracting like terms, and using $\limsup_{n\to\infty}|\theta_0|<\infty$, yield for $E[(y_t-\delta^{(0)\prime}x_{\delta,t})^4]$:

$$\left\|y_t-\delta^{(0)\prime}x_{\delta,t}\right\|_4\leq\left\|\beta_0' x_t\right\|_4+\left\|\epsilon_t\right\|_4+\left\|\delta^{(0)\prime}x_{\delta,t}\right\|_4\leq K.$$

Therefore:

$$\sup_{\zeta'\zeta=1} E\left[\max_{1\leq i\leq k_{\theta,n}}\left(\frac{1}{\sqrt{n}}\sum_{t=1}^{n}\eta_t\left(y_t-\delta^{(0)\prime\prime}x_{\delta,t}\right)\zeta' x_{(i),t}\right)^2\right]\leq K\ln\left(2k_{\theta,n}\right)\mathcal{M}_n^{1/2}.$$

Combine that with (C.11) to deduce:

$$\left|\tilde{\mathcal{Z}}_{(i)}^{(0)}\right|_\infty=O_p\left(\sqrt{\ln\left(k_{\theta,n}\right)}\mathcal{M}_n^{1/4}\right).$$

It remains to bound $|\sqrt{n}\widehat{\tilde{\theta}}_i-[\mathbf{0}_{k_\delta}',1]\tilde{\mathcal{Z}}_{(i)}^{(0)}|_\infty$ in (C.10). After adding and subtracting terms:

$$\begin{aligned}\left\|\sqrt{n}\left(\widehat{\tilde{\beta}}_{(i)}-\hat{\beta}_{(i)}^{(0)}\right)-\tilde{\mathcal{Z}}_{(i)}^{(0)}\right\|_\infty \;\leq\; & \left\|\mathcal{H}_{(i)}^{-1}\right\|_\infty\left\|\frac{1}{\sqrt{n}}\sum_{t=1}^{n}\eta_t x_{(i),t}x_{(i),t}'\right\|_\infty\left\|\hat{\beta}_{(i)}^{(0)}-\beta_{(i)}^{(0)}\right\|_\infty\\ &+\left\|\widehat{\mathcal{H}}_{(i)}^{-1}-\mathcal{H}_{(i)}^{-1}\right\|_\infty\left\|\frac{1}{\sqrt{n}}\sum_{t=1}^{n}\eta_t x_{(i),t}x_{(i),t}'\right\|_\infty\left\|\hat{\beta}_{(i)}^{(0)}-\beta_{(i)}^{(0)}\right\|_\infty\\ &+\left\|\widehat{\mathcal{H}}_{(i)}^{-1}-\mathcal{H}_{(i)}^{-1}\right\|_\infty\left\|\tilde{\mathcal{Z}}_{(i)}^{(0)}\right\|_\infty.\end{aligned}$$

By construction $\hat{\beta}_{(i)}^{(0)}-\beta_{(i)}^{(0)}=[\hat{\delta}^{(0)\prime}-\delta^{(0)\prime},0]'$. Hence$||\hat{\beta}_{(i)}^{(0)}-\beta_{(i)}^{(0)}||_\infty=O_p\left(1/\sqrt{n}\right)$ by standard arguments under Assumption 1. Moreover, $||\widehat{\mathcal{H}}_{(i)}^{-1}-\mathcal{H}_{(i)}^{-1}||_\infty=O_p(\sqrt{\ln\left(k_{\theta,n}\right)\mathcal{M}_n/n})$ by

Lemma A.1. Furthermore, by arguments in the proof of Lemma C.2,

$$\left\| \frac{1}{\sqrt{n}} \sum_{t=1}^{n} \eta_t x_{(i),t} x'_{(i),t} \right\|_\infty = O_p\left(\sqrt{\ln\left(k_{\theta,n}\right) \mathcal{M}_n}\right).$$

Therefore:

$$\begin{aligned}
&\left\| \sqrt{n}\left(\hat{\tilde{\beta}}_{(i)} - \hat{\beta}_{(i)}^{(0)}\right) - \tilde{\mathcal{Z}}_{(i)}^{(0)} \right\|_\infty \\
&\quad \leq O_p\left(\sqrt{\ln\left(k_{\theta,n}\right) \mathcal{M}_n}\right) \times O_p\left(1/\sqrt{n}\right) + O_p\left(\sqrt{\ln\left(k_{\theta,n}\right) \mathcal{M}_n/n}\right) \times O_p\left(\sqrt{\ln\left(k_{\theta,n}\right)} \mathcal{M}_n^{1/4}\right) \\
&\qquad + O_p\left(\sqrt{\ln\left(k_{\theta,n}\right) \mathcal{M}_n/n}\right) \times O_p\left(\sqrt{\ln\left(k_{\theta,n}\right) \mathcal{M}_n}\right) \times O_p\left(1/\sqrt{n}\right) \\
&\quad = O_p\left(\sqrt{\ln\left(k_{\theta,n}\right) \mathcal{M}_n/n}\right) + O_p\left(\ln\left(k_{\theta,n}\right) \mathcal{M}_n^{3/4}/n^{1/2}\right) + O_p\left(\ln\left(k_{\theta,n}\right) \mathcal{M}_n/n\right) \\
&\quad = O_p\left(\ln\left(k_{\theta,n}\right) \mathcal{M}_n^{3/4}/n^{1/2}\right).
\end{aligned}$$

Combine bounds to yield as claimed:

$$\begin{aligned}
&\left| \max_{1 \leq i \leq k_{\theta,n}} \left| \sqrt{n} \mathcal{W}_{n,i} \hat{\tilde{\theta}}_i \right| - \max_{1 \leq i \leq k_{\theta,n}} \left| \mathcal{W}_i [\mathbf{0}'_{k_\delta}, 1] \tilde{\mathcal{Z}}_{(i)}^{(0)} \right| \right| \\
&\qquad = O_p\left(\ln\left(k_{\theta,n}\right) \mathcal{M}_n^{3/4}/n^{1/2}\right) \\
&\qquad\quad + O_p\left(\sqrt{\ln\left(k_{\theta,n}\right) \mathcal{M}_n/n}\right) \times O_p\left(\ln\left(k_{\theta,n}\right) \mathcal{M}_n^{3/4}/n^{1/2}\right) \\
&\qquad\quad + O_p\left(\sqrt{\ln\left(k_{\theta,n}\right) \mathcal{M}_n/n}\right) \times O_p\left(\sqrt{\ln\left(k_{\theta,n}\right)} \mathcal{M}_n^{1/4}\right) \\
&\qquad = O_p\left(\ln\left(k_{\theta,n}\right) \mathcal{M}_n^{3/4}/n^{1/2}\right).
\end{aligned}$$

Now use the bounds on $\mathcal{M}_n$ from (C.5) to conclude $\ln\left(k_{\theta,n}\right) \mathcal{M}_n^{3/4}/n^{1/2} \to 0$ by case if $(i)$ $\mathcal{M}_n \leq \mathcal{C}$ hence $\ln(k_{\theta,n}) = o(n^{1/2})$; $(ii)$ $\mathcal{M}_n = O(\sqrt{\ln(k_{\theta,n})})$ hence $\ln(k_{\theta,n}) = o(n^{4/11})$; $(iii)$ $\mathcal{M}_n = O(\ln(k_{\theta,n}))$ hence $\ln(k_{\theta,n}) = o(n^{2/7})$; $(iv)$ $\mathcal{M}_n = O(k_{\theta,n}^{4/p})$ for $p \geq 4$ hence $k_{\theta,n} = o((n/[\ln(n)]^2)^{p/6})$ suffices. $\mathcal{QED}$.

## C.4 Lemma 3.2

Recall $\bar{\mathcal{H}}_{(i)}^{(0)}$ and $\tilde{\mathcal{Z}}_{(i)}^{(0)}$ in (C.7), and define:

$$\tilde{\sigma}_{(i)}^2(\lambda) \equiv \lambda' \left(\bar{\mathcal{H}}_{(i)}^{(0)}\right)^{-1} E\left[\left(y_t - \delta^{(0)\prime} x_{\delta,t}\right)^2 x_{\delta,t} x'_{\delta,t}\right] \left(\bar{\mathcal{H}}_{(i)}^{(0)}\right)^{-1} \lambda. \tag{C.12}$$

Recall $\eta_t$ is iid $N(0,1)$ and independent of the sample $\mathfrak{S}_n \equiv \{x_t, y_t\}_{t=1}^n$. $\Rightarrow^p$ denotes *weak convergence in probability* (Giné and Zinn, 1990: Section 3).

**Lemma 3.2.** *Let Assumption 1 hold. Let $\{\tilde{\boldsymbol{Z}}_{(i)}(\lambda)\}_{i\in\mathbb{N}}$, $\tilde{\boldsymbol{Z}}_{(i)}(\lambda) \sim N(0, \tilde{\sigma}_{(i)}^2(\lambda))$, be an independent copy of the Theorem 2.5.a null distribution process $\{\boldsymbol{Z}_{(i)}(\lambda)\}_{i\in\mathbb{N}}$, that is independent of the asymptotic draw $\{x_t, y_t\}_{t=1}^{\infty}$. Let $\{k_{\theta,n}\}$ satisfy by the (C.2) covariate cases (i)-(iii) $\ln(k_{\theta,n}) = o(n^{1/3})$, and (iv) $k_{\theta,n} = o((n/\ln(n))^{p/4})$. Then*

$$\sup_{z\geq 0}\left|P\left(\max_{1\leq i\leq k_{\theta,n}}\left|\lambda'\tilde{\mathcal{Z}}_{(i)}^{(0)}\right| \leq z|\mathfrak{S}_n\right) - P\left(\max_{1\leq i\leq k_{\theta,n}}\left|\tilde{\boldsymbol{Z}}_{(i)}(\lambda)\right| \leq z\right)\right| \xrightarrow{p} 0.$$

*Furthermore*

$$\max_{1\leq i\leq k_{\theta,n}}\left|\left[\mathbf{0}_{k_\delta}', 1\right]\tilde{\mathcal{Z}}_{(i)}^{(0)}\right| \Rightarrow^p \max_{i\in\mathbb{N}}\left|\tilde{\boldsymbol{Z}}_{(i)}([\mathbf{0}_{k_\delta}', 1])\right|.$$

**Proof.**

**Step 1.** Recall $\tilde{\mathcal{Z}}_{(i)}^{(0)} = -\bar{\mathcal{H}}_{(i)}^{(0)-1} n^{-1/2}\sum_{t=1}^n \eta_t G_{i,t}^{(0)}$ where $G_{i,t}^{(0)} \equiv (y_t - \delta^{(0)\prime}x_{\delta,t})x_{\delta,t}$ and $\bar{\mathcal{H}}_{(i)}^{(0)} \equiv E[x_{\delta,t}x_{\delta,t}']$. As always we only care about $\lambda = [\mathbf{0}_{k_\delta}', 1]'$, hence we may assume $||\lambda|| = 1$. Write

$$\hat{\sigma}_{n,(ij)}^2(\lambda) \equiv \lambda' E\left[\tilde{\mathcal{Z}}_{(i)}^{(0)}\tilde{\mathcal{Z}}_{(j)}^{(0)}|\mathfrak{S}_n\right]\lambda = \lambda'\bar{\mathcal{H}}_{(i)}^{(0)-1}\frac{1}{n}\sum_{t=1}^n \eta_t^2 G_{i,t}^{(0)}G_{j,t}^{(0)\prime}\bar{\mathcal{H}}_{(i)}^{(0)-1}\lambda$$
$$\tilde{\sigma}_{(ij)}^2(\lambda) \equiv \lambda'\bar{\mathcal{H}}_{(i)}^{(0)-1}E\left[G_{i,t}^{(0)}G_{j,t}^{(0)\prime}\right]\bar{\mathcal{H}}_{(j)}^{(0)-1}\lambda.$$

