Optimal Data-Driven Regression Discontinuity Plots*

Supplemental Appendix

Sebastian Calonico† Matias D. Cattaneo‡ Rocio Titiunik§

November 25, 2015

Abstract

This supplemental appendix contains the proofs of our main theorems, additional methodological and technical results, detailed simulation evidence, and further empirical illustrations not included in the main paper to conserve space.

*Financial support from the National Science Foundation (SES 1357561) is gratefully acknowledged.
†Department of Economics, University of Miami.
‡Department of Economics and Department of Statistics, University of Michigan.
§Department of Political Science, University of Michigan.

Contents

1 Implied Weights in Optimal WIMSE Approach **2**

2 Proofs of Main Theorems **3**

2.1 Lemma SA1 3

2.2 Lemma SA2 5

2.3 Lemma SA3 6

2.4 Proof of Theorem 1 8

2.5 Proof of Theorem 2 10

2.6 Proof of Theorem 3 11

2.7 Proof of Remark 1 11

2.8 Proof of Theorem 4 and Remark 2 12

3 Data-Driven Implementations with Arbitrary $w(x)$ **12**

3.1 Evenly Spaced RD Plots 13

3.2 Quantile Spaced RD Plots 16

4 Other Empirical Applications **18**

4.1 U.S. Senate Data 18

4.2 Progresa/Oportunidades Data 19

4.3 Head Start Data 20

4.4 Summary of Results 20

5 Complete Simulation Results **28**

6 Numerical Comparison of Partitioning Schemes **48**

1 Implied Weights in Optimal WIMSE Approach

Recall from the main paper that the optimal choices of number of bins based on a WIMSE can be written as

$$J_{\texttt{ES-}\omega,-,n} = \lceil \omega_- J_{\texttt{ES-}\mu,-,n} \rceil \quad \text{and} \quad J_{\texttt{ES-}\omega,+,n} = \lceil \omega_+ J_{\texttt{ES-}\mu,+,n} \rceil,$$

where $J_{\texttt{ES-}\mu,-,n}$ and $J_{\texttt{ES-}\mu,+,n}$ denote the IMSE-optimal choices and $\omega_- = (\omega_{\mathscr{B},-}/\omega_{\mathscr{V},-})^{1/3}$ and $\omega_+ = (\omega_{\mathscr{B},+}/\omega_{\mathscr{V},+})^{1/3}$. As discussed in the paper, this result may be used to justify ad-hoc rescalings chosen by the researchers when using the IMSE-optimal choices as a starting point. In particular, given a choice of rescaling factors ω_- and ω_+, we have:

$$(\omega_{\mathscr{V},-}, \omega_{\mathscr{B},-}) = \left(\frac{1}{1+\omega_-^3}, \frac{\omega_-^3}{1+\omega_-^3}\right) \qquad \text{and} \qquad (\omega_{\mathscr{V},+}, \omega_{\mathscr{B},+}) = \left(\frac{1}{1+\omega_+^3}, \frac{\omega_+^3}{1+\omega_+^3}\right),$$

which are the resulting weights entering the WIMSE objective function that would be compatible with such choices of rescale constants for the IMSE-optimal number of bins.

To gain some intuition on the relative weights emerging from manual rescaling of the IMSE-optimal choice, we present the implied weights in the optimal WIMSE approach for different, common choices of rescaling constants ω:

ω	$\omega_{\mathscr{V}}$	$\omega_{\mathscr{B}}$
0.1	0.999	0.001
0.2	0.992	0.008
0.5	0.889	0.111
1	0.500	0.500
2	0.111	0.889
5	0.008	0.992
10	0.001	0.999

As expected, the larger ω the smaller the weight on variance ($\omega_{\mathscr{V}}$) and the larger the weight on bias ($\omega_{\mathscr{B}}$) in the WIMSE objective function. Our software implementations in `R` and `Stata` compute this weights explicitly as part of the standard output; see Calonico, Cattaneo and Titiunik (2014a, 2015) for further details.

2 Proofs of Main Theorems

We state and prove results only for the treatment group (subindex "+") because for the control group the results and proofs are analogous. Here we only provide short, self-contained proofs of the main results presented in the paper. To this end, we first state three preliminary technical lemmas. We also offer short proofs of these lemmas, and provide references to the underlying results not reproduced here to conserve space.

Recall that the lower and upper end points of $P_{+,j}$ are denoted, respectively, by $p_{+,j-1}$ and $p_{+,j}$ for $j = 1, 2, \cdots, J_{+,n}$, which are nonrandom under ES partitioning and random under QS partitioning. Let $\bar{p}_{+,j} = (p_{+,j} + p_{+,j-1})/2$ be the middle point of bin $P_{+,j}$. Throughout the supplemental appendix C denotes an arbitrary positive, bounded constant taking different values in different places.

2.1 Lemma SA1

This lemma holds for any nonrandom partition $\mathcal{P}_{+,n}$ satisfying

$$\frac{C_1}{J_{+,n}} \leq \min_{1\leq j\leq J_{+,n}} |p_{+,j} - p_{+,j-1}| \leq \max_{1\leq j\leq J_{+,n}} |p_{+,j} - p_{+,j-1}| \leq \frac{C_2}{J_{+,n}},$$

for fixed positive constants C_1 and C_2. In particular, it holds for $\mathcal{P}_{\texttt{ES},+,n}$.

Note also that Lemma SA1(i) shows that $\mathbb{P}(N_{+,j} > 0) \to 1$ uniformly in j, which guarantees that the estimators for the ES partitioning scheme are well-behaved in large samples.

Lemma SA1. *Let Assumption 1 hold. For $\mathcal{P}_{\texttt{ES},+,n}$, if*

$$\frac{J_{+,n}\log(J_{+,n})}{n} \to 0 \qquad \textit{and} \qquad J_{+,n} \to \infty,$$

then the following results hold.

(i) $$\max_{1\le j\le J_{+,n}} |\mathbb{1}(N_{+,j}>0)-1| = o_{\mathbb{P}}(1).$$

(ii) $$\max_{1\le j\le J_{+,n}} |N_{+,j}/n - \mathbb{P}[X_i\in P_{+,j}]| = o_{\mathbb{P}}(J_{+,n}^{-1}).$$

(iii) $$\max_{1\le j\le J_{+,n}} \left|\frac{1}{n}\sum_{i=1}^{n}\mathbb{1}_{P_{+,j}}(X_i)\frac{X_i-\bar{p}_{+,j}}{p_{+,j}-p_{+,j-1}} - \mathbb{E}\left[\mathbb{1}_{P_{+,j}}(X_i)\frac{X_i-\bar{p}_{+,j}}{p_{+,j}-p_{+,j-1}}\right]\right| = o_{\mathbb{P}}(J_{+,n}^{-1}).$$

(iv) $$\max_{1\le j\le J_{+,n}} \left|\mathbb{E}\left[\mathbb{1}_{P_{+,j}}(X_i)\frac{X_i-\bar{p}_{+,j}}{p_{+,j}-p_{+,j-1}}\right]\right| = o(J_{+,n}^{-1}).$$

Proof of Lemma SA1. The proof of this lemma is very similar to the results given in the supplemental appendix of Cattaneo and Farrell (2013). Part (i) follows by properties of the Binomial distribution and simple bounding arguments, under the assumptions imposed. For part (ii), note that $\mathbb{E}[\mathbb{1}(X_i\in P_{+,j})] = \mathbb{P}[X_i\in P_{+,j}] = O(J_{+,n}^{-1})$ and $C_1/J_{+,n} \le \mathbb{V}[\mathbb{1}(X_i\in P_{+,j})] \le C_2/J_{+,n}$, uniformly in $j=1,2,\cdots,J_{+,n}$. For any $\varepsilon>0$, and using Bernstein inequality, we have

$$\begin{aligned}
&\mathbb{P}\left[J_{+,n}\max_{1\le j\le J_{+,n}}\left|\frac{N_j}{n}-\mathbb{P}[X_i\in P_{+,j}]\right|>\varepsilon\right]\\
&\quad\le J_{+,n}\max_{1\le j\le J_{+,n}}\mathbb{P}\left[\left|\sum_{i=1}^{n}(\mathbb{1}(X_i\in P_{+,j})-\mathbb{P}[X_i\in P_{+,j}])\right|>n\varepsilon/J_{+,n}\right]\\
&\quad\le J_{+,n}\max_{1\le j\le J_{+,n}} 2\exp\left\{-\frac{n^2\varepsilon^2/J_{+,n}^2}{2\sum_{i=1}^{n}\mathbb{V}[\mathbb{1}(X_i\in P_{+,j})]+2n\varepsilon}\right\}\\
&\quad\le C\exp\left\{-\frac{Cn}{J_{+,n}+J_{+,n}^2\varepsilon}+\log(J_{+,n})\right\}\le C\exp\left\{-\frac{Cn}{J_{+,n}}+\log(J_{+,n})\right\}\to 0,
\end{aligned}$$

provided that $J_{+,n}\log(J_{+,n})/n\to\infty$. Part (iii) follows by similar arguments.

Finally, to verify part (iv), using change of variables we obtain

$$\begin{aligned}
&\max_{1\le j\le J_{+,n}}\left|\mathbb{E}\left[\mathbb{1}_{P_{+,j}}(X_i)\frac{X_i-\bar{p}_{+,j}}{p_{+,j}-p_{+,j-1}}\right]\right|\\
&\quad=\max_{1\le j\le J_n}\left|\int_{\bar{x}}^{x_u}\mathbb{1}_{P_{+,j}}(x)\frac{x-\bar{p}_{+,j}}{p_{+,j}-p_{+,j-1}}f(x)dx\right|\\
&\quad=\max_{1\le j\le J_{+,n}}(p_{+,j}-p_{+,j-1})\left|\int_{-1}^{1}uf(u(p_{+,j}-p_{+,j-1})+\bar{p}_{+,j})du\right|\\
&\quad=\max_{1\le j\le J_{+,n}}\frac{x_u-\bar{x}}{J_{+,n}}\left|\int_{-1}^{1}uf(\bar{p}_{+,j})du+o(1)\right|,
\end{aligned}$$

and the result follows. $\square$

2.2 Lemma SA2

This second lemma characterizes the properties of the random partitioning scheme based on quantile estimates. These results will be used when handling the partitioning scheme $\mathcal{P}_{\texttt{QS},+,n}$: recall that $p_{+,j} = \hat{F}_+^{-1}(j/J_{+,n})$ in this case, $j = 1, 2, \cdots, J_{+,n}$, and thus set $q_{+,j} = F_+^{-1}(j/J_{+,n})$ with $F_+^{-1}(y) = \inf\{x : F_+(x) \geq y\}$ with

$$F_+(x) = \frac{\mathbb{P}[X_i \leq x, X_i \geq \bar{x}]}{\mathbb{P}[X_i \geq \bar{x}]} = F(x|X_i \geq \bar{x}).$$

Lemma SA2. *Let Assumption 1 hold. For $\mathcal{P}_{\texttt{QS},+,n}$, if*

$$\frac{J_{+,n}\log(J_{+,n})}{n} \to 0 \qquad \textit{and} \qquad \frac{J_{+,n}}{\log(n)} \to \infty,$$

then the following results hold.

$$\textbf{(i)} \qquad \max_{1\leq j\leq J_{+,n}} |N_{+,j}/N_+ - 1/J_{+,n}| = o_{\mathbb{P}}(J_{+,n}^{-1}).$$

$$\textbf{(ii)} \qquad \max_{1\leq j\leq J_{+,n}} |p_{+,j} - p_{+,j-1} - (q_{+,j} - q_{+,j-1})| = o_{\mathbb{P}}(J_{+,n}^{-1}).$$

Proof of Lemma SA2. Because the sample size N_+ is random, we employ the following result: if $N_+ \to_{\text{as}} \infty$ and $Z_n \to_{\text{as}} Z_\infty$, then $Z_{N_+} \to_{\text{as}} Z_\infty$. In our case, $N_+ = \sum_{i=1}^n \mathbb{1}(X_i \geq \bar{x})$ and thus $N_+/n \to_{\text{as}} P_+$. Hence, it suffices to assume $N_+ \to \infty$ is not random, but we need to prove the statements in an almost sure sense. The rest of the proof takes limits as $N_+ \to \infty$.

Part (i) now follows from properties of distribution function and quantile processes (e.g., Shorack and Wellner, 2009). Using continuity and boundedness of $f(x)$, we have

$$\begin{aligned} N_{+,j} &= \sum_{i=1}^n \mathbb{1}\left(\hat{F}_+^{-1}\left(\frac{j-1}{J_{+,n}}\right) \leq X_i < \hat{F}_+^{-1}\left(\frac{j}{J_{+,n}}\right)\right) \\ &= N_+\hat{F}_+\left(\hat{F}_+^{-1}\left(\frac{j}{J_{+,n}}\right)\right) - N_+\hat{F}_+\left(\hat{F}_+^{-1}\left(\frac{j-1}{J_{+,n}}\right)\right)\{1 + o_{\text{as}}(1)\} = \frac{N_+}{J_{+,n}}\{1 + o_{\text{as}}(1)\}, \end{aligned}$$

uniformly in $j = 1, 2, \cdots, J_{+,n}$, under the rate restrictions imposed.

Similarly, part (ii) follows from properties of the modulus of continuity of the sample quantile process (e.g., Mason (1984) and Shorack and Wellner (2009, Chapter 14)). We have

$$\begin{aligned}
&\max_{1\leq j\leq J_{+,n}} |p_{+,j} - p_{+,j-1} - (q_{+,j} - q_{+,j-1})| \\
&= \max_{1\leq j\leq J_{+,n}} \left| \hat{F}_{+}^{-1}\left(\frac{j}{J_{+,n}}\right) - F_{+}^{-1}\left(\frac{j}{J_{+,n}}\right) - \left(\hat{F}_{+}^{-1}\left(\frac{j-1}{J_{+,n}}\right) - F_{+}^{-1}\left(\frac{j-1}{J_{+,n}}\right)\right)\right| = o_{\mathrm{as}}(J_{+,n}^{-1}),
\end{aligned}$$

under the rate restrictions imposed. □

2.3 Lemma SA3

Our final third technical lemma gives the main convergence results for the spacings estimators used to construct data-driven choices of partition sizes. We employ the notation introduced in Section 5 of the main paper.