By construction $\lambda'\tilde{\mathcal{Z}}_{(i)}^{(0)}|\mathfrak{S}_n \sim N(0, \hat{\sigma}_{n,(ii)}^2(\lambda))$. Recall $\{\tilde{\boldsymbol{Z}}_{(i)}(\lambda)\}_{n\in\mathbb{N}}$ are sequences with $\tilde{\boldsymbol{Z}}_{(i)}(\lambda) \sim N(0, \tilde{\sigma}_{(ii)}^2(\lambda))$ independent of $\mathfrak{S}_n$. Define

$$\Delta_{k_{\theta,n}}(\lambda) \equiv \max_{1\leq i,j\leq k_{\theta,n}}\left|\hat{\sigma}_{n,(ij)}^2(\lambda) - \tilde{\sigma}_{(ij)}^2(\lambda)\right|.$$

Lemma 3.1 in Chernozhukov, Chetverikov, and Kato (2013) therefore yields the following conditional Slepian-type inequality (cf. Chernozhukov, Chetverikov, and Kato, 2015, Theorem 2, Proposition 1):

$$\begin{aligned}\mathcal{E}_{k_{\theta,n}} &\equiv \sup_{z\geq 0}\left|P\left(\max_{1\leq i\leq k_{\theta,n}}\left|[\mathbf{0}_{k_\delta}', 1]\tilde{\mathcal{Z}}_{(i)}^{(0)}\right| \leq z|\mathfrak{S}_n\right) - P\left(\max_{1\leq i\leq k_{\theta,n}}\left|\tilde{\boldsymbol{Z}}_{(i)}([\mathbf{0}_{k_\delta}', 1]')\right| \leq z\right)\right| \\ &= O_p\left(\Delta_{k_{\theta,n}}^{1/3}(\lambda)\max\left\{1, \ln\left(k_{\theta,n}/\Delta_{k_{\theta,n}}(\lambda)\right)\right\}^{2/3}\right).\end{aligned}$$

The proof is therefore complete if we show $\mathcal{E}_{k_{\theta,n}} = o_p(1)$.

We first need to bound $\Delta_{k_{\theta,n}}(\lambda)$. Add and subtract like terms to yield from the triangle inequality:

$$\Delta_{k_{\theta,n}}(\lambda) \leq \max_{i \in \mathbb{N}} \left\| \bar{\mathcal{H}}_{(i)}^{-1} \right\|^2 \times \max_{1 \leq i,j \leq k_{\theta,n}} \left| \frac{1}{n} \sum_{t=1}^{n} \left( \eta_t^2 G_{i,t}^{(0)} G_{j,t}^{(0)\prime} - E\left[ G_{i,t}^{(0)} G_{j,t}^{(0)\prime} \right] \right) \right|.$$

Under Assumption 1 $\max_{i \in \mathbb{N}} ||\bar{\mathcal{H}}_{(i)}^{-1}|| < \infty$, and $E[\eta_t^2 G_{i,t}^{(0)} G_{j,t}^{(0)\prime}] = E[G_{i,t}^{(0)} G_{j,t}^{(0)\prime}]$ by mutual independence and $\eta_t \sim N(0,1)$. Now combine Lemma C.1 with mutual independence, Gaussianicity, and the fact that $G_t^{(0)} \equiv G_{i,t}^{(0)} = (y_t - \delta^{(0)\prime} x_{\delta,t}) x_{\delta,t}$ does not depend on $i$, to deduce:

$$\begin{aligned}
\max_{1 \leq i \leq k_{\theta,n}} \left| \frac{1}{n} \sum_{t=1}^{n} \eta_t^2 \sup_{\omega'\omega=1} \left( \omega' G_{i,t}^{(0)} G_{j,t}^{(0)\prime} \omega \right) \right| &= O_p\left( \sqrt{\frac{1}{n} 8 \ln(k_{\theta,n}) E\left[ \sup_{\omega'\omega=1} \left( \omega' G_t^{(0)} \right)^2 \right]} \right) \\
&= O_p\left( \sqrt{\frac{1}{n} \ln(k_{\theta,n}) E\left\| G_t^{(0)} \right\|^2} \right)
\end{aligned}$$

Now, in view of $\limsup_{n \to \infty} |\theta_0| < \infty$:

$$\left\| G_t^{(0)} \right\| \leq \left| y_t - \delta^{(0)\prime} x_{\delta,t} \right| \|x_{\delta,t}\| \leq |\epsilon_t| \|x_{\delta,t}\| + K \left\| x_{\delta,t} x_{\delta,t}' \right\| + K \left\| x_{\delta,t} x_{\theta,t}' \right\|.$$

Hence, by applications of Minkowski's inequality, and Assumption 1.a:

$$E\left\| G_t^{(0)} \right\|^2 \leq E\left( |\epsilon_t| \|x_{\delta,t}\| + K \left\| x_{\delta,t} x_{\delta,t}' \right\| + K \left\| x_{\delta,t} x_{\theta,t}' \right\| \right)^2 \leq K \mathcal{M}_4.$$

Therefore

$$\max_{1 \leq i \leq k_{\theta,n}} \left| \frac{1}{n} \sum_{t=1}^{n} \eta_t^2 \sup_{\omega'\omega=1} \left( \omega' G_{i,t}^{(0)} G_{j,t}^{(0)\prime} \omega \right) \right| = O_p\left( \sqrt{\frac{\ln(k_{\theta,n}) \mathcal{M}_4}{n}} \right),$$

yielding

$$\Delta_{k_{\theta,n}}(\lambda) = O_p\left( \sqrt{\frac{\ln(k_{\theta,n}) \mathcal{M}_4}{n}} \right).$$

Now turn to $\mathcal{E}_{k_{\theta,n}}$ and note the presence of $\max\{1, \ln\left(k_{\theta,n}/\Delta_{k_{\theta,n}}(\lambda)\right)\}$. Using the $\mathcal{M}_4$ bounds in (C.5) by the covariate cases (C.2), it follows:

$i.$ $\dfrac{\Delta_{k_{\theta,n}}(\lambda)}{k_{\theta,n}} = O_p\left( \dfrac{1}{k_{\theta,n}} \sqrt{\dfrac{\ln(k_{\theta,n})}{n}} \right) = o_p(1)$ if $\ln(k_{\theta,n}) = o(n)$;

$ii.$ $\dfrac{\Delta_{k_{\theta,n}}(\lambda)}{k_{\theta,n}} = O_p\left( \dfrac{\ln(k_{\theta,n})}{k_{\theta,n}\sqrt{n}} \right) = o_p(1)$ if $\ln(k_{\theta,n}) = o(n^{1/2})$;

$iii.$ $\dfrac{\Delta_{k_{\theta,n}}(\lambda)}{k_{\theta,n}} = O_p\left( \dfrac{(\ln(k_{\theta,n}))^{3/2}}{k_{\theta,n}\sqrt{n}} \right) = o_p(1)$ if $\ln(k_{\theta,n}) = o(n^{1/3})$;

$$iv. \ \frac{\Delta_{k_{\theta,n}}(\lambda)}{k_{\theta,n}} = O_p\left(\frac{\sqrt{\ln(k_{\theta,n})}}{k_{\theta,n}^{1-4/p}\sqrt{n}}\right) = o_p(1) \text{ if } p \geq 4 \text{ and } \ln(k_{\theta,n}) = o(n).$$

Setting $\ln(k_{\theta,n}) = o(n^{1/3})$, the least upper bound from above, it follows

$$\begin{aligned}\mathcal{E}_{k_{\theta,n}} &= O_p\left(\Delta_{k_{\theta,n}}^{1/3}(\lambda)\max\left\{1, \ln\left(k_{\theta,n}/\Delta_{k_{\theta,n}}(\lambda)\right)\right\}^{2/3}\right) = O_p\left(\Delta_{k_{\theta,n}}^{1/3}(\lambda)\right) \\ &= O_p\left(\left(\frac{\ln(k_{\theta,n})\mathcal{M}_4}{n}\right)^{1/6}\right).\end{aligned}$$

Now use (C.5) to conclude by case $\mathcal{E}_{k_{\theta,n}} \stackrel{p}{\to} 0$ if $\ln(k_{\theta,n}) = o(n^{1/3})$ and $(i)$ $\ln(k_{\theta,n}) = o(n)$; $(ii)$ $\ln(k_{\theta,n}) = o(n^{2/3})$; $(iii)$ $\ln(k_{\theta,n}) = o(n^{1/2})$; and $(iv)$ $k_{\theta,n} = o(n^{p/4}/\ln(n)^{\omega})$ for any $\omega \geq p/4$, thus $k_{\theta,n} = o((n/\ln(n))^{p/4})$ suffices. In summary, by case: $(i)$-$(iii)$ $\ln(k_{\theta,n}) = o(n^{1/3})$, and $(iv)$ $k_{\theta,n} = o((n/\ln(n))^{p/4})$.

**Step 2.** Step 1 yields $\max_{1\leq i\leq k_{\theta,n}}|[\mathbf{0}'_{k_\delta}, 1]\tilde{\mathcal{Z}}_{(i)}^{(0)}|\ |\mathfrak{S}_n \stackrel{d}{\to} \max_{i\in\mathbb{N}}|\tilde{\boldsymbol{Z}}_{(i)}([\mathbf{0}'_{k_\delta}, 1])|$ $awp1$ with respect to $\{x_t, y_t\}_{t=1}^{\infty}$. Therefore $\max_{1\leq i\leq k_{\theta,n}}|[\mathbf{0}'_{k_\delta}, 1]\tilde{\mathcal{Z}}_{(i)}^{(0)}| \Rightarrow^p \max_{i\in\mathbb{N}}|\tilde{\boldsymbol{Z}}_{(i)}([\mathbf{0}'_{k_\delta}, 1])|$; see. Gine and Zinn (1990, Section 3). $\mathcal{QED}$.

## C.5 Lemma A.1

**Lemma A.1.** *Let Assumption 1 hold, and let $\{k_{\theta,n}\}$ be any monotonically increasing sequence of integers. Then:*

*a.* $||\sqrt{n}\widehat{\mathcal{G}}_{(i)}||_\infty = O_p(\sqrt{\ln(k_{\theta,n})\mathcal{M}_n})$.

*b.* $||\widehat{\mathcal{H}}_{(i)}^{-1} - \mathcal{H}_{(i)}^{-1}||_\infty = O_p(\sqrt{\ln(k_{\theta,n})\mathcal{M}_n/n})$.

*c.* $||\sqrt{n}(\hat{\beta}_{(i)} - \beta_{(i)}^*)||_\infty = O_p(\sqrt{\ln(k_{\theta,n})\mathcal{M}_n})$ *provided* $\ln(k_{\theta,n}) = o(\sqrt{n}/\mathcal{M}_n)$.