Lemma SA3. *Let Assumption 1 hold, and set $\ell \in \mathbb{Z}_+$. If $Y_i(1)$ is continuously distributed and $g : [\bar{x}, x_u] \to \mathbb{R}_+$ is continuous, then the following results hold.*

(i) $$N_{+}^{\ell-1}\sum_{i=2}^{N_+}(X_{+,(i)} - X_{+,(i-1)})^{\ell} g(\bar{X}_{+,(i)}) \to_{\mathbb{P}} \ell! P_{+}^{\ell-1}\int_{\bar{x}}^{x_u} f(x)^{1-\ell} g(x)dx.$$

(ii) $$N_{+}^{\ell-1}\sum_{i=2}^{N_+}(X_{+,(i)} - X_{+,(i-1)})^{\ell}(Y_{+,[i]} - Y_{+,[i-1]})^2 g(\bar{X}_{+,(i)}) \to_{\mathbb{P}} \ell! P_{+}^{\ell-1} 2\int_{\bar{x}}^{x_u} f(x)^{1-\ell}\sigma_{+}^{2}(x) g(x)dx.$$

Proof of Lemma SA3. We prove the result assuming that N_+ is nonrandom, and thus limits are taken as $N_+ \to \infty$. Set $U_i = F_+(X_{+,i}) \sim \text{Uniform}(0,1)$ and $U_{(i)} = F_+(X_{+,(i)})$, $i = 1, \cdots, N_+$. Recall that $\{N_+(U_{(i)} - U_{(i-1)}) : i = 2, \cdots, N_+\} =_{\mathrm{d}} \{E_i/\bar{E} : i = 2, \cdots, N_+\}$, where $\{E_i : i = 2, \cdots, N_+\}$ i.i.d. random variables with $E_i \sim \text{Exponential}(1)$ and $\bar{E} = \sum_{i=2}^{N_+} E_i/N_+$, and where $Z_1 =_{\mathrm{d}} Z_2$ denotes that Z_1 and Z_2 have the same probability law. Set $\bar{u}_i = (i-1/2)/N_+$ and recall that $\max_{2\leq i\leq N_+}\sup_{U_{(i-1)}\leq u\leq U_{(i)}} |u - \bar{u}_i| \to_{\mathbb{P}} 0$.

For part (i), using the above, $N_+^{-1}\sum_{i=2}^{N_+} E_i^{\ell} \to_{\mathbb{P}} \mathbb{E}[E_i^{\ell}] = \ell!$, and uniform continuity of $g(\cdot)$ and

$f(\cdot)$,

$$
\begin{aligned}
N_+^{\ell-1}&\sum_{i=2}^{N_+}(X_{+,(i)}-X_{+,(i-1)})^\ell g(\bar{X}_{+,(i)})\\
&=\frac{1}{N_+}\sum_{i=2}^{N_+}(N_+(U_{(i)}-U_{(i-1)}))^\ell \frac{g(F_+^{-1}(u_{n,i}))}{f_+(F_+^{-1}(u_{n,i}))^\ell}\{1+o_{\mathbb{P}}(1)\}\\
&=_{\mathrm{d}} \frac{1}{N_+}\sum_{i=2}^{N_+}\left(\frac{E_i}{\bar{E}}\right)^\ell \frac{g(F_+^{-1}(u_{n,i}))}{f_+(F_+^{-1}(u_{n,i}))^\ell}\{1+o_{\mathbb{P}}(1)\}\\
&=\frac{1}{N_+}\sum_{i=2}^{N_+}\mathbb{E}[E_i^\ell]\frac{g(F_+^{-1}(u_{n,i}))}{f_+(F_+^{-1}(u_{n,i}))^\ell}\{1+o_{\mathbb{P}}(1)\}\\
&\to_{\mathbb{P}} \ell!\int_0^1 \frac{g(F_+^{-1}(u))}{f_+(F_+^{-1}(u))^\ell}du,
\end{aligned}
$$

and the result follows by change of variables and because $f_+(x)=f(x)\mathbb{1}(x\geq \bar{x})/P_+$. This result implies, in particular, $\sum_{i=2}^{N_+}(X_{+,(i)}-X_{+,(i-1)})^\ell g(\bar{X}_{+,(i)})=O_{\mathbb{P}}(N_+^{1-\ell})$.

For part (ii), let $\mathbf{X}_{(+)}=(X_{+,(1)},X_{+,(2)},\cdots,X_{+,(N_+)})$. Recall that $(Y_{+,[1]},Y_{+,[2]},\cdots,Y_{+,[N_+]})$ are independent conditional on $\mathbf{X}_{(+)}$ and $\mathbb{E}[g(Y_{+,[i]})|\mathbf{X}_{(+)}]=\mathbb{E}[g(Y_{+,[i]})|X_{+,(i)}]=G(X_{+,(i)})$ with $G(x)=\mathbb{E}[g(Y_{+,i})|X_{+,i}=x]$. Therefore, $\mathbb{E}[(Y_{+,[i]}-Y_{+,[i-1]})^2|\mathbf{X}_{(+)}]=\sigma_+^2(X_{+,(i)})+\sigma_+^2(X_{+,(i-1)})+(\mathbb{E}[Y_{+,[i]}|\mathbf{X}_{(+)}]-\mathbb{E}[Y_{+,[i-1]}|\mathbf{X}_{(+)}])^2=\sigma_+^2(X_{+,(i)})+\sigma_+^2(X_{+,(i-1)})+O_{\mathbb{P}}(N_+^{-2})$, uniformly in i. This gives

$$
N_+^{\ell-1}\sum_{i=2}^{N_+}(X_{+,(i)}-X_{+,(i-1)})^\ell (Y_{+,[i]}-Y_{+,[i-1]})^2 g(\bar{X}_{+,(i)})=T_1+T_2,
$$

with

$$
T_1=N_+^{\ell-1}\sum_{i=2}^{N_+}(X_{+,(i)}-X_{+,(i-1)})^\ell(\sigma_+^2(X_{+,[i]})+\sigma_+^2(X_{+,[i-1]}))g(\bar{X}_{+,(i)})+o_{\mathbb{P}}(1),
$$

$$
T_2=N_+^{\ell-1}\sum_{i=2}^{N_+}(X_{+,(i)}-X_{+,(i-1)})^\ell\left[(Y_{+,[i]}-Y_{+,[i-1]})^2-\mathbb{E}[(Y_{+,[i]}-Y_{+,[i-1]})^2|\mathbf{X}_{[+]}]\right]g(\bar{X}_{+,(i)}).
$$

Noting that $\sigma_+^2(X_{+,(i)})+\sigma_+^2(X_{+,(i-1)})=2\sigma_+^2(\bar{X}_{+,(i)})\{1+o_{\mathbb{P}}(1)\}$, uniformly in i, it follows that $T_1\to_{\mathbb{P}}\ell! P_+^{\ell-1}2\int_{\bar{x}}^{x_u}f(x)^{1-\ell}\sigma_+^2(x)g(x)dx$, as in part (i). Thus, it remains to show that $T_2\to_{\mathbb{P}} 0$. To this end, first define $\tilde{Y}_i=(Y_{+,[i]}-Y_{+,[i-1]})^2-\mathbb{E}[(Y_{+,[i]}-Y_{+,[i-1]})^2|\mathbf{X}_{(+)}]$, and note that

$\mathbb{E}[\tilde{Y}_i, \tilde{Y}_{i-s}|\mathbf{X}_{(+)}] = 0$ whenever $s \geq 2$, which implies

$$\mathbb{V}[T_2|\mathbf{X}_{(+)}] \leq N_+^{2(\ell-1)} \sum_{i=2}^{N_+} (X_{+,(i)} - X_{+,(i-1)})^{2\ell} \mathbb{V}[\tilde{Y}_i|\mathbf{X}_{(+)}] g(\bar{X}_{+,(i)})^2$$
$$+ 2N_+^{2(\ell-1)} \sum_{i=2}^{N_+} (X_{+,(i)} - X_{+,(i-1)})^{\ell} (X_{+,(i-1)} - X_{+,(i-2)})^{\ell} \mathbb{E}[\tilde{Y}_i \tilde{Y}_{i-1}|\mathbf{X}_{(+)}] g(\bar{X}_{+,(i)}) g(\bar{X}_{+,(i-1)})$$
$$\leq C N_+^{-1},$$

and the result follows by the dominated convergence theorem.

The random sample size case ($N_+ = \sum_{i=1}^n \mathbb{1}(X_i \geq \bar{x})$) can be handled, for example, using the approach described in Aras et al. (1989) and references therein. □

2.4 Proof of Theorem 1

For the variance part, we have

$$\mathbb{V}[\hat{\mu}_+(x; J_{+,n})|\mathbf{X}_n] = \sum_{j=1}^{J_{+,n}} \frac{\mathbb{1}(N_{+,j} > 0) \mathbb{1}_{P_{+,j}}(x)}{N_{+,j}^2} \sum_{i=1}^{n} \mathbb{1}_{P_{+,j}}(X_i) \sigma_+^2(X_i),$$

and using uniform continuity of $w(\cdot)$ and $\sigma_+^2(\cdot)$ on $[\bar{x}, x_u]$ and Lemma SA1, we obtain

$$\int_{\bar{x}}^{x_u} \mathbb{V}[\hat{\mu}_+(x; J_{+,n})|\mathbf{X}_n] w(x) dx$$
$$= \sum_{j=1}^{J_{+,n}} \frac{\mathbb{1}(N_{+,j} > 0)}{N_{+,j}^2} \left(\int_{\bar{x}}^{x_u} \mathbb{1}_{P_{+,j}}(x) w(x) dx \right) \sum_{i=1}^{n} \mathbb{1}_{P_{+,j}}(X_i) \sigma_+^2(X_i)$$
$$= \sum_{j=1}^{J_{+,n}} \frac{\mathbb{1}(N_{+,j} > 0)}{N_{+,j}} (p_{+,j} - p_{+,j-1}) \sigma_+^2(\bar{p}_{+,j}) w(\bar{p}_{+,j}) \{1 + o_{\mathbb{P}}(1)\}$$
$$= \frac{1}{n} \sum_{j=1}^{J_{+,n}} \frac{\sigma_+^2(\bar{p}_{+,j}) w(\bar{p}_{+,j})}{f(\bar{p}_{+,j})} \{1 + o_{\mathbb{P}}(1)\},$$

because $\mathbb{P}[X_i \in P_{+,j}] = \int_{p_{+,j-1}}^{p_{+,j}} f(x)dx = (p_{+,j} - p_{+,j-1})f(\bar{p}_{+,j})\{1 + o(1)\}$ uniformly in j. Using properties of the Riemann integral it then follows that

$$
\begin{aligned}
\int_{\bar{x}}^{x_u} & \mathbb{V}[\hat{\mu}_{\mathtt{ES},+}(x; J_{+,n})|\mathbf{X}_n]w(x)dx \\
&= \frac{J_{+,n}}{n} \frac{1}{x_u - \bar{x}} \sum_{j=1}^{J_{+,n}} (p_{+,j} - p_{+,j-1}) \frac{\sigma_+^2(\bar{p}_{+,j})w(\bar{p}_{+,j})}{f(\bar{p}_{+,j})} \{1 + o_{\mathbb{P}}(1)\} \\
&= \frac{J_{+,n}}{n} \frac{1}{x_u - \bar{x}} \int_{\bar{x}}^{x_u} \frac{\sigma_+^2(x)}{f(x)} w(x)dx \{1 + o_{\mathbb{P}}(1)\} \\
&= \frac{J_{+,n}}{n} \mathscr{V}_{\mathtt{ES},+} \{1 + o_{\mathbb{P}}(1)\},
\end{aligned}
$$

because $p_{+,j+1} - p_{+,j} = (x_u - \bar{x})/J_{+,n}$ for the evenly spaced partition.

Next, for the bias term, note that $\int_{\bar{x}}^{x_u} (\mathbb{E}[\hat{\mu}_+(x; J_n)|\mathbf{X}_n] - \mu_+(x))^2 w(x)dx = T_1 + T_2 + T_3$ with

$$
T_1 = \int_{\bar{x}}^{x_u} T_1(x)^2 w(x)dx, \qquad T_2 = \int_{\bar{x}}^{x_u} T_2(x)^2 w(x)dx, \qquad T_3 = 2\int_{\bar{x}}^{x_u} T_1(x)T_2(x)w(x)dx,
$$

$$
T_1(x) = \sum_{j=1}^{J_{+,n}} \mathbb{1}_{P_{+,j}}(x)(\mathbb{1}(N_{+,j} > 0)\mu_+(\bar{p}_{+,j}) - \mu_+(x)),
$$

$$
T_2(x) = \sum_{j=1}^{J_{+,n}} \mathbb{1}_{P_{+,j}}(x) \frac{\mathbb{1}(N_{+,j} > 0)}{N_{+,j}} \left(\sum_{i=1}^{n} \mathbb{1}_{P_{+,j}}(X_i)(\mu_+(X_i) - \mu_+(\bar{p}_{+,j})) \right).
$$