**Proof.**

**Claim (a).** Apply Lemma C.1 element-wise to $v_{(i),t}x_{(i),t}$, and invoke Young's inequality, to yield (dropping superfluous multiplicative constants):

$$\begin{aligned}\left\|\sqrt{n}\widehat{\mathcal{G}}_{(i)}\right\|_\infty &= O_p\left(\sqrt{\ln(k_{\theta,n}) \times E\left[\max_{1\leq i\leq k_{\theta,n}}\left|v_{(i),t}x_{(i),t}\right|^2\right]}\right) \\ &= O_p\left(\sqrt{\ln(k_{\theta,n}) \times E\left[\max_{1\leq i\leq k_{\theta,n}} v_{(i),t}^4 + \max_{1\leq i\leq k_{\theta,n}}\left|x_{(i),t}\right|^4\right]}\right) \\ &= O_p\left(\sqrt{\ln(k_{\theta,n}) \times \left\{E\left[\max_{1\leq i\leq k_{\theta,n}} v_{(i),t}^4\right] + \mathcal{M}_n\right\}}\right)\end{aligned}$$

By Minkowski's equality and Assumption 1.a:

$$\begin{aligned}\left\|\max_{1\leq i\leq k_{\theta,n}}\left|v_{(i),t}\right|\right\|_4 &\leq \|\beta_0' x_t\|_4 + \left\|\max_{1\leq i\leq k_{\theta,n}}\left|\beta_{(i)}^{*\prime} x_{(i),t}\right|\right\|_4 + \|\epsilon_t\|_4 \\ &\leq K + \max_{1\leq j\leq k_{\theta,n}+k_\delta}\|x_{j,t}\|_4 \times |\beta_0| + K\mathcal{M}_n^{1/4} = O\left(\mathcal{M}_n^{1/4}\right).\end{aligned}$$

The claim follows instantly.

**Claim (b).** Observe that

$$\left\|\widehat{\mathcal{H}}_{(i)}^{-1} - \mathcal{H}_{(i)}^{-1}\right\|_\infty \leq \left\|\widehat{\mathcal{H}}_{(i)}^{-1}\right\|_\infty \times \left\|\mathcal{H}_{(i)}^{-1}\right\|_\infty \times \left\|\widehat{\mathcal{H}}_{(i)} - \mathcal{H}_{(i)}\right\|_\infty.$$

Under Assumption 1.b,c

$$\begin{aligned}\left\|\widehat{\mathcal{H}}_{(i)}^{-1}\right\|_\infty &= K\left(\min_{1\leq i\leq k_{\theta,n}}\{\underline{\lambda}_{(i),n}\}\right)^{-1} = O_p(1) \\ \left\|\mathcal{H}_{(i)}^{-1}\right\|_\infty &= K\left(\min_{1\leq i\leq k_{\theta,n}}\{\underline{\lambda}_{(i)}\}\right)^{-1} = O(1).\end{aligned}$$

Further, apply Lemma C.1 to $x_{(i),t}x_{(i),t}' - E[x_{(i),t}x_{(i),t}']$ element-wise to yield

$$\left\|\widehat{\mathcal{H}}_{(i)} - \mathcal{H}_{(i)}\right\|_\infty = O_p\left(\sqrt{\frac{\ln\left(k_{\theta,n}\right)}{n}E\left[\max_{1\leq i\leq k_{\theta,n}}\left|x_{(i),t}x_{(i),t}' - E\left[x_{(i),t}x_{(i),t}'\right]\right|^2\right]}\right). \tag{C.13}$$

Notice by Minkowski and Cauchy-Schwartz inequalities:

$$E\left[\max_{1\leq i\leq k_{\theta,n}}\left|x_{(i),t}x_{(i),t}' - E\left[x_{(i),t}x_{(i),t}'\right]\right|^2\right] \leq 2\mathcal{M}_n.$$

Therefore as claimed $||\widehat{\mathcal{H}}_{(i)}^{-1} - \mathcal{H}_{(i)}^{-1}||_\infty = O_p(\sqrt{\ln\left(k_{\theta,n}\right)\mathcal{M}_n/n})$.

**Claim (c).** The parsimonious first order condition and positive definiteness of $\widehat{\mathcal{H}}_{(i)}$ *awp*1, combined with (a) and (b), yield:

$$\begin{aligned}\left\|\sqrt{n}\left(\hat{\beta}_{(i)} - \beta_{(i)}^*\right)\right\|_\infty &\leq \left\|\mathcal{H}_{(i)}^{-1}\right\|_\infty\left\|\sqrt{n}\widehat{\mathcal{G}}_{(i)}\right\|_\infty + \left\|\widehat{\mathcal{H}}_{(i)}^{-1} - \mathcal{H}_{(i)}^{-1}\right\|_\infty\left\|\sqrt{n}\widehat{\mathcal{G}}_{(i)}\right\|_\infty \\ &= O_p\left(\sqrt{\ln\left(k_{\theta,n}\right)\mathcal{M}_n}\right) + O_p\left(\frac{\ln(k_{\theta,n})}{\sqrt{n}}\mathcal{M}_n\right).\end{aligned} \tag{C.14}$$

The last line is asymptotically dominated by $\sqrt{\ln\left(k_{\theta,n}\right)\mathcal{M}_n}$ provided $\ln(k_{\theta,n}) = o(\sqrt{n}/\mathcal{M}_n)$. $\mathcal{QED}$.

# D Max-Statistic Limit Law

The goal of this section is to characterize a setting in which the Theorem 2.4 (cf. (6)) max-statistic limit process $\{\boldsymbol{Z}_{(i)}([\mathbf{0}'_{k_\delta}, 1])\}_{i \in \mathbb{N}}$ belongs to the max-domain of attraction. In that case for some positive sequences $\{a_n, b_n\}$

$$\frac{1}{a_n}\left(\max_{1 \leq i \leq k_{\theta,n}} \left|\boldsymbol{Z}_{(i)}([\mathbf{0}'_{k_\delta}, 1])\right| - b_n\right)$$

has a well defined limit law. A direct route taken here is to ensure $\boldsymbol{Z}_{(i)}([\mathbf{0}'_{k_\delta}, 1])$ are iid $N(0,1)$, after a simple rescaling. Then $\sqrt{2\ln(k_{\theta,n})}(\max_{1 \leq i \leq k_{\theta,n}} |\boldsymbol{Z}_{(i)}([\mathbf{0}'_{k_\delta}, 1])| - b_n)$ is asymptotically Gumbel by classic arguments, thus $a_n = 1/\sqrt{2\ln(k_{\theta,n})}$ and $b_n \sim \sqrt{2\ln(k_{\theta,n})}$.

In order for $\boldsymbol{Z}_{(i)}([\mathbf{0}'_{k_\delta}, 1])$ to be mutually independent we must control the presence of nuisance parameters $\delta_0$ in the high dimensional regression model

$$y_t = \delta_0' x_{\delta,t} + \theta_0' x_{\theta,t} + \epsilon_t = \beta_0' x_t + \epsilon_t. \tag{D.1}$$

This follows because each $\sqrt{n}\hat{\theta}_i$ is a function of the covariate subset $x_{\delta,t}$ and therefore possibly asymptotically dependent, rendering $\boldsymbol{Z}_{(i)}([\mathbf{0}'_{k_\delta}, 1])$ possibly dependent across $i$. We therefore either require $k_\delta = 0$ or $(x_{\delta,t}, x_{\theta,t})$ to be mutually uncorrelated.

Consider $k_\delta = 0$. We can then always assume $x_{i,t} = x_{i,\theta,t}$ are zero-mean mutually uncorrelated by a trivial reparameterization of (D.1). Write $\tilde{y}_t \equiv y_t - E[y_t]$, $\tilde{x}_{\theta,t} \equiv \mathcal{H}_{\theta,\theta}^{-1}(x_{\theta,t} - E[x_{\theta,t}])$, and $\tilde{\theta}_0 \equiv \mathcal{H}_{\theta,\theta}\theta_0$, with $\mathcal{H}_{\theta,\theta} \equiv E[(x_{\theta,t} - E[x_{\theta,t}])(x_{\theta,t} - E[x_{\theta,t}])']$ assumed positive definite. Then (D.1) becomes

$$\tilde{y}_t = \tilde{\theta}_0' \tilde{x}_{\theta,t} + \epsilon_t \text{ where } E\left[\tilde{y}_t\right] = E\left[\tilde{x}_{\theta,t}\right] = 0 \text{ and } E\left[\tilde{x}_{\theta,t}\tilde{x}_{\theta,t}'\right] = I_{k_\theta}.$$

Clearly $\tilde{\theta}_0 = \mathbf{0}_{k_\theta}$ *if and only if* $\theta_0 = \mathbf{0}_{k_\theta}$ by positive definiteness of $\mathcal{H}_{\theta,\theta}$. In the proof of Lemma D.1, below, we show that $E[\tilde{x}_{\theta,t}\tilde{x}_{\theta,t}'] = I_{k_\theta}$, serial independence, and mutual independence between $x_{\theta,t}$ and $\epsilon_t$, yield the desired max-domain of attraction property.

Conversely suppose $k_\delta > 0$, and assume positive definite

$$\mathcal{H} \equiv E[(x_t - E[x_t])(x_t - E[x_t])']$$

is block-diagonal such that $(x_{\delta,t}, x_{\theta,t})$ are mutually uncorrelated:

$$\mathcal{H} = \begin{bmatrix} \mathcal{H}_{\delta,\delta} & \mathbf{0}_{k_\delta \times k_\theta} \\ \mathbf{0}_{k_\theta \times k_\delta} & \mathcal{H}_{\theta,\theta} \end{bmatrix}.$$

Now write $\tilde{y}_t \equiv y_t - E[y_t]$, $\tilde{x}_t \equiv \mathcal{H}^{-1}(x_t - E[x_t])$. Then by block-diagonality

$$\tilde{\beta}_0 \equiv \mathcal{H}\beta_0 = [(\mathcal{H}_{\delta,\delta}\delta)', (\mathcal{H}_{\theta,\theta}\theta)']' = \left[\tilde{\delta}', \tilde{\theta}'\right]',$$

say. Thus

$$\tilde{y}_t = \tilde{\beta}_0'\tilde{x}_t + \epsilon_t = \tilde{\delta}'\tilde{x}_{\delta,t} + \tilde{\theta}'\tilde{x}_{\theta,t} + \epsilon_t,$$

and again $\tilde{\theta}_0 = \mathbf{0}_{k_\theta}$ *if and only if* $\theta_0 = \mathbf{0}_{k_\theta}$.

Now let $\widehat{\tilde{\theta}}_i$ be the least squares estimator of $\tilde{\theta}_i^*$ in the (reparameterized) parsimonious models

$$\tilde{y}_t = \tilde{\delta}^{*\prime} x_{\delta,t} + \tilde{\theta}_i^* \tilde{x}_{\theta,i,t} + v_{(i),t}.$$

In order to reduce notation, assume $\mathcal{W}_{n,i} = \mathcal{W}_i = 1$. The general case is similar.