Using uniform continuity of $\mu_+(\cdot)$ and $w(\cdot)$ on $[\bar{x}, x_u]$ and Lemma SA1, we obtain

$$
\begin{aligned}
T_1 &= \frac{1}{12} \sum_{j=1}^{J_{+,n}} \left(\mu_+^{(1)}(\bar{p}_{+,j})\right)^2 w(\bar{p}_{+,j}) \int_{\bar{x}}^{x_u} \mathbb{1}_{P_{+,j}}(x)(\bar{p}_{+,j} - x)^2 dx \{1 + o_{\mathbb{P}}(1)\} \\
&= \frac{1}{12} \sum_{j=1}^{J_{+,n}} (p_{+,j} - p_{+,j-1})^3 \left(\mu_+^{(1)}(\bar{p}_{+,j})\right)^2 w(\bar{p}_{+,j}) \{1 + o_{\mathbb{P}}(1)\} \\
&= \frac{1}{J_{+,n}^2} \frac{(x_u - \bar{x})^2}{12} \int_{\bar{x}}^{x_u} \left(\mu_+^{(1)}(x)\right)^2 w(x)dx \{1 + o_{\mathbb{P}}(1)\} = J_{+,n}^{-2} \mathscr{B}_{\mathtt{ES},+} \{1 + o_{\mathbb{P}}(1)\},
\end{aligned}
$$

because $\int_a^b ((a+b)/2 - x)^2 dx = (b-a)^3/12$ and $p_{+,j+1} - p_{+,j} = (x_u - \bar{x})/J_{+,n}$ for the evenly spaced partition. This implies that $T_1 = O_{\mathbb{P}}(J_{+,n}^{-2})$. Thus, to finish the proof, we show that $T_2 = o_{\mathbb{P}}(J_{+,n}^{-2})$ and $T_3 = o_{\mathbb{P}}(J_{+,n}^{-2})$. For T_2, using uniform continuity of $\mu_+(\cdot)$ and $w(\cdot)$ on $[\bar{x}, x_u]$ and Lemma SA1 we have

$$|T_2| \leq C \sum_{j=1}^{J_{+,n}} \frac{\mathbb{1}(N_{+,j} > 0)}{J_{+,n}^2 N_{+,j}^2/n^2} \left(\frac{1}{n} \sum_{i=1}^{n} \mathbb{1}_{P_{+,j}}(X_i) \frac{X_i - \bar{p}_{+,j}}{p_{+,j} - p_{+,j-1}} \right)^2 \{1 + o_{\mathbb{P}}(1)\} = o_{\mathbb{P}}(J_{+,n}^{-2}),$$

while, for T_3, Cauchy-Swartz inequality implies $|T_3| \leq \sqrt{T_1}\sqrt{T_2} = O_{\mathbb{P}}(J_{+,n}^{-1}) o_{\mathbb{P}}(J_{+,n}^{-1}) = o_{\mathbb{P}}(J_{+,n}^{-2})$. $\square$

2.5 Proof of Theorem 2

Recall that $p_{+,j} = \hat{F}_{+}^{-1}(j/J_{+,n})$ and $q_{+,j} = F_{+}^{-1}(j/J_{+,n})$. If $J_{+,n} < N_+$, then $\mathbb{1}(N_{+,j} > 0) = 1$, but now the partitioning scheme $\mathcal{P}_{\texttt{QS},+,n}$ is random. For the variance part, letting $\bar{q}_{+,j} = (q_{+,j} + q_{+,j-1})/2$, we have

$$\begin{aligned}
&\int_{\bar{x}}^{x_u} \mathbb{V}[\hat{\mu}_{\texttt{QS},+}(x; J_{+,n})|\mathbf{X}_n] w(x) dx \\
&\quad = \sum_{j=1}^{J_{+,n}} \frac{1}{N_{+,j}^2} \left(\int_{\bar{x}}^{x_u} \mathbb{1}_{P_{+,j}}(x) w(x) dx \right) \sum_{i=1}^{n} \mathbb{1}_{P_{+,j}}(X_i) \sigma_+^2(X_i) \\
&\quad = \frac{J_{+,n}}{N_+} \sum_{j=1}^{J_{+,n}} (p_{+,j} - p_{+,j-1}) \sigma_+^2(\bar{p}_{+,j}) w(\bar{p}_{+,j}) \{1 + o_{\mathbb{P}}(1)\} \\
&\quad = \frac{J_{+,n}}{N_+} \sum_{j=1}^{J_{+,n}} (q_{+,j} - q_{+,j-1}) \sigma_+^2(\bar{q}_{+,j}) w(\bar{q}_{+,j}) \{1 + o_{\mathbb{P}}(1)\} \\
&\quad = \frac{J_{+,n}}{n} \frac{1}{P_+} \int_{\bar{x}}^{x_u} \sigma_+^2(x) w(x) dx \{1 + o_{\mathbb{P}}(1)\} = \frac{J_{+,n}}{n} \mathscr{V}_{\texttt{QS},+} \{1 + o_{\mathbb{P}}(1)\},
\end{aligned}$$

using Lemma SA2 and properties of the Riemann integral.

For the bias part, using the previous results and proceeding as in the proof of Theorem 1,

$$\begin{aligned}
&\int_{\bar{x}}^{x_u} (\mathbb{E}[\hat{\mu}_{\texttt{QS},+}(x; J_n)|\mathbf{X}_n] - \mu_+(x))^2 w(x) dx \\
&\quad = \frac{1}{12} \sum_{j=1}^{J_{+,n}} (p_{+,j} - p_{+,j-1})^3 \left(\mu_+^{(1)}(\bar{p}_{+,j}) \right)^2 w(\bar{p}_{+,j}) \{1 + o_{\mathbb{P}}(1)\} \\
&\quad = \frac{1}{12} \sum_{j=1}^{J_{+,n}} (q_{+,j} - q_{+,j-1})^3 \left(\mu_+^{(1)}(\bar{q}_{+,j}) \right)^2 w(\bar{q}_{+,j}) \{1 + o_{\mathbb{P}}(1)\} \\
&\quad = \frac{1}{J_{+,n}^2} \frac{P_+^2}{12} \int_{\bar{x}}^{x_u} \left(\frac{\mu_+^{(1)}(x)}{f(x)} \right)^2 w(x) dx \{1 + o_{\mathbb{P}}(1)\} = J_{+,n}^{-2} \mathscr{B}_{\texttt{QS},+} \{1 + o_{\mathbb{P}}(1)\},
\end{aligned}$$

because, for quantile spaced partitions, expanding $F_+^{-1}(u)$ around $\bar{u} = F_+(\bar{q}_{+,j}) \in [(j-1)/J_{+,n}, j/J_{+,n}])$,

$$q_{+,j} - q_{+,j-1} = F_+^{-1}\left(\frac{j}{J_{+,n}}\right) - F_+^{-1}\left(\frac{j-1}{J_{+,n}}\right) = \frac{1}{f_+(\bar{q}_{+,j})}\frac{1}{J_{+,n}}\{1 + o_{\mathbb{P}}(1)\},$$

uniformly in $j = 1, 2, \cdots, J_{+,n}$, where $f_+(x) = \partial F_+(x)/\partial x = f(x)\mathbb{1}(x \geq \bar{x})/P_+$. □

2.6 Proof of Theorem 3

Using Lemma SA3 with $\ell = 1$ and $g(x) = 1$,

$$\hat{\mathscr{V}}_{\mathtt{ES},+} = \frac{1}{x_u - \bar{x}}\frac{1}{2}\sum_{i=2}^{N_+}(X_{+,(i)} - X_{+,(i-1)})(Y_{+,[i]} - Y_{+,[i-1]})^2 = \frac{1}{x_u - \bar{x}}\int_{\bar{x}}^{x_u}\sigma_+^2(x)dx + o_{\mathbb{P}}(1),$$

which gives $\hat{\mathscr{V}}_{\mathtt{ES},+} \to_{\mathbb{P}} \mathscr{V}_{\mathtt{ES},+}$. Next, note that for power series estimators, Newey (1997, Theorem 4) gives

$$\sup_{x\in[\bar{x},x_u]}|\hat{\mu}_{+,k_n}^{(1)}(x) - \mu_+^{(1)}(x)|^2 = O_{\mathbb{P}}(k_n^7/n + k_n^{-2S+8}) = o_{\mathbb{P}}(1).$$

Using this uniform consistency result we have

$$\begin{aligned}
\hat{\mathscr{B}}_{\mathtt{ES},+} &= \frac{(x_u - \bar{x})^2}{12n}\sum_{i=1}^{n}\mathbb{1}(X_i < \bar{x})\left(\hat{\mu}_{+,k_n}^{(1)}(X_i)\right)^2 = \frac{(x_u - \bar{x})^2}{12}\frac{1}{n}\sum_{i=1}^{n}\mathbb{1}(X_i < \bar{x})\left(\mu_+^{(1)}(X_i)\right)^2 + o_{\mathbb{P}}(1) \\
&= \frac{(x_u - \bar{x})^2}{12}\int_{\bar{x}}^{x_u}\left(\mu_+^{(1)}(x)\right)^2 w(x)dx + o_{\mathbb{P}}(1),
\end{aligned}$$

which gives $\hat{\mathscr{B}}_{\mathtt{ES},+} \to_{\mathbb{P}} \mathscr{B}_{\mathtt{ES},+}$.

Putting the above together, consistency of all the data-driven selectors follows. □

2.7 Proof of Remark 1

Note that for power series estimators, Newey (1997, Theorem 4) gives

$$\sup_{x\in[\bar{x},x_u]}|\hat{\mu}_{+,k_n,p}(x) - \mathbb{E}[Y(1)^p|X_i = x]|^2 = O_{\mathbb{P}}(k_n^3/n + k_n^{-2S+2}) = o_{\mathbb{P}}(1)$$

for $p = 1, 2$, under the assumptions imposed, which implies

$$\sup_{x\in[\bar{x},x_u]} |\hat{\sigma}^2_+(x) - \sigma^2_+|^2 = O_{\mathbb{P}}(k_n^3/n + k_n^{-2S+2}) = o_{\mathbb{P}}(1).$$

Using this result, and Lemma SA3 with $\ell = 1$ and $g(x) = \sigma^2_+(x)$,

$$\begin{aligned}
\check{\mathscr{V}}_{\texttt{ES},+} &= \frac{1}{x_u - \bar{x}} \sum_{i=2}^{N_+} (X_{+,(i)} - X_{+,(i-1)}) \hat{\sigma}^2_{+,k}(\bar{X}_{+,(i)}) \\
&= \frac{1}{x_u - \bar{x}} \sum_{i=2}^{N_+} (X_{+,(i)} - X_{+,(i-1)}) \sigma^2_{+,k}(\bar{X}_{+,(i)}) + o_{\mathbb{P}}(1) \to_{\mathbb{P}} \frac{1}{x_u - \bar{x}} \int_{\bar{x}}^{x_u} \sigma^2_+(x) dx = \mathscr{V}_{\texttt{ES},+}.
\end{aligned}$$

Combining this with Theorem SA1, the different consistency results follow. □

2.8 Proof of Theorem 4 and Remark 2

Proceeding as in the proofs of Theorem 3 and Remark 1, the results are established using Lemma SA3, $N_+/n \to_{\mathbb{P}} P_+$, and uniform consistency of power series estimators, as appropriate for each case. □

3 Data-Driven Implementations with Arbitrary $w(x)$

In this section we provide data-driven implementations for all of our number of bins selectors when $w(x)$ is taken as given. As discussed in the main text, we estimate the unknown constants using ideas related to spacings estimators whenever possible, but we also discuss series (polynomial) nonparametric regression estimates for completeness (to handle the non-continuous outcome case).

Recall the notation introduced in the main paper related to order statistics and concomitants. For a collection of continuous random variables $\{(Z_i, W_i) : i = 1, 2, \cdots, n\}$ we let $W_{(i)}$ be the i-th order statistic of W_i and $Z_{[i]}$ its corresponding concomitant. That is, $W_{(1)} < W_{(2)} < \cdots < W_{(n)}$ and $(Z_{[i]}, W_{(i)}) = (Z_i, W_{(i)})$ for all $i = 1, 2, \cdots, n$. Letting $\{(Y_{-,i}, X_{-,i}) : i = 1, 2, \cdots, N_-\}$ and $\{(Y_{+,i}, X_{+,i}) : i = 1, 2, \cdots, N_+\}$ be the subsamples of control $(X_i < \bar{x})$ and treatment $(X_i \geq \bar{x})$

units, respectively. We also have:

$$\bar{X}_{-,(i)} = \frac{X_{-,(i)} + X_{-,(i-1)}}{2}, \qquad i = 2, 3, \cdots, N_-, \qquad \hat{\mu}^{(1)}_{-,k}(x) = \mathbf{r}^{(1)}_k(x)'\hat{\boldsymbol{\beta}}_{-,k},$$

$$\bar{X}_{+,(i)} = \frac{X_{+,(i)} + X_{+,(i-1)}}{2}, \qquad i = 2, 3, \cdots, N_+, \qquad \hat{\mu}^{(1)}_{+,k}(x) = \mathbf{r}^{(1)}_k(x)'\hat{\boldsymbol{\beta}}_{+,k},$$

and $\mathbf{r}^{(1)}_k(x) = \partial\mathbf{r}_k(x)/\partial x = (0, 1, 2x, 3x^2, \cdots, kx^{k-1})'$.