**Lemma D.1.** *Let Assumption 1 holds and assume $\mathcal{H}$ is positive definite. Let $\mathfrak{G}$ be a Gumbel random variable. Then for some $\{b_n\}$, $b_n \sim \sqrt{2\ln(k_{\theta,n})}$, and any $\{k_{\theta,n}\}$, $\ln(k_{\theta,n}) = o([\sqrt{n}/\mathcal{M}_n]^{2/3})$, we have under under $H_0$*

$$\sqrt{2\ln(k_{\theta,n})}\left(\frac{\mathcal{T}_n - b_n}{\sqrt{E[\epsilon_t^2]}}\right) \xrightarrow{d} \mathfrak{G}.$$

**Proof.** Let $H_0$ hold. Let $E[\epsilon_t^2] = 1$, $E[x_{j,t}] = 0$ and $E[x_{\theta,t}x_{\theta,t}'] = I_{k_\theta}$ to reduce notation, allowing us simply to work in model (D.1), where $\mathcal{T}_n = \max_{1\leq i\leq k_{\theta,n}} |\sqrt{n}\mathcal{W}_{n,i}\hat{\theta}_i|$, etc. Assume $k_\delta = 0$. The proof is nearly the same if $k_\delta > 0$ and $(x_{\delta,t}, x_{\theta,t})$ are mutually uncorrelated.

We have:

$$\begin{aligned}\sqrt{2\ln(k_{\theta,n})}\,(\mathcal{T}_n - b_n) \;&=\; \sqrt{2\ln(k_{\theta,n})}\left(\left|\sqrt{n}\hat{\theta}_i\right|_\infty - \left|\frac{1}{\sqrt{n}}\sum_{t=1}^n \epsilon_t x_{\theta,i,t}\right|_\infty\right) \\ &\quad+\sqrt{2\ln(k_{\theta,n})}\left(\left|\frac{1}{\sqrt{n}}\sum_{t=1}^n \epsilon_t x_{\theta,i,t}\right|_\infty - b_n\right).\end{aligned} \tag{D.2}$$

**Step 1.** Consider the second term in (D.2). Theorem 2.4 and $E[\epsilon_t^2] = E[x_{\theta,i,t}^2] = 1$ yields that for some Gaussian sequence $\{\boldsymbol{Z}_{(i)} : n \in \mathbb{N}\}_{i=1}^\infty$, $\boldsymbol{Z}_{(i)} \sim N(0,1)$, and any slowly varying $g(n) \to \infty$ (note: $z \geq 0$),

$$\sup_{z\geq 0}\left|P\left(\left|\frac{1}{\sqrt{n}}\sum_{t=1}^n \epsilon_t x_{\theta,i,t}\right|_\infty \leq z\right) - P\left(\left|\boldsymbol{Z}_{(i)}\right|_\infty \leq z\right)\right| = O(1/g(n)),$$

where by case: $(i)$ $\ln(k_{\theta,n}) = o(n^{1/7})$; $(ii)$ $\limsup_{n\to\infty} k_{\theta,n}/n^a > 0$ and $\ln(k_{\theta,n}) = o(n^{1/7})$; $(iii)$ $\limsup_{n\to\infty} k_{\theta,n}/n^a > 0$ and $\ln(k_{\theta,n}) = o(n^{1/7}/g(n)^{4/7})$; and $(iv)$ $k_{\theta,n} = o\left(n^{p/4}/g(n)^p\right)$ where

$p \geq 4$. Trivially therefore (note $z \in \mathbb{R}$):

$$\sup_{z\in\mathbb{R}} \left| P\left( \left| \frac{1}{\sqrt{n}} \sum_{t=1}^{n} \epsilon_t x_{\theta,i,t} \right|_\infty \leq z \right) - P\left( \left| \boldsymbol{Z}_{(i)} \right|_\infty \leq z \right) \right| \to 0.$$

Uniformity implies convergence holds for any sequence of real numbers $\{z_n\}$:

$$\left| P\left( \left| \frac{1}{\sqrt{n}} \sum_{t=1}^{n} \epsilon_t x_{\theta,i,t} \right|_\infty \leq z_n \right) - P\left( \left| \boldsymbol{Z}_{(i)} \right|_\infty \leq z_n \right) \right| \to 0.$$

Now, for any positive real sequence $\{b_n\}$ let

$$z_n = b_n + \frac{1}{\sqrt{2\ln(k_{\theta,n})}} z \text{ for arbitrary } z \in \mathbb{R}.$$

Rearrange terms to deduce:

$$\begin{aligned} &\left| P\left( \sqrt{2\ln(k_{\theta,n})} \left( \left| \frac{1}{\sqrt{n}} \sum_{t=1}^{n} \epsilon_t x_{\theta,i,t} \right|_\infty - b_n \right) \leq z \right) \right. \\ &\qquad\qquad \left. - P\left( \sqrt{2\ln(k_{\theta,n})} \left( \left| \boldsymbol{Z}_{(i)} \right|_\infty - b_n \right) \leq z \right) \right| \to 0. \end{aligned} \tag{D.3}$$

The iid property over $t$, square integrability, mutual independence between $\epsilon_t$ and $x_{\theta,i,t}$, mutual uncorrelatedness between (re-scaled) $x_{\theta,i,t}$ and $x_{\theta,j,t}$ $\forall i \neq j$, and $E[\epsilon_t^2] = E[x_{\theta,i,t}^2] = 1$ imply for any $h \in \mathbb{N}$, and any $h$-tuple $\{i_1, ..., i_h\}$ of positive integers $i_l \in \{1, ..., k_{\theta,n}\}$,

$$\frac{1}{\sqrt{n}} \sum_{t=1}^{n} \epsilon_t \sum_{l=1}^{h} \lambda_l x_{\theta,i,t} \xrightarrow{d} N(0,1) \;\; \forall \lambda'\lambda = 1.$$

This yields convergence in finite dimensional distributions to a multivariate standard normal law:

$$\left\{ \frac{1}{\sqrt{n}} \sum_{t=1}^{n} \epsilon_t x_{\theta,i,t} : 1 \leq i \leq h \right\} \xrightarrow{d} N(0, I_h).$$

The normalized Theorem 2.4 Gaussian limit laws $\boldsymbol{Z}_{(i)}(\lambda)$ are therefore iid $N(0,1)$. It thus follows instantly that for some $b_n \sim \sqrt{2\ln(k_{\theta,n})}$, and a Gumbel random variable $\mathfrak{G}$ (e.g. de Haan and Ferreira, 2000, Theorem 1.1.2, Example 1.1.7):

$$\left| P\left( \sqrt{2\ln(k_{\theta,n})} \left( \left| \boldsymbol{Z}_{(i)}(\lambda) \right|_\infty - b_n \right) \leq z \right) - P\left( \mathfrak{G} \leq z \right) \right| \to 0 \;\forall z \in \mathbb{R}. \tag{D.4}$$

Combine (D.3) and (D.4) to obtain for $k_{\theta,n}$ bounded above by case:

$$\sqrt{2\ln(k_{\theta,n})}\left(\left|\frac{1}{\sqrt{n}}\sum_{t=1}^{n}\epsilon_t x_{\theta,i,t}\right| P\left(\sqrt{2\ln(k_{\theta,n})}\left(\left|\boldsymbol{Z}_{(i)}\right|_{\infty}-b_n\right)\leq z\right)-b_n\right)\overset{d}{\to}\mathfrak{G}.$$

**Step 2.** It remains to prove that the first term (D.2) satisfies for certain $\{k_{\theta,n}\}$:

$$\sqrt{2\ln(k_{\theta,n})}\left(\left|\sqrt{n}\hat{\theta}_i\right|_{\infty}-\left|\frac{1}{\sqrt{n}}\sum_{t=1}^{n}\epsilon_t x_{\theta,i,t}\right|_{\infty}\right)\overset{p}{\to}0.$$

Use arguments in the proof of Lemma 2.2, with $\mathcal{W}_{n,i}=\mathcal{W}_i=1$, $k_\delta=0$, mutual uncorrelatedness and $E[x^2_{\theta,i,t}]=1$, to deduce

$$\sqrt{2\ln(k_{\theta,n})}\left|\left|\sqrt{n}\hat{\theta}_i\right|_{\infty}-\left|\frac{1}{\sqrt{n}}\sum_{t=1}^{n}\epsilon_t x_{\theta,i,t}\right|_{\infty}\right|=O_p\left(\frac{(\ln(k_{\theta,n}))^{3/2}\mathcal{M}_n}{\sqrt{n}}\right). \tag{D.5}$$

Therefore $\ln(k_{\theta,n})=o([\sqrt{n}/\mathcal{M}_n]^{2/3})$ ensures the desired result. $\mathcal{QED}$.

**Remark 2.** The requirement $\ln(k_{\theta,n})=o([\sqrt{n}/\mathcal{M}_n]^{2/3})$ holds in each covariate case $(i)$ – $(iv)$ by exploiting the $\mathcal{M}_n$ bounds in (C.5). In order to see this, notice that $(i)$ is trivial.

For $(ii)$ we have $\ln(k_{\theta,n})=o(n^{1/7})$. But $n^{1/7}=o(n^{1/3}/\mathcal{M}_n^{2/3})$ because $\mathcal{M}_n=o(n^{2/7})$ follows from $\mathcal{M}_n=O(\sqrt{\ln(k_{\theta,n})})$ in (C.5). Hence $\ln(k_{\theta,n})=o([\sqrt{n}/\mathcal{M}_n]^{2/3})$. Case $(iii)$ is essentially identical.

Finally, under $(iv)$ we have $k_{\theta,n}=o\left(n^{p/4}/g(n)^p\right)$ for slowly varying $g(n)$, hence $\ln(k_{\theta,n})=o\left(\ln(n)\right)$. Further $\mathcal{M}_n=O(k_{\theta,n}^{4/p})$ from (C.5). Therefore $\mathcal{M}_n=o\left(n/g(n)^4\right)=O(n^{1/2}/\left(\ln(n)\right)^{3/2})$. Re-arrange terms to yield $\ln(k_{\theta,n})=o([\sqrt{n}/\mathcal{M}_n]^{2/3})$.

# E Simulation Robustness Checks

As robustness checks we allow heterogeneous covariates, and large(r) $k_\delta$. We focus exclusively on the purposed parsimonious max-tests given its dominance in the baseline simulations. In order to focus attention on the robustness checks themselves, we work under the null with unbounded iid covariates.

## E.1 Regressor Dispersion

While the covariates $x_{\theta,i,t}$ are not homogeneous in all benchmark cases, their variances are fairly similar or identical depending on case. The dispersion of $x_{\theta,i,t}$, however, has a direct impact on the dispersion of the key parsimonious model estimators $\hat{\theta}_i$.

We therefore add three covariate cases allowing for different variances across $x_{\theta,t} = [x_{\theta,i,t}] \sim N(0, \Psi)$. We bypass covariate dependence and let $[x_{\delta,t}, x_{\theta,t}]$ be within and across block iid, in order to focus on the pure effects of covariate dispersion. When present, each nuisance $x_{\delta,i,t}$ is iid $N(0,1)$. In each case we use a diagonal variance matrix $\Psi = [\Psi_{i,j}]$. In case $(a)$ $\Psi_{i,i} = 1 + 100(i-1)/k_{\theta,n} \in [1, 100]$ hence $x_{\theta,i,t}$ has a monotonically larger variance as $i$ increases. In this setting a max-t-test may perform better due to self-scaling. In the remaining two cases diagonal $\Psi$ has entries $(b)$ $\Psi_{1,1} = 10$ or $(c)$ $\Psi_{1,1} = 100$, and all other $\Psi_{i,i} = 1$, hence there is a single potentially influential regressor.