3.1 Evenly Spaced RD Plots

For the case of ES RD Plots with generic $w(x)$ weighting scheme, we propose the following estimators:

$$\hat{\mathscr{V}}_{\texttt{ES},-} = \frac{1}{\bar{x} - x_l}\frac{n}{4}\sum_{i=2}^{N_-}(X_{-,(i)} - X_{-,(i-1)})^2(Y_{-,[i]} - Y_{-,[i-1]})^2 w(\bar{X}_{-,(i)}), \qquad \text{(SA-1)}$$

$$\hat{\mathscr{B}}_{\texttt{ES},-} = \frac{(\bar{x} - x_l)^2}{12}\sum_{i=2}^{N_-}(X_{-,(i)} - X_{-,(i-1)})\left(\hat{\mu}^{(1)}_{-,k}(\bar{X}_{-,[i]})\right)^2 w(\bar{X}_{-,(i)}), \qquad \text{(SA-2)}$$

and

$$\hat{\mathscr{V}}_{\texttt{ES},+} = \frac{1}{x_u - \bar{x}}\frac{n}{4}\sum_{i=2}^{N_+}(X_{+,(i)} - X_{+,(i-1)})^2(Y_{+,[i]} - Y_{+,[i-1]})^2 w(\bar{X}_{+,(i)}), \qquad \text{(SA-3)}$$

$$\hat{\mathscr{B}}_{\texttt{ES},+} = \frac{(x_u - \bar{x})^2}{12}\sum_{i=2}^{N_+}(X_{+,(i)} - X_{+,(i-1)})\left(\hat{\mu}^{(1)}_{+,k}(\bar{X}_{+,[i]})\right)^2 w(\bar{X}_{+,(i)}). \qquad \text{(SA-4)}$$

Thus, our proposed data-driven selectors for ES RD Plots take the form:

$$\hat{J}_{\texttt{ES-}\mu,-,n} = \left\lceil \left(\frac{2\hat{\mathscr{B}}_{\texttt{ES},-}}{\hat{\mathscr{V}}_{\texttt{ES},-}}\right)^{1/3} n^{1/3} \right\rceil \quad \text{and} \quad \hat{J}_{\texttt{ES-}\mu,+,n} = \left\lceil \left(\frac{2\hat{\mathscr{B}}_{\texttt{ES},+}}{\hat{\mathscr{V}}_{\texttt{ES},+}}\right)^{1/3} n^{1/3} \right\rceil, \qquad \text{(SA-5)}$$

$$\hat{J}_{\texttt{ES-}\omega,-,n} = \left\lceil \omega_- \left(\frac{2\hat{\mathscr{B}}_{\texttt{ES},-}}{\hat{\mathscr{V}}_{\texttt{ES},-}}\right)^{1/3} n^{1/3} \right\rceil \quad \text{and} \quad \hat{J}_{\texttt{ES-}\omega,+,n} = \left\lceil \omega_+ \left(\frac{2\hat{\mathscr{B}}_{\texttt{ES},+}}{\hat{\mathscr{V}}_{\texttt{ES},+}}\right)^{1/3} n^{1/3} \right\rceil, \qquad \text{(SA-6)}$$

$$\hat{J}_{\texttt{ES-}\vartheta,-,n} = \left\lceil \frac{\hat{\mathcal{V}}_-}{\hat{\mathscr{V}}_{\texttt{ES},-}}\frac{n}{\log(n)^2} \right\rceil \quad \text{and} \quad \hat{J}_{\texttt{ES-}\vartheta,+,n} = \left\lceil \frac{\hat{\mathcal{V}}_+}{\hat{\mathscr{V}}_{\texttt{ES},+}}\frac{n}{\log(n)^2} \right\rceil, \qquad \text{(SA-7)}$$

using the estimators in (SA-1)–(SA-4), and where $\hat{\mathcal{V}}_-$ and $\hat{\mathcal{V}}_+$ are consistent estimators of their population counterparts $\mathcal{V}_-$ and $\mathcal{V}_+$. The following theorem shows that, when the polynomial fits are viewed as nonparametric approximations with $k = k_n \to \infty$, the different number of bins selectors are nonparametric consistent.

Theorem SA1. *Suppose Assumption 1 holds with $S \geq 5$, $w : [x_l, x_u] \mapsto \mathbb{R}_+$ is continuous, and $Y_i(0)$ and $Y_i(1)$ are continuously distributed. If $k_n^7/n \to 0$ and $k_n \to \infty$, then*

$$\frac{\hat{J}_{\texttt{ES-}\omega,-,n}}{J_{\texttt{ES-}\omega,-,n}} \to_{\mathbb{P}} 1, \quad \frac{\hat{J}_{\texttt{ES-}\vartheta,-,n}}{J_{\texttt{ES-}\vartheta,-,n}} \to_{\mathbb{P}} 1, \quad \frac{\hat{J}_{\texttt{ES-}\omega,+,n}}{J_{\texttt{ES-}\omega,+,n}} \to_{\mathbb{P}} 1, \quad \frac{\hat{J}_{\texttt{ES-}\vartheta,+,n}}{J_{\texttt{ES-}\vartheta,+,n}} \to_{\mathbb{P}} 1,$$

provided that $\hat{\mathcal{V}}_- \to_{\mathbb{P}} \mathcal{V}_-$ and $\hat{\mathcal{V}}_+ \to_{\mathbb{P}} \mathcal{V}_+$.

Proof of Theorem SA1. Using Lemma A3 with $k = 2$ and $N_+/n \to_{\mathbb{P}} P_+$,

$$\begin{aligned}
\hat{\mathscr{V}}_{\texttt{ES},+} &= \frac{1}{x_u - \bar{x}} \frac{n}{4} \sum_{i=2}^{N_+} (X_{+,(i)} - X_{+,(i-1)})^2 (Y_{+,[i]} - Y_{+,[i-1]})^2 w(\bar{X}_{+,(i)}) \\
&= \frac{1}{x_u - \bar{x}} \frac{N_+}{4P_+} \sum_{i=2}^{N_+} (X_{+,(i)} - X_{+,(i-1)})^2 (Y_{+,[i]} - Y_{+,[i-1]})^2 w(\bar{X}_{+,(i)}) + o_{\mathbb{P}}(1) \\
&= \frac{1}{x_u - \bar{x}} \int_{\bar{x}}^{x_u} \frac{\sigma_+^2(x)}{f_+(x)} w(x) dx + o_{\mathbb{P}}(1),
\end{aligned}$$

which gives $\hat{\mathscr{V}}_{\texttt{ES},+} \to_{\mathbb{P}} \mathscr{V}_{\texttt{ES},+}$. Similarly, $\hat{\mathscr{V}}_{\texttt{ES},-} \to_{\mathbb{P}} \mathscr{V}_{\texttt{ES},-}$.

Next, recall that for power series estimators $\sup_{x \in [\bar{x}, x_u]} |\hat{\mu}^{(1)}_{+,k_n}(x) - \mu^{(1)}_+(x)|^2 = O_{\mathbb{P}}(k_n^7/n + k_n^{-2S+8}) = o_{\mathbb{P}}(1)$. Using this uniform consistency result, and Lemma A3 with $k = 1$, we have

$$\begin{aligned}
\hat{\mathscr{B}}_{\texttt{ES},+} &= \frac{(x_u - \bar{x})^2}{12} \sum_{i=2}^{N_+} (X_{+,(i)} - X_{+,(i-1)}) \left(\hat{\mu}^{(1)}_{+,k_n}(\bar{X}_{+,(i)})\right)^2 w(\bar{X}_{+,(i)}) \\
&= \frac{(x_u - \bar{x})^2}{12} \sum_{i=2}^{N_+} (X_{+,(i)} - X_{+,(i-1)}) \left(\mu^{(1)}_{+}(\bar{X}_{+,(i)})\right)^2 w(\bar{X}_{+,(i)}) + o_{\mathbb{P}}(1) \\
&= \frac{(x_u - \bar{x})^2}{12} \int_{\bar{x}}^{x_u} \left(\mu^{(1)}_{+}(x)\right)^2 w(x) dx + o_{\mathbb{P}}(1),
\end{aligned}$$

which gives $\hat{\mathscr{B}}_{\texttt{ES},+} \to_{\mathbb{P}} \mathscr{B}_{\texttt{ES},+}$. Similarly, $\hat{\mathscr{B}}_{\texttt{ES},-} \to_{\mathbb{P}} \mathscr{B}_{\texttt{ES},-}$. ■

Recall that the special case $\omega_{\mathscr{V},-} = \omega_{\mathscr{V},+} = 1/2$ gives $\hat{J}_{\texttt{ES-}\mu,-,n} = \hat{J}_{\texttt{ES-}\omega,-,n}$ and $\hat{J}_{\texttt{ES-}\mu,+,n} = \hat{J}_{\texttt{ES-}\omega,+,n}$. Theorem SA1 therefore gives a formal justification for employing any of the selectors introduced in our paper for the number of bins in ES RD Plots constructed with a known, arbitrary weight function $w(x)$; a particular choice being $w(x) = 1$.

As discussed in the main text, when $Y_i(0)$ and $Y_i(1)$ are not continuously distributed, the concomitant-based estimation method becomes invalid. In this case, we need to employ other more standard nonparametric techniques. For example, assuming that $\mathbb{E}[Y_i(t)^2|X_i = x]$, $t = 0, 1$, are twice continuously differentiable, we can use the following estimators:

$$\check{\mathscr{V}}_{\mathtt{ES},-} = \frac{1}{\bar{x} - x_l} \frac{n}{2} \sum_{i=2}^{N_-} (X_{-,(i)} - X_{-,(i-1)})^2 \hat{\sigma}^2_{-,k}(\bar{X}_{-,(i)}) w(\bar{X}_{-,(i)}),$$

$$\check{\mathscr{V}}_{\mathtt{ES},+} = \frac{1}{x_u - \bar{x}} \frac{n}{2} \sum_{i=2}^{N_+} (X_{+,(i)} - X_{+,(i-1)})^2 \hat{\sigma}^2_{+,k}(\bar{X}_{+,(i)}) w(\bar{X}_{+,(i)}),$$

$$\hat{\sigma}^2_{-,k}(x) = \hat{\mu}_{-,k,2}(x) - (\hat{\mu}_{-,k,1}(x))^2, \qquad \hat{\sigma}^2_{+,k}(x) = \hat{\mu}_{+,k,2}(x) - (\hat{\mu}_{+,k,1}(x))^2,$$

where, for $k \in \mathbb{Z}_+$ and $p \in \mathbb{Z}_{++}$,

$$\hat{\mu}_{-,k,p}(x) = \mathbf{r}_k(x)'\hat{\boldsymbol{\beta}}_{-,k,p}, \qquad \hat{\boldsymbol{\beta}}_{-,k,p} = \arg\min_{\boldsymbol{\beta}\in\mathbb{R}^{k+1}} \sum_{i=1}^{n} \mathbb{1}(X_i < \bar{x})(Y_i^p - \mathbf{r}_k(X_i)'\boldsymbol{\beta})^2,$$

$$\hat{\mu}_{+,k,p}(x) = \mathbf{r}_k(x)'\hat{\boldsymbol{\beta}}_{+,k,p}, \qquad \hat{\boldsymbol{\beta}}_{+,k,p} = \arg\min_{\boldsymbol{\beta}\in\mathbb{R}^{k+1}} \sum_{i=1}^{n} \mathbb{1}(X_i \geq \bar{x})(Y_i^p - \mathbf{r}_k(X_i)'\boldsymbol{\beta})^2,$$

and note that $\hat{\mu}_{-,k}(x) = \hat{\mu}_{-,k,1}(x)$ and $\hat{\mu}_{+,k}(x) = \hat{\mu}_{+,k,1}(x)$ with our notation.

From results for power series estimators,

$$\sup_{x\in[\bar{x},x_u]} |\hat{\mu}_{+,k_n,p}(x) - \mathbb{E}[Y(1)^p|X_i = x]|^2 = O_{\mathbb{P}}(k_n^3/n + k_n^{-2S+2}) = o_{\mathbb{P}}(1)$$

for $p = 1, 2$, under the assumptions imposed, which implies

$$\sup_{x\in[\bar{x},x_u]} |\hat{\sigma}^2_+(x) - \sigma^2_+|^2 = O_{\mathbb{P}}(k_n^3/n + k_n^{-2S+2}) = o_{\mathbb{P}}(1).$$

Therefore, Lemma A3 with $k = 2$ and $N_+/n \to_{\mathbb{P}} P_+$,

$$\begin{aligned}\check{\mathscr{V}}_{\mathtt{ES},+} &= \frac{1}{x_u - \bar{x}} \frac{n}{2} \sum_{i=2}^{N_+} (X_{+,(i)} - X_{+,(i-1)})^2 \hat{\sigma}_+^2(\bar{X}_{+,(i)}) w(\bar{X}_{+,(i)}) \\ &= \frac{1}{x_u - \bar{x}} \frac{N_+}{2P_+} \sum_{i=2}^{N_+} (X_{+,(i)} - X_{+,(i-1)})^2 \sigma_+^2(\bar{X}_{+,(i)}) w(\bar{X}_{+,(i)}) + o_{\mathbb{P}}(1) \\ &= \frac{1}{x_u - \bar{x}} \int_{\bar{x}}^{x_u} \frac{\sigma_+^2(x)}{f_+(x)} w(x) dx + o_{\mathbb{P}}(1),\end{aligned}$$

which gives $\check{\mathscr{V}}_{\mathtt{ES},+} \to_{\mathbb{P}} \mathscr{V}_{\mathtt{ES},+}$. Similarly, $\check{\mathscr{V}}_{\mathtt{ES},-} \to_{\mathbb{P}} \mathscr{V}_{\mathtt{ES},-}$.

Combining these results with Theorem SA1, it can easily be shown that the following selectors are consistent for any continuous, arbitrary choice of $w(x)$:

$$\check{J}_{\mathtt{ES}\text{-}\mu,-,n} = \left\lceil \left(\frac{2\hat{\mathscr{B}}_{\mathtt{ES},-}}{\check{\mathscr{V}}_{\mathtt{ES},-}} \right)^{1/3} n^{1/3} \right\rceil \quad \text{and} \quad \check{J}_{\mathtt{ES}\text{-}\mu,+,n} = \left\lceil \left(\frac{2\hat{\mathscr{B}}_{\mathtt{ES},+}}{\check{\mathscr{V}}_{\mathtt{ES},+}} \right)^{1/3} n^{1/3} \right\rceil, \tag{SA-8}$$

$$\check{J}_{\mathtt{ES}\text{-}\omega,-,n} = \left\lceil \omega_- \left(\frac{2\hat{\mathscr{B}}_{\mathtt{ES},-}}{\check{\mathscr{V}}_{\mathtt{ES},-}} \right)^{1/3} n^{1/3} \right\rceil \quad \text{and} \quad \check{J}_{\mathtt{ES}\text{-}\omega,+,n} = \left\lceil \omega_+ \left(\frac{2\hat{\mathscr{B}}_{\mathtt{ES},+}}{\check{\mathscr{V}}_{\mathtt{ES},+}} \right)^{1/3} n^{1/3} \right\rceil, \tag{SA-9}$$

$$\check{J}_{\mathtt{ES}\text{-}\vartheta,-,n} = \left\lceil \frac{\hat{\mathcal{V}}_-}{\check{\mathscr{V}}_{\mathtt{ES},-}} \frac{n}{\log(n)^2} \right\rceil \quad \text{and} \quad \check{J}_{\mathtt{ES}\text{-}\vartheta,+,n} = \left\lceil \frac{\hat{\mathcal{V}}_+}{\check{\mathscr{V}}_{\mathtt{ES},+}} \frac{n}{\log(n)^2} \right\rceil, \tag{SA-10}$$

provided that $\hat{\mathcal{V}}_- \to_{\mathbb{P}} \mathcal{V}_-$ and $\hat{\mathcal{V}}_+ \to_{\mathbb{P}} \mathcal{V}_+$.

3.2 Quantile Spaced RD Plots

We discuss generic estimators for QS RD Plots employing an arbitrary, known weighting function $w(x)$, paralleling the results given above for ES RD Plots. The underlying estimators are:

$$\hat{\mathscr{V}}_{\mathtt{QS},-} = \frac{n}{2N_-} \sum_{i=2}^{N_-} (X_{-,(i)} - X_{-,(i-1)})(Y_{-,[i]} - Y_{-,[i-1]})^2 w(\bar{X}_{-,(i)}), \tag{SA-11}$$

$$\hat{\mathscr{B}}_{\mathtt{QS},-} = \frac{N_-^2}{72} \sum_{i=2}^{N_-} (X_{-,(i)} - X_{-,(i-1)})^3 \left(\hat{\mu}_{-,k}^{(1)}(\bar{X}_{-,(i)}) \right)^2 w(\bar{X}_{-,(i)}), \tag{SA-12}$$

and

$$\hat{\mathscr{V}}_{\mathtt{QS},+} = \frac{n}{2N_+}\sum_{i=2}^{N_+}(X_{+,(i)} - X_{+,(i-1)})(Y_{+,[i]} - Y_{+,[i-1]})^2 w(\bar{X}_{+,(i)}), \tag{SA-13}$$

$$\hat{\mathscr{B}}_{\mathtt{QS},+} = \frac{N_+^2}{72}\sum_{i=2}^{N_+}(X_{+,(i)} - X_{+,(i-1)})^3 \left(\hat{\mu}^{(1)}_{+,k}(\bar{X}_{+,(i)})\right)^2 w(\bar{X}_{+,(i)}). \tag{SA-14}$$

Therefore, the resulting selectors for QS partitions take the form:

$$\hat{J}_{\mathtt{QS}\text{-}\mu,-,n} = \left\lceil \left(\frac{2\hat{\mathscr{B}}_{\mathtt{QS},-}}{\hat{\mathscr{V}}_{\mathtt{QS},-}}\right)^{1/3} n^{1/3}\right\rceil \quad \text{and} \quad \hat{J}_{\mathtt{QS}\text{-}\mu,+,n} = \left\lceil \left(\frac{2\hat{\mathscr{B}}_{\mathtt{QS},+}}{\hat{\mathscr{V}}_{\mathtt{QS},+}}\right)^{1/3} n^{1/3}\right\rceil, \tag{SA-15}$$

$$\hat{J}_{\mathtt{QS}\text{-}\omega,-,n} = \left\lceil \omega_- \left(\frac{2\hat{\mathscr{B}}_{\mathtt{QS},-}}{\hat{\mathscr{V}}_{\mathtt{QS},-}}\right)^{1/3} n^{1/3}\right\rceil \quad \text{and} \quad \hat{J}_{\mathtt{QS}\text{-}\omega,+,n} = \left\lceil \omega_+ \left(\frac{2\hat{\mathscr{B}}_{\mathtt{QS},+}}{\hat{\mathscr{V}}_{\mathtt{QS},+}}\right)^{1/3} n^{1/3}\right\rceil, \tag{SA-16}$$

$$\hat{J}_{\mathtt{QS}\text{-}\vartheta,-,n} = \left\lceil \frac{\hat{\mathcal{V}}_-}{\hat{\mathscr{V}}_{\mathtt{QS},-}} \frac{n}{\log(n)^2}\right\rceil \quad \text{and} \quad \hat{J}_{\mathtt{QS}\text{-}\vartheta,+,n} = \left\lceil \frac{\hat{\mathcal{V}}_+}{\hat{\mathscr{V}}_{\mathtt{QS},+}} \frac{n}{\log(n)^2}\right\rceil, \tag{SA-17}$$

using the estimators in (SA-11)–(SA-14), and appropriate consistent estimators $\hat{\mathcal{V}}_-$ and $\hat{\mathcal{V}}_+$. As in the case of Theorem SA1 for ES RD plots, the following theorem shows that these automatic partition-size selectors are nonparametric consistent if the polynomial fits are viewed as nonparametric approximations with $k = k_n \to \infty$.

Theorem SA2. *Suppose Assumption 1 holds with $S \geq 5$, $w : [x_l, x_u] \mapsto \mathbb{R}_+$ is continuous, and $Y_i(0)$ and $Y_i(1)$ are continuously distributed. If $k_n^7/n \to 0$ and $k_n \to \infty$, then*

$$\frac{\hat{J}_{\mathtt{QS}\text{-}\omega,-,n}}{J_{\mathtt{QS}\text{-}\omega,-,n}} \to_{\mathbb{P}} 1, \quad \frac{\hat{J}_{\mathtt{QS}\text{-}\vartheta,-,n}}{J_{\mathtt{QS}\text{-}\vartheta,-,n}} \to_{\mathbb{P}} 1, \quad \frac{\hat{J}_{\mathtt{QS}\text{-}\omega,+,n}}{J_{\mathtt{QS}\text{-}\omega,+,n}} \to_{\mathbb{P}} 1, \quad \frac{\hat{J}_{\mathtt{QS}\text{-}\vartheta,+,n}}{J_{\mathtt{QS}\text{-}\vartheta,+,n}} \to_{\mathbb{P}} 1,$$

provided that $\hat{\mathcal{V}}_- \to_{\mathbb{P}} \mathcal{V}_-$ and $\hat{\mathcal{V}}_+ \to_{\mathbb{P}} \mathcal{V}_+$.

In practice, the choice $w(x) = 1$ is arguably the simplest one, but our results permit any continuous function $w(x)$. The proof of Theorem SA2 is very similar to the proof of Theorem SA1 given above, and hence omitted here to conserve space.

Next, for the case of non-continuous potential outcomes $Y_i(0)$ and $Y_i(1)$, we use the series polynomial estimation approach already introduced. Assuming that $\mathbb{E}[Y_i(t)^2|X_i = x]$, $t = 0, 1$, are twice continuously differentiable, we may use the following estimators:

$$\check{\mathscr{V}}_{\texttt{QS},-} = \frac{n}{N_-} \sum_{i=2}^{N_-} (X_{-,(i)} - X_{-,(i-1)}) \hat{\sigma}^2_{-,k}(\bar{X}_{-,(i)}) w(\bar{X}_{-,(i)}),$$

$$\check{\mathscr{V}}_{\texttt{QS},+} = \frac{n}{N_+} \sum_{i=2}^{N_+} (X_{+,(i)} - X_{+,(i-1)}) \hat{\sigma}^2_{+,k}(\bar{X}_{+,(i)}) w(\bar{X}_{+,(i)}),$$

where $\hat{\sigma}^2_{-,k}(x)$ and $\hat{\sigma}^2_{+,k}(x)$ are the polynomial approximations already discussed. The associated data-driven partition-size selectors are

$$\check{J}_{\texttt{QS-}\mu,-,n} = \left\lceil \left(\frac{2\hat{\mathscr{B}}_{\texttt{QS},-}}{\check{\mathscr{V}}_{\texttt{QS},-}} \right)^{1/3} n^{1/3} \right\rceil \quad \text{and} \quad \check{J}_{\texttt{QS-}\mu,+,n} = \left\lceil \left(\frac{2\hat{\mathscr{B}}_{\texttt{QS},+}}{\check{\mathscr{V}}_{\texttt{QS},+}} \right)^{1/3} n^{1/3} \right\rceil, \tag{SA-18}$$

$$\check{J}_{\texttt{QS-}\omega,-,n} = \left\lceil \omega_- \left(\frac{2\hat{\mathscr{B}}_{\texttt{QS},-}}{\check{\mathscr{V}}_{\texttt{QS},-}} \right)^{1/3} n^{1/3} \right\rceil \quad \text{and} \quad \check{J}_{\texttt{QS-}\omega,+,n} = \left\lceil \omega_+ \left(\frac{2\hat{\mathscr{B}}_{\texttt{QS},+}}{\check{\mathscr{V}}_{\texttt{QS},+}} \right)^{1/3} n^{1/3} \right\rceil, \tag{SA-19}$$

$$\check{J}_{\texttt{QS-}\vartheta,-,n} = \left\lceil \frac{\hat{\mathcal{V}}_-}{\check{\mathscr{V}}_{\texttt{QS},-}} \frac{n}{\log(n)^2} \right\rceil \quad \text{and} \quad \check{J}_{\texttt{QS-}\vartheta,+,n} = \left\lceil \frac{\hat{\mathcal{V}}_+}{\check{\mathscr{V}}_{\texttt{QS},+}} \frac{n}{\log(n)^2} \right\rceil, \tag{SA-20}$$

which are easily shown to be consistent in the sense of Theorem SA2, provided the conditions in that theorem hold.

4 Other Empirical Applications

In this section we include three additional empirical applications to illustrate the performance of our proposed methods when applied to different real datasets. Software packages in `R` and `STATA` are described in Calonico et al. (2015, 2014a).

4.1 U.S. Senate Data

We employ an extract of the dataset constructed by Cattaneo et al. (2015), who study several measures of incumbency advantage in U.S. Senate elections for the period 1914–2010. In particular, we focus here on the RD effect of the Democratic party winning a U.S. Senate seat on the vote share obtained in the following election for that same seat. This empirical illustration is analogous to the one presented by Lee (2008) for U.S. House elections: the running variable is the state-level margin of victory of the Democratic party in an election for a Senate seat, the threshold is $\bar{x} = 0$ and the outcome is the vote share of the Democratic party in the following election for the same Senate seat in the state, which occurs six years later. The unit of observation is the state, and the data set has a total of $n = 1,297$ state-year complete observations.

Results are presented in Figures SA-1 and SA-2.

4.2 Progresa/Oportunidades Data

We illustrate the performance of our methods employing household data from Oportunidades (formerly known as Progresa), a well-known large-scale anti-poverty conditional cash transfer program in Mexico. This conditional cash transfer program targeted poor households in rural and urban areas in Mexico. The program started in 1998 under the name of Progresa in rural areas. The most important elements of the program are the nutrition, health and education components. The nutrition component consists of a cash grant for all treated households and an additional supplement for households with young children and pregnant or lactating mothers. The educational grant is linked to regular attendance in school and starts on the third grade of primary school and continues until the last grade of secondary school. The transfer constituted a significant contribution to the income of eligible families.

This social program is best known for its experimental design: treatment was initially randomly assigned at the locality level in rural areas. Progresa was expanded to urban areas urban in 2003. Unlike the rural program, the allocation across treatment and control areas was not random. Instead, it was first offered in blocks with the highest density of poor households. In order to accurately target the program to poor households, in both rural and urban areas Mexican officials constructed a pre-intervention (at baseline) household poverty-index that determined each household's eligibility. Thus, Progresa/Oportunidades' eligibility assignment rule naturally leads to sharp (intention-to-treat) regression-discontinuity designs. For additional details for data construction, empirical analysis and related literature, see Calonico et al. (2014b, Section S.4).

Our empirical exercise investigates the program treatment effect on household non-food consumption expenditures two years after its implementation. In this application, X_i denotes the household's poverty-index, $\bar{x} = 0$ denotes the centered cutoff for each RD design, and Y_i denotes per capita non-food consumption. Our final database contains 691 control households ($X_i < 0$) and $2,118$ intention-to-treat households ($X_i \geq 0$) in the urban RD design ($n = 2,809$, $X_i \in [-2.25\ ,\ 4.11]$).

Results are presented in Figures SA-3 and SA-4.

4.3 Head Start Data

Head Start is a program of the United States Department of Health and Human Services that provides early childhood education, health, nutrition, and parent involvement services to low-income children and their families. It was established in 1965 as part of the War on Poverty, in order to foster stable family relationships, enhance children's physical and emotional well-being, and establish an environment to develop cognitive skills.

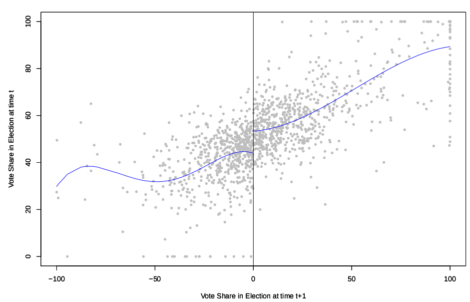
For each county, eligibility is based on the county's poverty rate, inducing a natural RD design. Ludwig and Miller (2007) uses this to identify the program's effects on health and schooling. For each county $i = 1, 2, ..., n$, the forcing variable is the county's 1960 poverty rate with treatment assignment given by $T_i = \mathbf{1}(X_i \geq \bar{x})$, where X_i represents the county's poverty rate in 1960 and $\bar{x}$ is the fixed threshold level. The cutoff is set to the poverty rate value of the 300th poorest county in 1960, which in this dataset is given by $\bar{x} = 59.198$. Here we consider as outcome variable the mortality rates per $100,000$ for children between 5–9 years old, with Head Start-related causes, for $1973 - 1983$ (see Panel A, Figure IV in Ludwig and Miller (2007)).

Results are presented in Figures SA-5 and SA-6.

4.4 Summary of Results

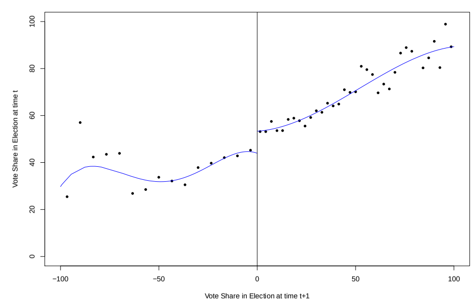
In all the empirical applications we considered, the data-driven selectors introduced in the main paper seemed to perform very well. The mimicking variance selector for the number of bins consistently delivered a disciplined "cloud of points", which appears to be substantially more useful than the scatter plot of the raw data. In addition, the IMSE-optimal choice of number of bins also performed well, in all cases "tracing out" the estimated smooth polynomial regression fits. As for the implementations, spacings estimators perform on par with polynomial estimators in all the applications considered. Finally, it is worth noting that ES and QS RD plots do not necessarily deliver different number of bins. For example, in the Head Start data set, the mimicking variance choices are essentially identical for both types of RD plots.