See Tables 9-11.Covariate dispersion matters more when there are nuisance parameters, which also applied in the benchmark simulations. Overall, irrespective of the three variance cases, the max-tests are slightly over-sized when $k_\delta = 10$, in particular at $n = 100$. The max-t-test is generally better able to control for covariate dispersion, whether across covariates, or in one influential covariate.

## E.2 Large $k_\delta$

We study two nuisance parameter cases, $k_\delta \in \{20, 40\}$, under the null. Thus $k_\delta$ is larger than in the baseline cases $\{0, 10\}$. The covariates are within and across group iid standard normal. Results are contained in Table 12.

In the proceeding baseline simulations we saw all max-tests (parsimonious, and debiased Lasso where applicable) logically perform better when there are zero nuisance parameters. It is not surprising, then, that the tests perform precipitously less well when there are many nuisance parameters, in particular at small sample sizes (amounting to a classic signal/noise problem). The tests exhibit increasing size distortions from $k_\delta = 10$ to 20 to 40, in particular at $n = 100$, but also $n = 250$ when $k_\delta = 40$. Test performance improves as $n$ increases, with fairly accurate size under $k_\delta = 20$ and $n = 500$, and monotonic (in $n$) improvements when $k_\delta = 40$. In experiments not reported here we see competitive size when $k_\delta = 40$ with $n = 1000$. A plausible solution for future consideration is to partial out a large, or even high, dimension nuisance parameter, similar to Cattaneo, Jansson, and Newey (2018)..

# F Complete Simulation Results

## F.1 Benchmark Results

### F.1.1 Bounded Regressors

Table 1: Rejection Frequencies under $H_0 : \theta = 0$

| | $k_\delta = 0$ | | | | | | | | | $k_\delta = 10$ | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $n = 100$ | | | | | | | | | | | | | | | | | | |
| | $k_{\theta,n} = 35$ | | | $k_{\theta,n} = 200$ | | | $k_{\theta,n} = 482$ | | | $k_{\theta,n} = 35$ | | | $k_{\theta,n} = 200$ | | | $k_{\theta,n} = 482$ | | |
| Test / Size | 1% | 5% | 10% | 1% | 5% | 10% | 1% | 5% | 10% | 1% | 5% | 10% | 1% | 5% | 10% | 1% | 5% | 10% |
| *p*-Max | .009 | .048 | .930 | .011 | .061 | .113 | .010 | .055 | .110 | .014 | .061 | .128 | .013 | .062 | .131 | .004 | .046 | .112 |
| *p*-Max-t | .012 | .055 | .103 | .014 | .065 | .127 | .012 | .064 | .125 | .018 | .065 | .142 | .019 | .125 | .142 | .017 | .122 | .162 |
| *dbl*-Max | .010 | .060 | .116 | .004 | .034 | .068 | .014 | .068 | .122 | .000 | .000 | .000 | .008 | .030 | .056 | .016 | .160 | .242 |
| *dbl*-Max-t | .012 | .046 | .096 | .012 | .034 | .090 | .042 | .096 | .232 | .000 | .002 | .002 | .000 | .000 | .000 | .000 | .000 | .000 |
| Wald(a) | .584 | 1.00 | 1.00 | - | - | - | - | - | - | .774 | 1.00 | 1.00 | - | - | - | - | - | - |
| Wald(b) | .000 | .000 | .024 | - | - | - | - | - | - | .000 | .022 | .072 | - | - | - | - | - | - |
| $n = 250$ | | | | | | | | | | | | | | | | | | |
| | $k_{\theta,n} = 35$ | | | $k_{\theta,n} = 1144$ | | | $k_{\theta,n} = 1250$ | | | $k_{\theta,n} = 35$ | | | $k_{\theta,n} = 1144$ | | | $k_{\theta,n} = 1250$ | | |
| Test / Size | 1% | 5% | 10% | 1% | 5% | 10% | 1% | 5% | 10% | 1% | 5% | 10% | 1% | 5% | 10% | 1% | 5% | 10% |
| *p*-Max | .010 | .047 | .094 | .011 | .045 | .103 | .006 | .044 | .096 | .015 | .063 | .126 | .008 | .046 | .093 | .007 | .048 | .101 |
| *p*-Max-t | .022 | .054 | .207 | .011 | .053 | .108 | .009 | .052 | .107 | .015 | .068 | .128 | .015 | .061 | .125 | .014 | .056 | .131 |
| Wald(a) | .226 | 1.00 | 1.00 | - | - | - | - | - | - | .026 | 1.00 | 1.00 | - | - | - | - | - | - |
| Wald(b) | .002 | .029 | .067 | - | - | - | - | - | - | .006 | .043 | .097 | - | - | - | - | - | - |
| $n = 500$ | | | | | | | | | | | | | | | | | | |
| | $k_{\theta,n} = 35$ | | | $k_{\theta,n} = 2381$ | | | $k_{\theta,n} = 5000$ | | | $k_{\theta,n} = 35$ | | | $k_{\theta,n} = 2381$ | | | $k_{\theta,n} = 5000$ | | |
| Test / Size | 1% | 5% | 10% | 1% | 5% | 10% | 1% | 5% | 10% | 1% | 5% | 10% | 1% | 5% | 10% | 1% | 5% | 10% |
| *p*-Max | .008 | .047 | .095 | .013 | .050 | .101 | .008 | .047 | .092 | .013 | .057 | .116 | .006 | .053 | .107 | .007 | .052 | .106 |
| *p*-Max-t | .013 | .049 | .101 | .013 | .057 | .100 | .008 | .048 | .091 | .014 | .056 | .117 | .013 | .061 | .118 | .012 | .056 | .108 |
| Wald(a) | .109 | 1.00 | 1.00 | - | - | - | - | - | - | .143 | 1.00 | 1.00 | - | - | - | - | - | - |
| Wald(b) | .003 | .031 | .072 | - | - | - | - | - | - | .002 | .047 | .099 | - | - | - | - | - | - |

"b" = bootstrapped p-value; "a" = asymptotic test. $p$ = parsimonious; *dbl* = de-biased Lasso. All test p-values are bootstrapped, based on 1,000 ($p$-max, Wald) or 250 (*dbl*-max) independently drawn samples.

Table 2: Rejection Frequencies under $H_1(i) : \theta_1 = .0015$ and $\theta_i = 0$, $i \geq 2$

| | $k_\delta = 0$ | | | | | | | | | $k_\delta = 10$ | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | $n = 100$ | | | | | | | | | | | | | | | | | |
| | $k_{\theta,n} = 35$ | | | $k_{\theta,n} = 200$ | | | $k_{\theta,n} = 482$ | | | $k_{\theta,n} = 35$ | | | $k_{\theta,n} = 200$ | | | $k_{\theta,n} = 482$ | | |
| Test / Size | 1% | 5% | 10% | 1% | 5% | 10% | 1% | 5% | 10% | 1% | 5% | 10% | 1% | 5% | 10% | 1% | 5% | 10% |
| $p$-Max | .921 | .977 | .989 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | .048 | .200 | .329 | .850 | .962 | .978 | 1.00 | 1.00 | 1.00 |
| $p$-Max-t | .946 | .988 | .993 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | .145 | .340 | .496 | .941 | .975 | .988 | 1.00 | 1.00 | 1.00 |
| *dbl*-Max | .028 | .052 | .108 | .008 | .056 | .108 | .028 | .128 | .244 | .000 | .004 | .004 | .008 | .040 | .084 | .016 | .048 | .092 |
| *dbl*-Max-t | .004 | .048 | .084 | .116 | .144 | .192 | .736 | .828 | .848 | .000 | .001 | .002 | .002 | .002 | .003 | .004 | .005 | .006 |
| Wald(a) | .989 | 1.00 | 1.00 | - | - | - | - | - | - | .887 | 1.00 | 1.00 | - | - | - | - | - | - |
| Wald(b) | .008 | .182 | .442 | - | - | - | - | - | - | .001 | .038 | .164 | - | - | - | - | - | - |
| | $n = 250$ | | | | | | | | | | | | | | | | | |
| | $k_{\theta,n} = 35$ | | | $k_{\theta,n} = 1144$ | | | $k_{\theta,n} = 1250$ | | | $k_{\theta,n} = 35$ | | | $k_{\theta,n} = 1144$ | | | $k_{\theta,n} = 1250$ | | |
| Test / Size | 1% | 5% | 10% | 1% | 5% | 10% | 1% | 5% | 10% | 1% | 5% | 10% | 1% | 5% | 10% | 1% | 5% | 10% |
| $p$-Max | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | .559 | .812 | .886 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 |
| $p$-Max-t | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | .773 | .902 | .943 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 |
| Wald(a) | .998 | 1.00 | 1.00 | - | - | - | - | - | - | .854 | 1.00 | 1.00 | - | - | - | - | - | - |
| Wald(b) | .764 | .955 | .981 | - | - | - | - | - | - | .164 | .460 | .651 | - | - | - | - | - | - |
| | $n = 500$ | | | | | | | | | | | | | | | | | |
| | $k_{\theta,n} = 35$ | | | $k_{\theta,n} = 2381$ | | | $k_{\theta,n} = 5000$ | | | $k_{\theta,n} = 35$ | | | $k_{\theta,n} = 2381$ | | | $k_{\theta,n} = 5000$ | | |
| Test / Size | 1% | 5% | 10% | 1% | 5% | 10% | 1% | 5% | 10% | 1% | 5% | 10% | 1% | 5% | 10% | 1% | 5% | 10% |
| $p$-Max | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | .952 | .991 | .998 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 |
| $p$-Max-t | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | .995 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 |
| Wald(a) | 1.00 | 1.00 | 1.00 | - | - | - | - | - | - | .972 | 1.00 | 1.00 | - | - | - | - | - | - |
| Wald(b) | 1.00 | 1.00 | 1.00 | - | - | - | - | - | - | .765 | .920 | .964 | - | - | - | - | - | - |

"b" = bootstrapped p-value; "a" = asymptotic test. $p$ = parsimonious; *dbl* = de-biased Lasso. All test p-values are bootstrapped, based on 1,000 ($p$-max, Wald) or 250 (*dbl*-max) independently drawn samples.