Figure SA-1: Scatter Plot and Automatic Data-driven ES RD Plots for Senate Elections Data



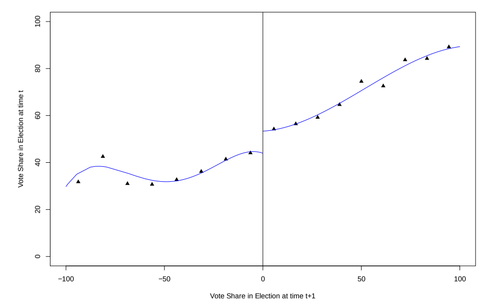
(a) Scatter Plot of Raw Data

$N_{-} = 595$; $N_{+} = 702$



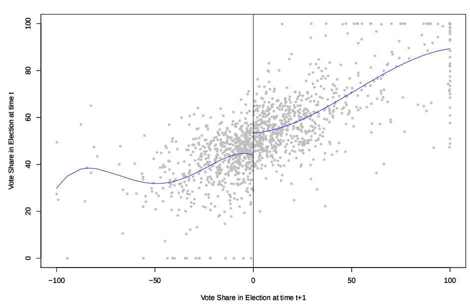
(b) Mimicking Variance, Spacings

$\hat{J}_{\texttt{ES-}\vartheta,-,n} = 15$; $\hat{J}_{\texttt{ES-}\vartheta,+,n} = 35$



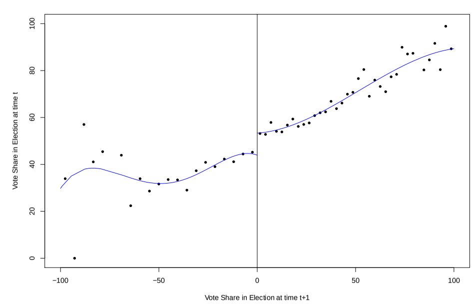
(c) IMSE-optimal, Spacings

$\hat{J}_{\texttt{ES-}\mu,-,n} = 8$; $\hat{J}_{\texttt{ES-}\mu,+,n} = 9$



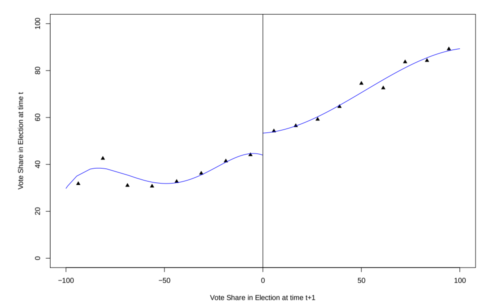
(d) Scatter Plot of Raw Data

$N_{-} = 595$; $N_{+} = 702$



(e) Mimicking Variance, Series

$\check{J}_{\texttt{ES-}\vartheta,-,n} = 21$; $\check{J}_{\texttt{ES-}\vartheta,+,n} = 36$

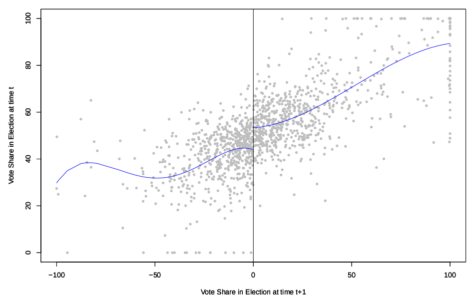


(f) IMSE-optimal, Series

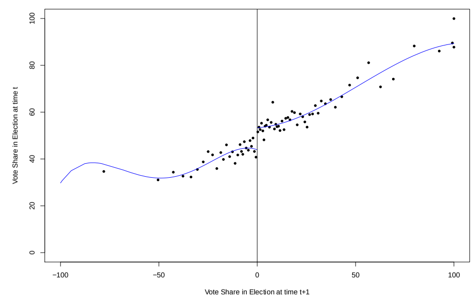
$\check{J}_{\texttt{ES-}\mu,-,n} = 8$; $\check{J}_{\texttt{ES-}\mu,+,n} = 9$

Notes: (i) sample size is $n = 1,297$; (ii) N_{-} and N_{+} denote the sample sizes for control and treatment units, respectively; (iii) solid blue lines depict 4th order polynomial fits using control and treated units separately.

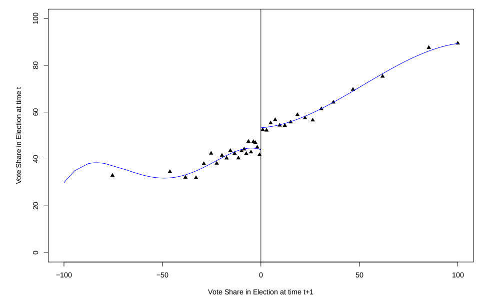
Figure SA-2: Scatter Plot and Automatic Data-driven QS RD Plots for Senate Elections Data



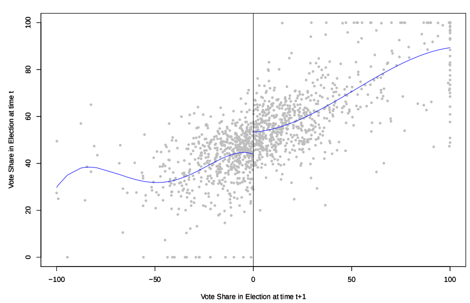
(a) Scatter Plot of Raw Data
$N_{-} = 595$; $N_{+} = 702$



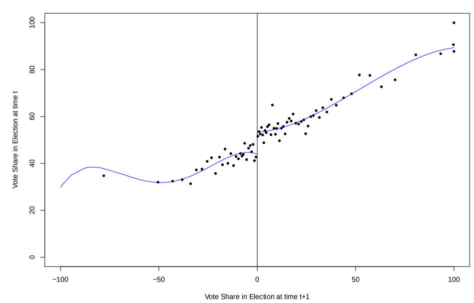
(b) Mimicking Variance, Spacings
$\hat{J}_{\texttt{QS-}\vartheta,-,n} = 28$; $\hat{J}_{\texttt{QS-}\vartheta,+,n} = 49$



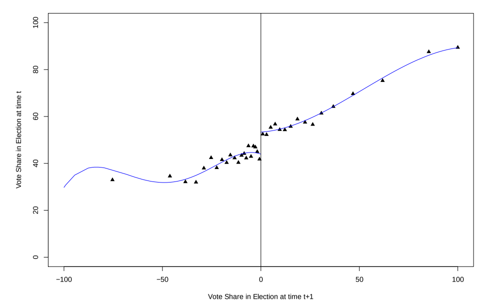
(c) IMSE-optimal, Spacings
$\hat{J}_{\texttt{QS-}\mu,-,n} = 21$; $\hat{J}_{\texttt{QS-}\mu,+,n} = 16$



(d) Scatter Plot of Raw Data
$N_{-} = 595$; $N_{+} = 702$



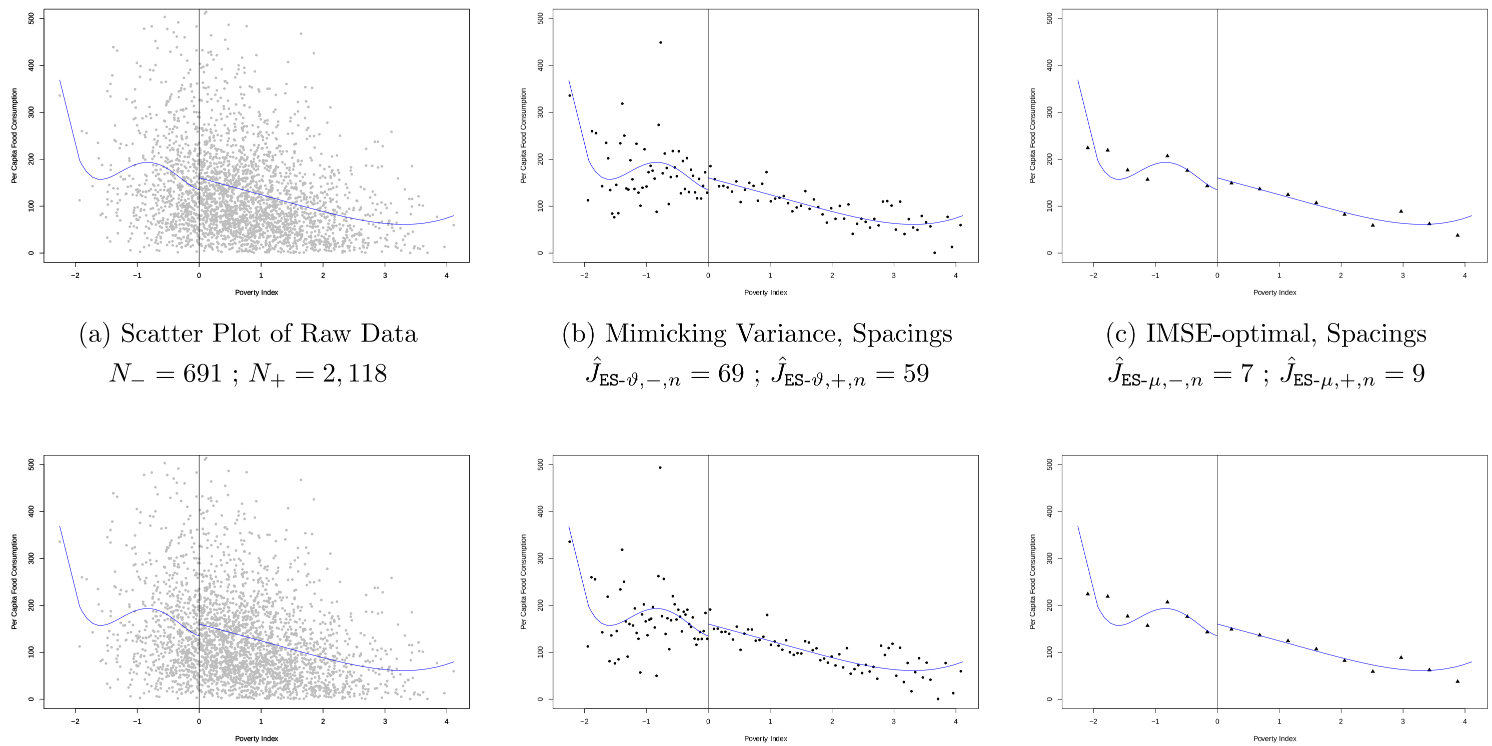
(e) Mimicking Variance, Series
$\check{J}_{\texttt{QS-}\vartheta,-,n} = 29$; $\check{J}_{\texttt{QS-}\vartheta,+,n} = 50$



(f) IMSE-optimal, Series
$\check{J}_{\texttt{QS-}\mu,-,n} = 21$; $\check{J}_{\texttt{QS-}\mu,+,n} = 16$

Notes: (i) sample size is $n = 1,297$; (ii) N_{-} and N_{+} denote the sample sizes for control and treatment units, respectively; (iii) solid blue lines depict 4th order polynomial fits using control and treated units separately.

Figure SA-3: Scatter Plot and Automatic Data-driven ES RD Plots for Progresa/Oportunidades (Urban Localities).



(a) Scatter Plot of Raw Data
$N_- = 691$; $N_+ = 2,118$

(b) Mimicking Variance, Spacings
$\hat{J}_{\texttt{ES-}\vartheta,-,n} = 69$; $\hat{J}_{\texttt{ES-}\vartheta,+,n} = 59$

(c) IMSE-optimal, Spacings
$\hat{J}_{\texttt{ES-}\mu,-,n} = 7$; $\hat{J}_{\texttt{ES-}\mu,+,n} = 9$

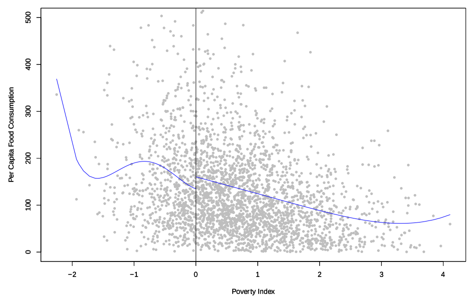
(d) Scatter Plot of Raw Data
$N_- = 691$; $N_+ = 2,118$

(e) Mimicking Variance, Series
$\check{J}_{\texttt{ES-}\vartheta,-,n} = 77$; $\check{J}_{\texttt{ES-}\vartheta,+,n} = 67$

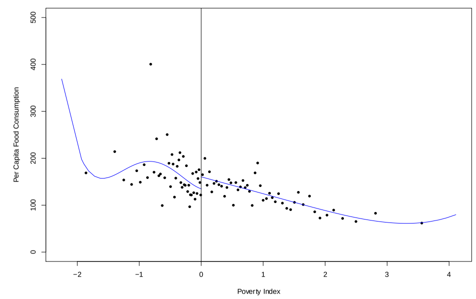
(f) IMSE-optimal, Series
$\check{J}_{\texttt{ES-}\mu,-,n} = 7$; $\check{J}_{\texttt{ES-}\mu,+,n} = 9$

Notes: (i) sample size is $n = 2,809$; (ii) N_- and N_+ denote the sample sizes for control and treatment units, respectively; (iii) solid blue lines depict 4th order polynomial fits using control and treated units separately.

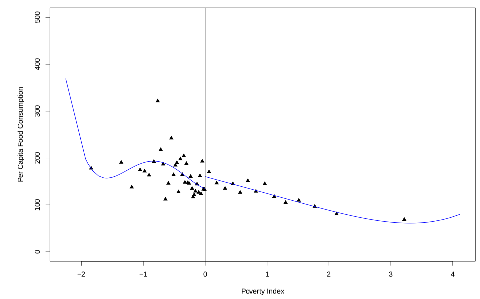
Figure SA-4: Scatter Plot and Automatic Data-driven QS RD Plots for Progresa/Oportunidades (Urban Localities).



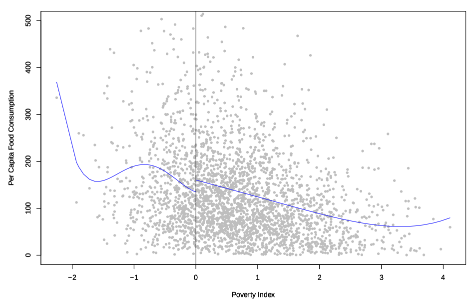
(a) Scatter Plot of Raw Data
$N_- = 691$; $N_+ = 2,118$



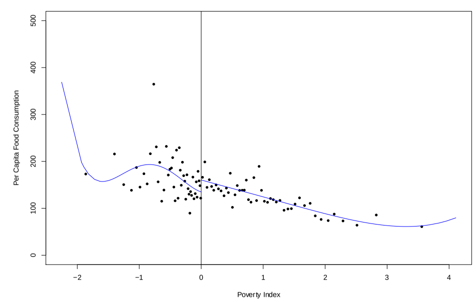
(b) Mimicking Variance, Spacings
$\hat{J}_{\texttt{QS-}\vartheta,-,n} = 45$; $\hat{J}_{\texttt{QS-}\vartheta,+,n} = 46$



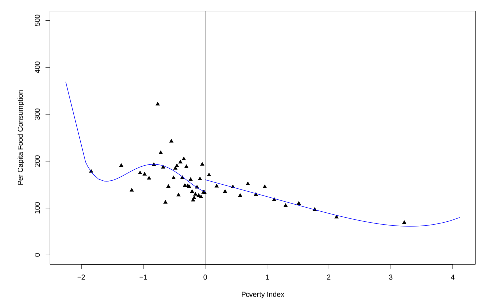
(c) IMSE-optimal, Spacings
$\hat{J}_{\texttt{QS-}\mu,-,n} = 36$; $\hat{J}_{\texttt{QS-}\mu,+,n} = 14$



(d) Scatter Plot of Raw Data
$N_- = 691$; $N_+ = 2,118$



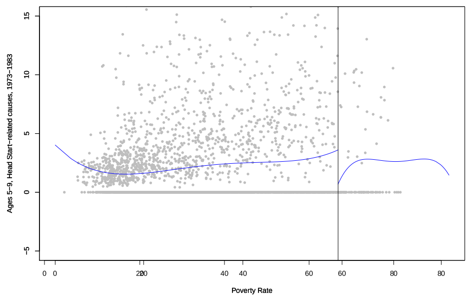
(e) Mimicking Variance, Series
$\check{J}_{\texttt{QS-}\vartheta,-,n} = 46$; $\check{J}_{\texttt{QS-}\vartheta,+,n} = 47$



(f) IMSE-optimal, Series
$\check{J}_{\texttt{QS-}\mu,-,n} = 36$; $\check{J}_{\texttt{QS-}\mu,+,n} = 14$

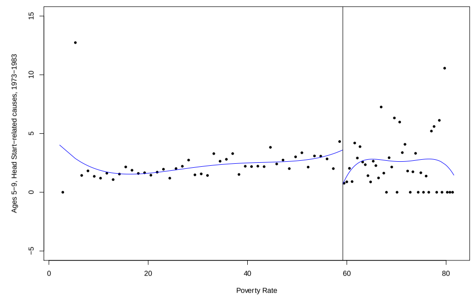
Notes: (i) sample size is $n = 2,809$; (ii) N_- and N_+ denote the sample sizes for control and treatment units, respectively; (iii) solid blue lines depict 4th order polynomial fits using control and treated units separately.

Figure SA-5: Scatter Plot and Automatic Data-driven ES RD Plots for Head Start Assistance



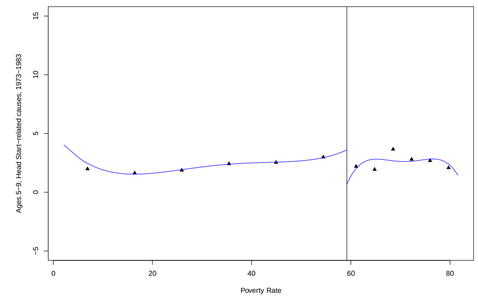
(a) Scatter Plot of Raw Data

$N_{-} = 2,810$; $N_{+} = 294$



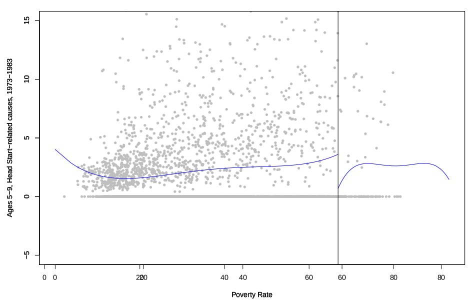
(b) Mimicking Variance, Spacings

$\hat{J}_{\texttt{ES-}\vartheta,-,n} = 45$; $\hat{J}_{\texttt{ES-}\vartheta,+,n} = 42$



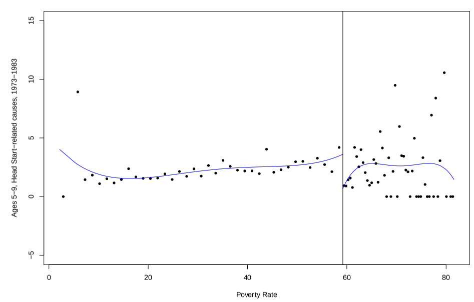
(c) IMSE-optimal, Spacings

$\hat{J}_{\texttt{ES-}\mu,-,n} = 6$; $\hat{J}_{\texttt{ES-}\mu,+,n} = 6$



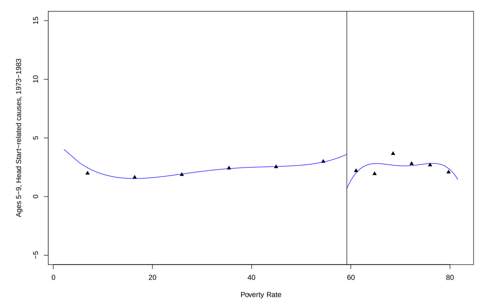
(d) Scatter Plot of Raw Data

$N_{-} = 2,810$; $N_{+} = 294$



(e) Mimicking Variance, Series

$\check{J}_{\texttt{ES-}\vartheta,-,n} = 39$; $\check{J}_{\texttt{ES-}\vartheta,+,n} = 52$

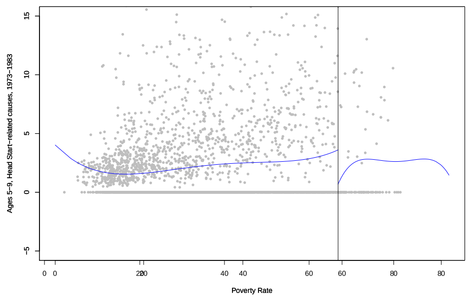


(f) IMSE-optimal, Series

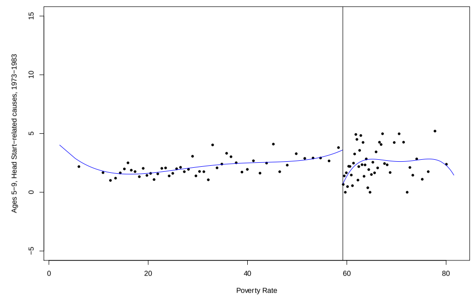
$\check{J}_{\texttt{ES-}\mu,-,n} = 6$; $\check{J}_{\texttt{ES-}\mu,+,n} = 6$

Notes: (i) sample size is $n = 3,104$; (ii) N_{-} and N_{+} denote the sample sizes for control and treatment units, respectively; (iii) solid blue lines depict 4th order polynomial fits using control and treated units separately.

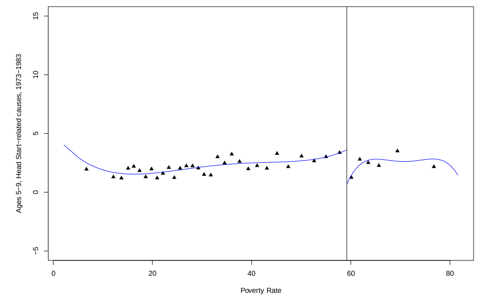
Figure SA-6: Scatter Plot and Automatic Data-driven QS RD Plots for Head Start Assistance



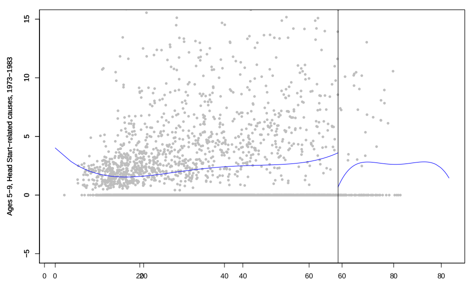
(a) Scatter Plot of Raw Data
$N_{-} = 2,810$; $N_{+} = 294$



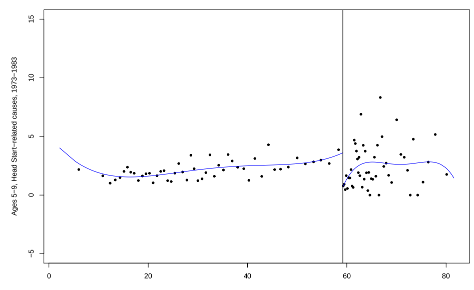
(b) Mimicking Variance, Spacings
$\hat{J}_{\texttt{QS}\text{-}\vartheta,-,n} = 48$; $\hat{J}_{\texttt{QS}\text{-}\vartheta,+,n} = 47$



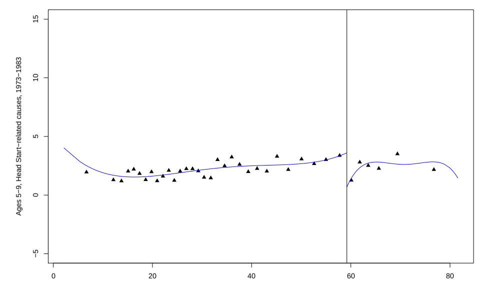
(c) IMSE-optimal, Spacings
$\hat{J}_{\texttt{QS}\text{-}\mu,-,n} = 31$; $\hat{J}_{\texttt{QS}\text{-}\mu,+,n} = 6$



(d) Scatter Plot of Raw Data
$N_{-} = 2,810$; $N_{+} = 294$



(e) Mimicking Variance, Series
$\check{J}_{\texttt{QS}\text{-}\vartheta,-,n} = 49$; $\check{J}_{\texttt{QS}\text{-}\vartheta,+,n} = 49$



(f) IMSE-optimal, Series
$\check{J}_{\texttt{QS}\text{-}\mu,-,n} = 31$; $\check{J}_{\texttt{QS}\text{-}\mu,+,n} = 6$

Notes: (i) sample size is $n = 3,104$; (ii) N_{-} and N_{+} denote the sample sizes for control and treatment units, respectively; (iii) solid blue lines depict 4th order polynomial fits using control and treated units separately.

5 Complete Simulation Results

We report the results from a Monte Carlo experiment to study the finite-sample behavior of our proposed methods. We consider several data generating processes, which vary in the distribution of the running variable, the conditional variance, and the distribution of the unobserved error term in the regression function.

Specifically, the data is generated as i.i.d. draws, $\{(Y_i, X_i)' : i = 1, 2, ..., n\}$ following

$$Y_i = \mu(X_i) + \varepsilon_i, \qquad X_i \sim \mathcal{F}_x, \qquad \varepsilon_i \sim \sigma(X_i)\mathcal{F}_\varepsilon,$$

where

$$\mu(x) = \begin{cases} 0.48 + 1.27x + 7.18x^2 + 20.21x^3 + 21.54x^4 + 7.33x^5 & \text{if } x < 0 \\ 0.52 + 0.84x - 3.00x^2 + 7.99x^3 - 9.01x^4 + 3.56x^5 & \text{if } x \geq 0 \end{cases},$$

and $\mathcal{F}_x$ equals either $(2\mathcal{B}(p_1, p_2) - 1)$, with $\mathcal{B}(p_1, p_2)$ denoting a Beta distribution with parameters p_1 and p_2, or equals a mixture of two normal distributions with means μ_1 and μ_2, respectively, same standard deviations set to 1/4 and mixing weights ω_1 and ω_2, respectively. In addition, $\sigma(x)$ is either equal to 1 (homoskedasticity) or equal to $\exp(-|x|/2)$ (heteroskedasticity), and $\mathcal{F}_\varepsilon$ is either $\mathcal{N}(0, 1)$ or $(\chi_4 - 4)/\sqrt{8}$. The functional form of $\mu(x)$ is obtained by fitting a 5-th order global polynomial with different coefficients for control and treatment units separately using the original data of Lee (2008), after discarding observations with past margin of victory greater than 99 and less than -99 percentage points. Figure SA-7 plots the regression function $\mu(x)$ and the different choices for the density of X_i. Notice that some of these densities take on "low" values in some regions of the support of X_i, in same cases near the RD cutoff.

Our Monte Carlo experiment considers 16 models that combine different choices of $\mathcal{F}_x$, $\sigma(x)$ and $\mathcal{F}_\varepsilon$, as described in Table SA-1. For each model in Table SA-1, we set $n = 5,000$ and generate $5,000$ simulations to compute the IMSE of both ES and QS partitioning schemes for different possible number of bins, as well as for the IMSE-optimal data-driven selector proposed. In each case considered, we also computed the mimicking variance selectors introduced in the paper, both infeasible and data-driven versions.

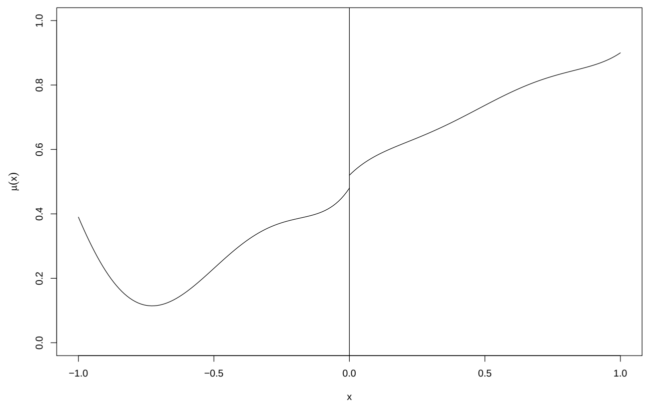
All tables include results for both ES and QS RD plots organized in two distinct panels. Panel

A focuses attention on the IMSE of different partitioning schemes in finite samples, as well as the performance of the associated IMSE-optimal data-driven selectors. All IMSEs are normalized relative to the IMSE evaluated at the optimal partition-size choice to avoid any scaling issue. Panel B reports several features of the empirical (finite-sample) distribution of the different data-driven number of bins selectors introduced in this paper: (i) spacings-based selectors for ES RD plots, (ii) polynomial-based selectors for ES RD plots, (iii) spacings-based selectors for QS RD plots, and (iv) polynomial-based selectors for QS RD plots. Therefore, our Monte Carlo experiment is designed to capture the finite-sample performance of Theorems 1 and 2 in terms of providing a good approximation to the IMSE (Panel A), and the finite-sample performance of Theorems 3 and 4 as well as the other consistency results discussed in the remarks in the paper (Panel B).