Table 3: Rejection Frequencies under $H_1(ii)$ : $\theta_{0,i} = .0002i/k_{\theta,n}$ for $i = 1, ..., k_{\theta,n}$

| | $k_\delta = 0$ | | | | | | | | | $k_\delta = 10$ | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | $n = 100$ | | | | | | | | | | | | | | | | | |
| | $k_{\theta,n} = 35$ | | | $k_{\theta,n} = 200$ | | | $k_{\theta,n} = 482$ | | | $k_{\theta,n} = 35$ | | | $k_{\theta,n} = 200$ | | | $k_{\theta,n} = 482$ | | |
| Test / Size | 1% | 5% | 10% | 1% | 5% | 10% | 1% | 5% | 10% | 1% | 5% | 10% | 1% | 5% | 10% | 1% | 5% | 10% |
| $p$-Max | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | .023 | .127 | .243 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 |
| $p$-Max-t | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | .068 | .214 | .342 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 |
| *dbl*-Max | .008 | .036 | .096 | .152 | .380 | .568 | .380 | .724 | .812 | .000 | .000 | .004 | .000 | .008 | .020 | .004 | .012 | .028 |
| *dbl*-Max-t | .028 | .060 | .112 | .340 | .504 | .552 | .402 | .789 | .821 | .008 | .012 | .012 | .000 | .000 | .004 | .004 | .004 | .016 |
| Wald(a) | 1.00 | 1.00 | 1.00 | - | - | - | - | - | - | .870 | 1.00 | 1.00 | - | - | - | - | - | - |
| Wald(b) | .467 | .974 | .998 | - | - | - | - | - | - | .000 | .025 | .111 | - | - | - | - | - | - |
| | $n = 250$ | | | | | | | | | | | | | | | | | |
| | $k_{\theta,n} = 35$ | | | $k_{\theta,n} = 1144$ | | | $k_{\theta,n} = 1250$ | | | $k_{\theta,n} = 35$ | | | $k_{\theta,n} = 1144$ | | | $k_{\theta,n} = 1250$ | | |
| Test / Size | 1% | 5% | 10% | 1% | 5% | 10% | 1% | 5% | 10% | 1% | 5% | 10% | 1% | 5% | 10% | 1% | 5% | 10% |
| $p$-Max | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | .081 | .237 | .382 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 |
| $p$-Max-t | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | .131 | .396 | .438 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 |
| Wald(a) | 1.00 | 1.00 | 1.00 | - | - | - | - | - | - | .526 | 1.00 | 1.00 | - | - | - | - | - | - |
| Wald(b) | 1.00 | 1.00 | 1.00 | - | - | - | - | - | - | .020 | .128 | .257 | - | - | - | - | - | - |
| | $n = 500$ | | | | | | | | | | | | | | | | | |
| | $k_{\theta,n} = 35$ | | | $k_{\theta,n} = 2381$ | | | $k_{\theta,n} = 5000$ | | | $k_{\theta,n} = 35$ | | | $k_{\theta,n} = 2381$ | | | $k_{\theta,n} = 5000$ | | |
| Test / Size | 1% | 5% | 10% | 1% | 5% | 10% | 1% | 5% | 10% | 1% | 5% | 10% | 1% | 5% | 10% | 1% | 5% | 10% |
| $p$-Max | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | .254 | .560 | .720 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 |
| $p$-Max-t | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | .454 | .722 | .844 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 |
| Wald(a) | 1.00 | 1.00 | 1.00 | - | - | - | - | - | - | .715 | 1.00 | 1.00 | - | - | - | - | - | - |
| Wald(b) | 1.00 | 1.00 | 1.00 | - | - | - | - | - | - | .237 | .499 | .650 | - | - | - | - | - | - |

"b" = bootstrapped p-value; "a" = asymptotic test. $p$ = parsimonious; *dbl* = de-biased Lasso. All test p-values are bootstrapped, based on 1,000 ($p$-max, Wald) or 250 (*dbl*-max) independently drawn samples.

Table 4: Rejection Frequencies under $H_1(iii) : \theta_i = .00015$ for $i = 1, ..., 10$

| | $k_\delta = 0$ | | | | | | | | | $k_\delta = 10$ | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | $n = 100$ | | | | | | | | | | | | | | | | | |
| | $k_{\theta,n} = 35$ | | | $k_{\theta,n} = 200$ | | | $k_{\theta,n} = 482$ | | | $k_{\theta,n} = 35$ | | | $k_{\theta,n} = 200$ | | | $k_{\theta,n} = 482$ | | |
| Test / Size | 1% | 5% | 10% | 1% | 5% | 10% | 1% | 5% | 10% | | | | 1% | 5% | 10% | 1% | 5% | 10% |
| $p$-Max | .878 | .970 | .986 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | .022 | .098 | .196 | .039 | .174 | .342 | .058 | .253 | .450 |
| $p$-Max-t | .920 | .978 | .990 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | .034 | .149 | .242 | .118 | .362 | .540 | .213 | .508 | .705 |
| *dbl*-Max | .004 | .052 | .100 | .008 | .020 | .040 | .008 | .008 | .012 | .000 | .000 | .000 | .004 | .036 | .056 | .004 | .016 | .028 |
| *dbl*-Max-t | .000 | .000 | .000 | .000 | .000 | .000 | .000 | .000 | .000 | .000 | .000 | .000 | .000 | .000 | .004 | .000 | .000 | .012 |
| Wald(a) | .963 | 1.00 | 1.00 | - | - | - | - | - | - | .809 | 1.00 | 1.00 | - | - | - | - | - | - |
| Wald(b) | .003 | .079 | .285 | - | - | - | - | - | - | .000 | .019 | .081 | - | - | - | - | - | - |
| | $n = 250$ | | | | | | | | | | | | | | | | | |
| | $k_{\theta,n} = 35$ | | | $k_{\theta,n} = 1144$ | | | $k_{\theta,n} = 1250$ | | | $k_{\theta,n} = 35$ | | | $k_{\theta,n} = 1144$ | | | $k_{\theta,n} = 1250$ | | |
| Test / Size | 1% | 5% | 10% | 1% | 5% | 10% | 1% | 5% | 10% | 1% | 5% | 10% | 1% | 5% | 10% | 1% | 5% | 10% |
| $p$-Max | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | .019 | .115 | .218 | .902 | .991 | 1.00 | .955 | 1.00 | 1.00 |
| $p$-Max-t | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | .041 | .173 | .282 | .946 | .996 | 1.00 | 1.00 | 1.00 | 1.00 |
| Wald(a) | .994 | 1.00 | 1.00 | - | - | - | - | - | - | .394 | 1.00 | 1.00 | - | - | - | - | - | - |
| Wald(b) | .694 | .923 | .974 | - | - | - | - | - | - | .009 | .067 | .157 | - | - | - | - | - | - |
| | $n = 500$ | | | | | | | | | | | | | | | | | |
| | $k_{\theta,n} = 35$ | | | $k_{\theta,n} = 2381$ | | | $k_{\theta,n} = 5000$ | | | $k_{\theta,n} = 35$ | | | $k_{\theta,n} = 2381$ | | | $k_{\theta,n} = 5000$ | | |
| Test / Size | 1% | 5% | 10% | 1% | 5% | 10% | 1% | 5% | 10% | 1% | 5% | 10% | 1% | 5% | 10% | 1% | 5% | 10% |
| $p$-Max | .987 | .997 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | .044 | .167 | .298 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 |
| $p$-Max-t | .986 | .998 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | .071 | .211 | .344 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 |
| Wald(a) | .922 | 1.00 | 1.00 | - | - | - | - | - | - | .289 | 1.00 | 1.00 | - | - | - | - | - | - |
| Wald(b) | .795 | .932 | .974 | - | - | - | - | - | - | .019 | .118 | .223 | - | - | - | - | - | - |

"b" = bootstrapped p-value; "a" = asymptotic test. $p$ = parsimonious; *dbl* = de-biased Lasso. All test p-values are bootstrapped, based on 1,000 ($p$-max, Wald) or 250 (*dbl*-max) independently drawn samples.

#### F.1.2 Unbounded Regressors

Table 5: Rejection Frequencies under $H_0 : \theta = 0$

| | $k_\delta = 0$ | | | | | | | | | $k_\delta = 10$ | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | $n = 100$ | | | | | | | | | | | | | | | | | |
| | $k_{\theta,n} = 35$ | | | $k_{\theta,n} = 200$ | | | $k_{\theta,n} = 482$ | | | $k_{\theta,n} = 35$ | | | $k_{\theta,n} = 200$ | | | $k_{\theta,n} = 482$ | | |
| Test / Size | 1% | 5% | 10% | 1% | 5% | 10% | 1% | 5% | 10% | 1% | 5% | 10% | 1% | 5% | 10% | 1% | 5% | 10% |
| $p$-Max | .010 | .042 | .116 | .013 | .053 | .103 | .008 | .055 | .096 | .015 | .063 | .132 | .009 | .056 | .123 | .006 | .048 | .125 |
| $p$-Max-t | .007 | .058 | .111 | .013 | .061 | .117 | .012 | .059 | .112 | .017 | .071 | .176 | .018 | .067 | .169 | .021 | .075 | .174 |
| *dbl*-Max | .014 | .054 | .108 | .000 | .022 | .068 | .021 | .078 | .176 | .000 | .000 | .000 | .010 | .040 | .084 | .014 | .120 | .194 |
| *dbl*-Max-t | .014 | .032 | .082 | .004 | .044 | .090 | .028 | .084 | .196 | .002 | .002 | .002 | .000 | .000 | .000 | .000 | .000 | .000 |
| Wald(a) | .590 | 1.00 | 1.00 | - | - | - | - | - | - | .797 | 1.00 | 1.00 | - | - | - | - | - | - |
| Wald(b) | .000 | .000 | .018 | - | - | - | - | - | - | .000 | .014 | .072 | - | - | - | - | - | - |
| | $n = 250$ | | | | | | | | | | | | | | | | | |
| | $k_{\theta,n} = 35$ | | | $k_{\theta,n} = 1144$ | | | $k_{\theta,n} = 1250$ | | | $k_{\theta,n} = 35$ | | | $k_{\theta,n} = 1144$ | | | $k_{\theta,n} = 1250$ | | |
| Test / Size | 1% | 5% | 10% | 1% | 5% | 10% | 1% | 5% | 10% | 1% | 5% | 10% | 1% | 5% | 10% | 1% | 5% | 10% |
| $p$-Max | .010 | .056 | .106 | .012 | .058 | .112 | .007 | .046 | .095 | .013 | .057 | .109 | .006 | .042 | .108 | .008 | .045 | .103 |
| $p$-Max-t | .012 | .055 | .106 | .015 | .061 | .116 | .009 | .057 | .097 | .016 | .060 | .122 | .014 | .062 | .127 | .014 | .067 | .132 |
| Wald(a) | .237 | 1.00 | 1.00 | - | - | - | - | - | - | .260 | 1.00 | 1.00 | - | - | - | - | - | - |
| Wald(b) | .000 | .015 | .066 | - | - | - | - | - | - | .004 | .045 | .093 | - | - | - | - | - | - |
| | $n = 500$ | | | | | | | | | | | | | | | | | |
| | $k_{\theta,n} = 35$ | | | $k_{\theta,n} = 2381$ | | | $k_{\theta,n} = 2500$ | | | $k_{\theta,n} = 35$ | | | $k_{\theta,n} = 2381$ | | | $k_{\theta,n} = 2500$ | | |
| Test / Size | 1% | 5% | 10% | 1% | 5% | 10% | 1% | 5% | 10% | 1% | 5% | 10% | 1% | 5% | 10% | 1% | 5% | 10% |
| $p$-Max | .012 | .053 | .105 | .011 | .055 | .103 | .007 | .052 | .103 | .011 | .058 | .106 | .011 | .051 | .107 | .014 | .056 | .106 |
| $p$-Max-t | .012 | .051 | .107 | .009 | .054 | .111 | .009 | .054 | .111 | .015 | .061 | .109 | .016 | .067 | .124 | .017 | .068 | .131 |
| Wald(a) | .116 | 1.00 | 1.00 | - | - | - | - | - | - | .156 | 1.00 | 1.00 | - | - | - | - | - | - |
| Wald(b) | .003 | .043 | .076 | - | - | - | - | - | - | .009 | .043 | .100 | - | - | - | - | - | - |

"b" = bootstrapped p-value; "a" = asymptotic test. $p$ = parsimonious; *dbl* = de-biased Lasso. All test p-values are bootstrapped, based on 1,000 ($p$-max, Wald) or 250 (*dbl*-max) independently drawn samples.

Table 6: Rejection Frequencies under $H_1(i) : \theta_1 = .0015$ and $\theta_i = 0,\ i \geq 2]$

| | $k_\delta = 0$ | | | | | | | | | $k_\delta = 10$ | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | $n = 100$ | | | | | | | | | | | | | | | | | |
| | $k_{\theta,n} = 35$ | | | $k_{\theta,n} = 200$ | | | $k_{\theta,n} = 482$ | | | $k_{\theta,n} = 35$ | | | $k_{\theta,n} = 200$ | | | $k_{\theta,n} = 482$ | | |
| Test / Size | 1% | 5% | 10% | 1% | 5% | 10% | 1% | 5% | 10% | 1% | 5% | 10% | 1% | 5% | 10% | 1% | 5% | 10% |
| $p$-Max | .923 | .981 | .993 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | .059 | .269 | .399 | .912 | .982 | .991 | 1.00 | 1.00 | 1.00 |
| $p$-Max-t | .946 | .986 | .994 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | .290 | .496 | .612 | .974 | .993 | .997 | 1.00 | 1.00 | 1.00 |
| *dbl*-Max | .012 | .048 | .092 | .020 | .044 | .100 | .036 | .180 | .332 | .000 | .004 | .008 | .008 | .044 | .084 | .008 | .036 | .064 |
| *dbl*-Max-t | .016 | .056 | .084 | .132 | .188 | .220 | .804 | .900 | .904 | .000 | .001 | .001 | .001 | .002 | .002 | .004 | .005 | .005 |
| Wald(a) | .979 | 1.00 | 1.00 | - | - | - | - | - | - | .914 | 1.00 | 1.00 | - | - | - | - | - | - |
| Wald(b) | .002 | .113 | .341 | - | - | - | - | - | - | .003 | .062 | .208 | - | - | - | - | - | - |
| | $n = 250$ | | | | | | | | | | | | | | | | | |
| | $k_{\theta,n} = 35$ | | | $k_{\theta,n} = 1144$ | | | $k_{\theta,n} = 1250$ | | | $k_{\theta,n} = 35$ | | | $k_{\theta,n} = 1144$ | | | $k_{\theta,n} = 1250$ | | |
| Test / Size | 1% | 5% | 10% | 1% | 5% | 10% | 1% | 5% | 10% | 1% | 5% | 10% | 1% | 5% | 10% | 1% | 5% | 10% |
| $p$-Max | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | .583 | .783 | .853 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 |
| $p$-Max-t | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | .569 | .737 | .831 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 |
| Wald(a) | 1.00 | 1.00 | 1.00 | - | - | - | - | - | - | .743 | 1.00 | 1.00 | - | - | - | - | - | - |
| Wald(b) | .961 | .996 | 1.00 | - | - | - | - | - | - | .083 | .335 | .497 | - | - | - | - | - | - |
| | $n = 500$ | | | | | | | | | | | | | | | | | |
| | $k_{\theta,n} = 35$ | | | $k_{\theta,n} = 2381$ | | | $k_{\theta,n} = 2500$ | | | $k_{\theta,n} = 35$ | | | $k_{\theta,n} = 2381$ | | | $k_{\theta,n} = 2500$ | | |
| Test / Size | 1% | 5% | 10% | 1% | 5% | 10% | 1% | 5% | 10% | 1% | 5% | 10% | 1% | 5% | 10% | 1% | 5% | 10% |
| $p$-Max | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | .989 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 |
| $p$-Max-t | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 |
| Wald(a) | 1.00 | 1.00 | 1.00 | - | - | - | - | - | - | .981 | 1.00 | 1.00 | - | - | - | - | - | - |
| Wald(b) | .998 | 1.00 | 1.00 | - | - | - | - | - | - | .830 | .953 | .971 | - | - | - | - | - | - |

"b" = bootstrapped p-value; "a" = asymptotic test. $p$ = parsimonious; *dbl* = de-biased Lasso. All test p-values are bootstrapped, based on 1,000 ($p$-max, Wald) or 250 (*dbl*-max) independently drawn samples.

Table 7: Rejection Frequencies under $H_1(ii) : \theta_{0,i} = .0002i/k_{\theta,n}$ for $i = 1, ..., k_{\theta,n}$

| | $k_\delta = 0$ | | | | | | | | | $k_\delta = 10$ | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $n = 100$ | | | | | | | | | | | | | | | | | | |
| | $k_{\theta,n} = 35$ | | | $k_{\theta,n} = 200$ | | | $k_{\theta,n} = 482$ | | | $k_{\theta,n} = 35$ | | | $k_{\theta,n} = 200$ | | | $k_{\theta,n} = 482$ | | |
| Test / Size | 1% | 5% | 10% | 1% | 5% | 10% | 1% | 5% | 10% | 1% | 5% | 10% | 1% | 5% | 10% | 1% | 5% | 10% |
| $p$-Max | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | .066 | .229 | .387 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 |
| $p$-Max-t | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | .129 | .336 | .493 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 |
| *dbl*-Max | .016 | .060 | .116 | .188 | .436 | .596 | .468 | .748 | .864 | .000 | .000 | .000 | .008 | .016 | .044 | .008 | .028 | .040 |
| *dbl*-Max-t | .008 | .036 | .088 | .380 | .536 | .584 | .503 | .810 | .900 | .000 | .000 | .000 | .000 | .000 | .000 | .000 | .008 | .012 |
| Wald(a) | 1.00 | 1.00 | 1.00 | - | - | - | - | - | - | .892 | 1.00 | 1.00 | - | - | - | - | - | - |
| Wald(b) | .542 | .982 | .998 | - | - | - | - | - | - | .001 | .051 | .180 | - | - | - | - | - | - |
| $n = 250$ | | | | | | | | | | | | | | | | | | |
| | $k_{\theta,n} = 35$ | | | $k_{\theta,n} = 1144$ | | | $k_{\theta,n} = 1250$ | | | $k_{\theta,n} = 35$ | | | $k_{\theta,n} = 1144$ | | | $k_{\theta,n} = 1250$ | | |
| Test / Size | 1% | 5% | 10% | 1% | 5% | 10% | 1% | 5% | 10% | 1% | 5% | 10% | 1% | 5% | 10% | 1% | 5% | 10% |
| $p$-Max | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | .093 | .309 | .450 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 |
| $p$-Max-t | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | .206 | .456 | .603 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 |
| Wald(a) | 1.00 | 1.00 | 1.00 | - | - | - | - | - | - | .547 | 1.00 | 1.00 | - | - | - | - | - | - |
| Wald(b) | 1.00 | 1.00 | 1.00 | - | - | - | - | - | - | .023 | .155 | .273 | - | - | - | - | - | - |
| $n = 500$ | | | | | | | | | | | | | | | | | | |
| | $k_{\theta,n} = 35$ | | | $k_{\theta,n} = 2381$ | | | $k_{\theta,n} = 2500$ | | | $k_{\theta,n} = 35$ | | | $k_{\theta,n} = 2381$ | | | $k_{\theta,n} = 2500$ | | |
| Test / Size | 1% | 5% | 10% | 1% | 5% | 10% | 1% | 5% | 10% | 1% | 5% | 10% | 1% | 5% | 10% | 1% | 5% | 10% |
| $p$-Max | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | .756 | .954 | .989 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 |
| $p$-Max-t | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | .909 | .983 | .995 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 |
| Wald(a) | 1.00 | 1.00 | 1.00 | - | - | - | - | - | - | .958 | 1.00 | 1.00 | - | - | - | - | - | - |
| Wald(b) | 1.00 | 1.00 | 1.00 | - | - | - | - | - | - | .700 | .898 | .941 | - | - | - | - | - | - |

"b" = bootstrapped p-value; "a" = asymptotic test. $p$ = parsimonious; *dbl* = de-biased Lasso. All test p-values are bootstrapped, based on 1,000 ($p$-max, Wald) or 250 (*dbl*-max) independently drawn samples.

Table 8: Rejection Frequencies under $H_1(iii) : \theta_i = .00015$ for $i = 1, ..., 10$

| | $k_\delta = 0$ | | | | | | | | | $k_\delta = 10$ | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | $n = 100$ | | | | | | | | | | | | | | | | | |
| | $k_{\theta,n} = 35$ | | | $k_{\theta,n} = 200$ | | | $k_{\theta,n} = 482$ | | | $k_{\theta,n} = 35$ | | | $k_{\theta,n} = 200$ | | | $k_{\theta,n} = 482$ | | |
| Test / Size | 1% | 5% | 10% | 1% | 5% | 10% | 1% | 5% | 10% | 1% | 5% | 10% | 1% | 5% | 10% | 1% | 5% | 10% |
| $p$-Max | .934 | .988 | .993 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | .019 | .112 | .190 | .046 | .183 | .339 | .087 | .308 | .523 |
| $p$-Max-t | .954 | .989 | .993 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | .051 | .165 | .275 | .267 | .573 | .750 | .267 | .578 | .750 |
| *dbl*-Max | .016 | .048 | .092 | .000 | .036 | .084 | .020 | .072 | .120 | .000 | .000 | .000 | .004 | .028 | .056 | .000 | .000 | .044 |
| *dbl*-Max-t | .012 | .036 | .076 | .000 | .008 | .016 | .016 | .024 | .028 | .000 | .000 | .000 | .000 | .000 | .000 | .000 | .012 | .016 |
| Wald(a) | .970 | 1.00 | 1.00 | - | - | - | - | - | - | .829 | 1.00 | 1.00 | - | - | - | - | - | - |
| Wald(b) | .003 | .101 | .307 | - | - | - | - | - | - | .000 | .025 | .110 | - | - | - | - | - | - |
| | $n = 250$ | | | | | | | | | | | | | | | | | |
| | $k_{\theta,n} = 35$ | | | $k_{\theta,n} = 1144$ | | | $k_{\theta,n} = 1250$ | | | $k_{\theta,n} = 35$ | | | $k_{\theta,n} = 1144$ | | | $k_{\theta,n} = 1250$ | | |
| Test / Size | 1% | 5% | 10% | 1% | 5% | 10% | 1% | 5% | 10% | 1% | 5% | 10% | 1% | 5% | 10% | 1% | 5% | 10% |
| $p$-Max | .998 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | .049 | .166 | .268 | .924 | .991 | .997 | .916 | .995 | 1.00 |
| $p$-Max-t | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | .055 | .178 | .299 | .977 | .997 | 1.00 | .979 | 1.00 | 1.00 |
| Wald(a) | .998 | 1.00 | 1.00 | - | - | - | - | - | - | .368 | 1.00 | 1.00 | - | - | - | - | - | - |
| Wald(b) | .730 | .944 | .980 | - | - | - | - | - | - | .005 | .062 | .157 | - | - | - | - | - | - |
| | $n = 500$ | | | | | | | | | | | | | | | | | |
| | $k_{\theta,n} = 35$ | | | $k_{\theta,n} = 2381$ | | | $k_{\theta,n} = 2500$ | | | $k_{\theta,n} = 35$ | | | $k_{\theta,n} = 2381$ | | | $k_{\theta,n} = 2500$ | | |
| Test / Size | 1% | 5% | 10% | 1% | 5% | 10% | 1% | 5% | 10% | 1% | 5% | 10% | 1% | 5% | 10% | 1% | 5% | 10% |
| $p$-Max | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | .016 | .116 | .213 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 |
| $p$-Max-t | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | .063 | .182 | .306 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 |
| Wald(a) | 1.00 | 1.00 | 1.00 | - | - | - | - | - | - | .281 | 1.00 | 1.00 | - | - | - | - | - | - |
| Wald(b) | 1.00 | 1.00 | 1.00 | - | - | - | - | - | - | .027 | .130 | .215 | - | - | - | - | - | - |

"b" = bootstrapped p-value; "a" = asymptotic test. $p$ = parsimonious; *dbl* = de-biased Lasso. All test p-values are bootstrapped, based on 1,000 ($p$-max, Wald) or 250 (*dbl*-max) independently drawn samples.

### F.2 Robustness Checks

Table 9: Robustness Check a.i: $V[x_{\theta,i,t}] \in [1, 100]$ increasing over $i$

| | $k_\delta = 0$ | | | | | | | | | $k_\delta = 10$ | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | $n = 100$ | | | | | | | | | | | | | | | | | |
| Test / Size | 1% | 5% | 10% | 1% | 5% | 10% | 1% | 5% | 10% | 1% | 5% | 10% | 1% | 5% | 10% | 1% | 5% | 10% |
| $k_{\theta,n}$ | 35 | | | 200 | | | 482 | | | 35 | | | 200 | | | 482 | | |
| $p$-Max | .014 | .057 | .100 | .011 | .051 | .104 | .010 | .051 | .099 | .024 | .071 | .123 | .023 | .090 | .154 | .022 | .091 | .157 |
| $p$-Max-t | .015 | .068 | .125 | .013 | .065 | .147 | .012 | .072 | .155 | .017 | .118 | .194 | .026 | .124 | .234 | .023 | .127 | .247 |
| | $n = 250$ | | | | | | | | | | | | | | | | | |
| $k_{\theta,n}$ | 35 | | | 1144 | | | 1250 | | | 35 | | | 1144 | | | 1250 | | |
| $p$-Max | .018 | .050 | .095 | .005 | .044 | .093 | .009 | .038 | .076 | .015 | .057 | .106 | .015 | .068 | .115 | .012 | .056 | .109 |
| $p$-Max-t | .010 | .050 | .104 | .006 | .034 | .114 | .013 | .056 | .105 | .012 | .076 | .141 | .015 | .067 | .149 | .016 | .069 | .144 |
| | $n = 500$ | | | | | | | | | | | | | | | | | |
| $k_{\theta,n}$ | 35 | | | 2381 | | | 5000 | | | 35 | | | 2381 | | | 5000 | | |
| $p$-Max | .015 | .055 | .104 | .004 | .043 | .086 | .008 | .041 | .092 | .015 | .053 | .103 | .012 | .049 | .099 | .009 | .049 | .091 |
| $p$-Max-t | .012 | .050 | .097 | .011 | .059 | .095 | .011 | .053 | .105 | .009 | .049 | .103 | .011 | .055 | .105 | .011 | .052 | .108 |

Table 10: Robustness Check a.ii: $V[x_{\theta,1,t}] = 10$, and $\{\{V[x_{\theta,i,t}]\}_{i=2}^{\mathring{k}_{n,\theta}}, V[x_{\delta,i,t}]\} = 1$

| | $k_\delta = 0$ | | | | | | | | | $k_\delta = 10$ | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | $n = 100$ | | | | | | | | | | | | | | | | | |
| Test / Size | 1% | 5% | 10% | 1% | 5% | 10% | 1% | 5% | 10% | 1% | 5% | 10% | 1% | 5% | 10% | 1% | 5% | 10% |
| $k_{\theta,n}$ | 35 | | | 200 | | | 482 | | | 35 | | | 200 | | | 482 | | |
| $p$-Max | .013 | .036 | .078 | .006 | .027 | .079 | .003 | .020 | .055 | .016 | .090 | .155 | .005 | .070 | .134 | .007 | .040 | .116 |
| $p$-Max-t | .009 | .058 | .113 | .011 | .067 | .138 | .014 | .062 | .132 | .014 | .122 | .218 | .035 | .133 | .248 | .028 | .139 | .203 |
| | $n = 250$ | | | | | | | | | | | | | | | | | |
| $k_{\theta,n}$ | 35 | | | 1144 | | | 1250 | | | 35 | | | 1144 | | | 1250 | | |
| $p$-Max | .009 | .050 | .105 | .005 | .040 | .074 | .004 | .026 | .058 | .009 | .051 | .107 | .005 | .035 | .097 | .008 | .053 | .096 |
| $p$-Max-t | .013 | .059 | .105 | .019 | .052 | .101 | .009 | .047 | .108 | .008 | .065 | .132 | .018 | .068 | .127 | .012 | .080 | .145 |
| | $n = 500$ | | | | | | | | | | | | | | | | | |
| $k_{\theta,n}$ | 35 | | | 2381 | | | 5000 | | | 35 | | | 2381 | | | 5000 | | |
| $p$-Max | .006 | .053 | .097 | .003 | .036 | .073 | .005 | .032 | .069 | .006 | .050 | .103 | .006 | .044 | .088 | .006 | .043 | .092 |
| $p$-Max-t | .007 | .056 | .101 | .009 | .046 | .103 | .008 | .053 | .109 | .008 | .056 | .103 | .012 | .055 | .113 | .010 | .063 | .117 |

Table 11: Robustness Check a.iii: $V[x_{\theta,1,t}] = 100$, and $\{\{V[x_{\theta,i,t}]\}_{i=2}^{\mathring{k}_{n,\theta}}, V[x_{\delta,i,t}]\} = 1$

| | $k_\delta = 0$ | | | | | | | | | $k_\delta = 10$ | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | $n = 100$ | | | | | | | | | | | | | | | | | |
| Test / Size | 1% | 5% | 10% | 1% | 5% | 10% | 1% | 5% | 10% | 1% | 5% | 10% | 1% | 5% | 10% | 1% | 5% | 10% |
| $k_{\theta,n}$ | 35 | | | 200 | | | 482 | | | 35 | | | 200 | | | 482 | | |
| $p$-Max | .005 | .026 | .076 | .001 | .019 | .064 | .005 | .024 | .056 | .014 | .072 | .138 | .004 | .054 | .145 | .004 | .054 | .126 |
| $p$-Max-t | .008 | .054 | .122 | .015 | .067 | .131 | .015 | .057 | .134 | .025 | .108 | .175 | .028 | .124 | .231 | .023 | .138 | .248 |
| | $n = 250$ | | | | | | | | | | | | | | | | | |
| $k_{\theta,n}$ | 35 | | | 1144 | | | 1250 | | | 35 | | | 1144 | | | 1250 | | |
| $p$-Max | .006 | .038 | .079 | .003 | .022 | .079 | .003 | .024 | .071 | .015 | .057 | .122 | .013 | .053 | .117 | .007 | .042 | .089 |
| $p$-Max-t | .008 | .051 | .091 | .008 | .047 | .125 | .008 | .046 | .113 | .018 | .066 | .132 | .016 | .085 | .163 | .011 | .066 | .138 |
| | $n = 500$ | | | | | | | | | | | | | | | | | |
| $k_{\theta,n}$ | 35 | | | 2381 | | | 5000 | | | 35 | | | 2381 | | | 5000 | | |
| $p$-Max | .008 | .055 | .111 | .013 | .042 | .091 | .004 | .032 | .066 | .009 | .056 | .108 | .008 | .046 | .099 | .007 | .048 | .097 |
| $p$-Max-t | .008 | .056 | .119 | .013 | .049 | .110 | .008 | .048 | .095 | .013 | .060 | .119 | .014 | .060 | .127 | .013 | .062 | .124 |

Table 12: Robustness Check b: $k_\delta \in \{20, 40\}$, for $n \in \{100, 250, 500\}$

| | $k_\delta = 20$ | | | | | | | | | $k_\delta = 40$ | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $n = 100$ | | | | | | | | | | | | | | | | | | |
| Test / Size | 1% | 5% | 10% | 1% | 5% | 10% | 1% | 5% | 10% | 1% | 5% | 10% | 1% | 5% | 10% | 1% | 5% | 10% |
| $k_{\theta,n}$ | | 35 | | | 200 | | | 482 | | | 35 | | | 200 | | | 482 | |
| $p$-Max | .023 | .116 | .237 | .014 | .114 | .249 | .010 | .118 | .243 | .107 | .335 | .491 | .120 | .406 | .649 | .125 | .498 | .727 |
| $p$-Max-t | .042 | .116 | .305 | .058 | .233 | .385 | .058 | .255 | .450 | .183 | .399 | .579 | .304 | .668 | .837 | .401 | .807 | .929 |
| $n = 250$ | | | | | | | | | | | | | | | | | | |
| $k_{\theta,n}$ | | 35 | | | 1144 | | | 1250 | | | 35 | | | 1144 | | | 1250 | |
| $p$-Max | .016 | .075 | .151 | .007 | .071 | .145 | .010 | .064 | .144 | .026 | .102 | .190 | .019 | .131 | .259 | .020 | .137 | .259 |
| $p$-Max-t | .014 | .093 | .167 | .017 | .095 | .205 | .024 | .112 | .203 | .036 | .118 | .216 | .051 | .211 | .354 | .054 | .220 | .361 |
| $n = 500$ | | | | | | | | | | | | | | | | | | |
| $k_{\theta,n}$ | | 35 | | | 2381 | | | 5000 | | | 35 | | | 2381 | | | 5000 | |
| $p$-Max | .013 | .060 | .117 | .004 | .056 | .114 | .012 | .057 | .108 | .016 | .070 | .145 | .019 | .079 | .156 | .017 | .079 | .152 |
| $p$-Max-t | .015 | .069 | .121 | .014 | .073 | .153 | .015 | .064 | .122 | .020 | .075 | .156 | .028 | .113 | .208 | .024 | .117 | .212 |

# References

Bühlmann, P., and S. Van De Geer (2011): *Statistics for High-Dimensional Data: Methods, Theory and Applications*. Springer, Heidelberg.

Cattaneo, M. D., M. Jansson, and W. Newey (2018): "Inference in Linear Regression Models with Many Covariates and Heteroscedasticity," *Journal of the American Statistical Association*, 113, 1350–1361.

Chernozhukov, V., D. Chetverikov, and K. Kato (2013): "Gaussian Approximations and Multiplier Bootstrap for Maxima of Sums of High-Dimensional Random Vectors," *Annals of Statistics*, 41, 2786–2819.

——— (2015): "Comparison and Anti-Concentration Bounds for Maxima of Gaussian Random Vectors," *Probability Theory and Related Fields*, 162, 47–70.

de Haan, L., and A. Ferreira (2000): *Extreme Value Theory*. Springer.

Gine, E., and J. Zinn (1990): "Bootstrapping General Empirical Measures," *Annals of Probability*, 18, 851–869.

Vershynin, R. (2018): *High-Dimensional Probability*. Cambridge University Press, Cambridge, UK.