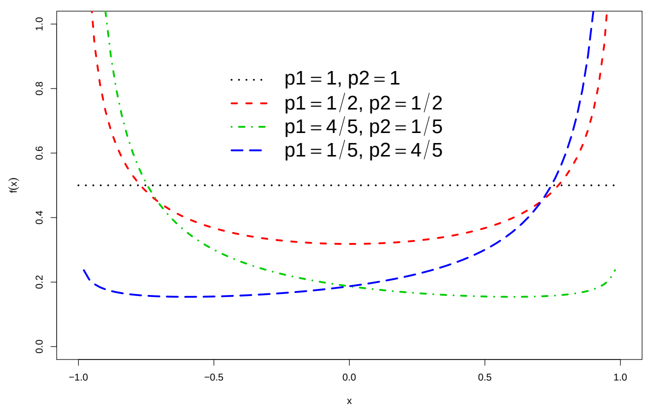
The results of our simulation experiment are very encouraging. First, in all cases the IMSE is minimized at the corresponding IMSE-optimal number of bins choice derived in the paper, suggesting that Theorems 1 and 2 provide a good finite-sample approximation. The theoretical IMSE-optimal number of bins almost always exactly coincides with the simulated IMSE-optimal number of bins. Second, in all models we find that our proposed data-driven implementations of the different number of bins selectors perform quite well, exhibiting a concentrated finite-sample distribution centered at the target population (optimal) choice introduced in this paper. That is, the summary statistics in Panel B of each table show that our data-driven implementations of the population selectors choices have a finite sample distribution well centered and concentrated around their population targets, when using either spacings estimators or polynomial estimators.

In sum, our extensive simulation study indicates that the different data-driven number of bins selectors underlying the construction of the RD plots perform well in finite samples.

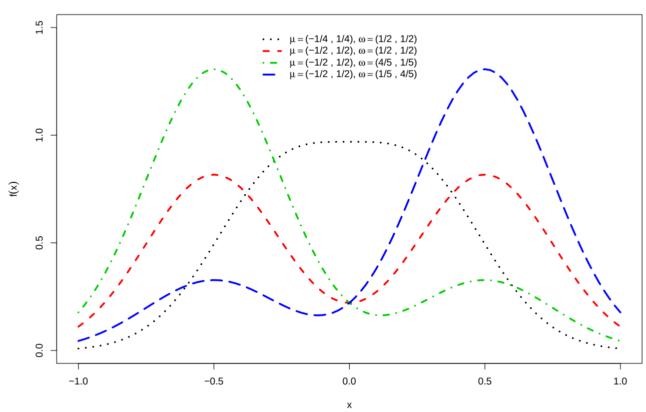
Figure SA-7: Data Generating Processes



(a) Regression function, $\mu(x)$.



(b) X_i's distribution, $\mathcal{B}(p_1, p_2)$.



(c) X_i's distribution, Mixture of Normals

Table SA-1: Data Generating Processes

Panel A: Models 1 to 8

Model	p_1	p_2	$\sigma^2(x)$	$\mathcal{F}_\varepsilon$
1	1	1	1	$\mathcal{N}(0,1)$
2	0.5	0.5	1	$\mathcal{N}(0,1)$
3	0.2	0.8	$\exp(-\lvert x\rvert/2)$	$\mathcal{N}(0,1)$
4	0.8	0.2	$\exp(-\lvert x\rvert/2)$	$\mathcal{N}(0,1)$
5	1	1	1	$(\chi_4-4)/\sqrt{8}$
6	0.5	0.5	1	$(\chi_4-4)/\sqrt{8}$
7	0.2	0.8	$\exp(-\lvert x\rvert/2)$	$(\chi_4-4)/\sqrt{8}$
8	0.8	0.2	$\exp(-\lvert x\rvert/2)$	$(\chi_4-4)/\sqrt{8}$

Panel B: Models 9 to 16

Model	μ_1	μ_2	ω_1	ω_2	$\sigma^2(x)$	$\mathcal{F}_\varepsilon$
9	-0.25	0.25	0.5	0.5	1	$\mathcal{N}(0,1)$
10	-0.5	0.5	0.5	0.5	1	$\mathcal{N}(0,1)$
11	-0.5	0.5	0.8	0.2	$\exp(-\lvert x\rvert/2)$	$\mathcal{N}(0,1)$
12	-0.5	0.5	0.2	0.8	$\exp(-\lvert x\rvert/2)$	$\mathcal{N}(0,1)$
13	-0.25	0.25	0.5	0.5	1	$(\chi_4-4)/\sqrt{8}$
14	-0.5	0.5	0.5	0.5	1	$(\chi_4-4)/\sqrt{8}$
15	-0.5	0.5	0.8	0.2	$\exp(-\lvert x\rvert/2)$	$(\chi_4-4)/\sqrt{8}$
16	-0.5	0.5	0.2	0.8	$\exp(-\lvert x\rvert/2)$	$(\chi_4-4)/\sqrt{8}$

Table SA-2: Simulations Results for Model 1

Panel A: IMSE for Grid of Number of Bins and Estimated Choices

$J_{-,n}$	$\frac{\mathsf{IMSE}_{\mathsf{ES},-}(J_{-,n})}{\mathsf{IMSE}^*_{\mathsf{ES},-}}$	$J_{+,n}$	$\frac{\mathsf{IMSE}_{\mathsf{ES},+}(J_{+,n})}{\mathsf{IMSE}^*_{\mathsf{ES},+}}$	$J_{-,n}$	$\frac{\mathsf{IMSE}_{\mathsf{QS},-}(J_{-,n})}{\mathsf{IMSE}^*_{\mathsf{QS},-}}$	$J_{+,n}$	$\frac{\mathsf{IMSE}_{\mathsf{QS},+}(J_{+,n})}{\mathsf{IMSE}^*_{\mathsf{QS},+}}$
20	1.047	11	1.148	20	1.047	11	1.148
21	1.027	12	1.081	21	1.027	12	1.081
22	1.013	13	1.039	22	1.013	13	1.039
23	1.005	14	1.014	23	1.005	14	1.014
24	1.000	15	1.002	24	1.000	15	1.002
25	1.000	16	1.000	25	1.000	16	1.000
26	1.003	17	1.006	26	1.003	17	1.006
27	1.008	18	1.017	27	1.008	18	1.017
28	1.016	19	1.033	28	1.016	19	1.033
29	1.025	20	1.053	29	1.025	20	1.053
30	1.036	21	1.076	30	1.036	21	1.076
$\hat{J}_{\mathsf{ES}\text{-}\mu,-,n}$	1.033	$\hat{J}_{\mathsf{ES}\text{-}\mu,+,n}$	0.9435	$\hat{J}_{\mathsf{QS}\text{-}\mu,-,n}$	1.072	$\hat{J}_{\mathsf{QS}\text{-}\mu,+,n}$	0.9351
$\check{J}_{\mathsf{ES}\text{-}\mu,-,n}$	1.034	$\check{J}_{\mathsf{ES}\text{-}\mu,+,n}$	0.9428	$\check{J}_{\mathsf{QS}\text{-}\mu,-,n}$	1.073	$\check{J}_{\mathsf{QS}\text{-}\mu,+,n}$	0.9347

Panel B: Summary Statistics for the Estimated Number of Bins

Pop. Par.		Min.	1st Qu.	Median	Mean	3rd Qu.	Max.	Std. Dev.
$J_{\mathsf{ES}\text{-}\mu,-,n} = 25$	$\hat{J}_{\mathsf{ES}\text{-}\mu,-,n}$	22	25	26	25.95	27	29	0.93
	$\check{J}_{\mathsf{ES}\text{-}\mu,-,n}$	23	25	26	25.93	26	29	0.87
$J_{\mathsf{ES}\text{-}\vartheta,-,n} = 118$	$\hat{J}_{\mathsf{ES}\text{-}\vartheta,-,n}$	105	116	120	119.6	123	139	5.05
	$\check{J}_{\mathsf{ES}\text{-}\vartheta,-,n}$	110	117	119	119.3	121	131	2.72
$J_{\mathsf{ES}\text{-}\mu,+,n} = 16$	$\hat{J}_{\mathsf{ES}\text{-}\mu,+,n}$	14	15	15	15.34	16	17	0.57
	$\check{J}_{\mathsf{ES}\text{-}\mu,+,n}$	14	15	15	15.34	16	17	0.55
$J_{\mathsf{ES}\text{-}\vartheta,+,n} = 116$	$\hat{J}_{\mathsf{ES}\text{-}\vartheta,+,n}$	103	113	117	116.7	120	139	4.71
	$\check{J}_{\mathsf{ES}\text{-}\vartheta,+,n}$	107	115	117	116.7	118	128	2.65
$J_{\mathsf{QS}\text{-}\mu,-,n} = 25$	$\hat{J}_{\mathsf{QS}\text{-}\mu,-,n}$	23	26	27	26.91	27	30	0.92
	$\check{J}_{\mathsf{QS}\text{-}\mu,-,n}$	23	26	27	26.89	27	30	0.90
$J_{\mathsf{QS}\text{-}\vartheta,-,n} = 118$	$\hat{J}_{\mathsf{QS}\text{-}\vartheta,-,n}$	108	117	120	119.6	122	134	3.66
	$\check{J}_{\mathsf{QS}\text{-}\vartheta,-,n}$	110	117	119	119.3	121	131	2.71
$J_{\mathsf{QS}\text{-}\mu,+,n} = 16$	$\hat{J}_{\mathsf{QS}\text{-}\mu,+,n}$	14	15	15	15.21	15	17	0.51
	$\check{J}_{\mathsf{QS}\text{-}\mu,+,n}$	14	15	15	15.21	15	17	0.50
$J_{\mathsf{QS}\text{-}\vartheta,+,n} = 116$	$\hat{J}_{\mathsf{QS}\text{-}\vartheta,+,n}$	106	114	117	116.6	119	130	3.50
	$\check{J}_{\mathsf{QS}\text{-}\vartheta,+,n}$	107	115	117	116.7	118	128	2.65

Notes:

(i) Population quantities:
$J_{\mathsf{ES}\text{-}\mu,\cdot,n}$ = IMSE-optimal partition size for ES RD Plot.
$J_{\mathsf{ES}\text{-}\vartheta,\cdot,n}$ = Mimicking variance partition size for ES RD Plot.
$J_{\mathsf{QS}\text{-}\mu,\cdot,n}$ = IMSE-optimal partition size for QS RD Plot.
$J_{\mathsf{QS}\text{-}\vartheta,\cdot,n}$ = Mimicking variance partition size for QS RD Plot.
$\mathsf{IMSE}^*_{\mathsf{ES},\cdot} = \mathsf{IMSE}_{\mathsf{ES},\cdot}(J_{\mathsf{ES}\text{-}\mu,\cdot,n})$ = ES IMSE function evaluated at optimal choice.
$\mathsf{IMSE}^*_{\mathsf{QS},\cdot} = \mathsf{IMSE}_{\mathsf{QS},\cdot}(J_{\mathsf{QS}\text{-}\mu,\cdot,n})$ = QS IMSE function evaluated at optimal choice.

(ii) Estimators:
$\hat{J}_{\mathsf{ES}\text{-}\mu,\cdot,n}$ = spacings estimator of $J_{\mathsf{ES}\text{-}\mu,\cdot,n}$; $\check{J}_{\mathsf{ES}\text{-}\mu,\cdot,n}$ = polynomial estimator of $J_{\mathsf{ES}\text{-}\mu,\cdot,n}$.
$\hat{J}_{\mathsf{ES}\text{-}\vartheta,\cdot,n}$ = spacings estimator of $J_{\mathsf{ES}\text{-}\vartheta,\cdot,n}$; $\check{J}_{\mathsf{ES}\text{-}\vartheta,\cdot,n}$ = polynomial estimator of $J_{\mathsf{ES}\text{-}\vartheta,\cdot,n}$.
$\hat{J}_{\mathsf{QS}\text{-}\mu,\cdot,n}$ = spacings estimator of $J_{\mathsf{QS}\text{-}\mu,\cdot,n}$; $\check{J}_{\mathsf{QS}\text{-}\mu,\cdot,n}$ = polynomial estimator of $J_{\mathsf{QS}\text{-}\mu,\cdot,n}$.
$\hat{J}_{\mathsf{QS}\text{-}\vartheta,\cdot,n}$ = spacings estimator of $J_{\mathsf{QS}\text{-}\vartheta,\cdot,n}$; $\check{J}_{\mathsf{QS}\text{-}\vartheta,\cdot,n}$ = polynomial estimator of $J_{\mathsf{QS}\text{-}\vartheta,\cdot,n}$.

Table SA-3: Simulations Results for Model 2

Panel A: IMSE for Grid of Number of Bins and Estimated Choices

$J_{-,n}$	$\frac{\mathsf{IMSE}_{\mathsf{ES},-}(J_{-,n})}{\mathsf{IMSE}^*_{\mathsf{ES},-}}$	$J_{+,n}$	$\frac{\mathsf{IMSE}_{\mathsf{ES},+}(J_{+,n})}{\mathsf{IMSE}^*_{\mathsf{ES},+}}$	$J_{-,n}$	$\frac{\mathsf{IMSE}_{\mathsf{QS},-}(J_{-,n})}{\mathsf{IMSE}^*_{\mathsf{QS},-}}$	$J_{+,n}$	$\frac{\mathsf{IMSE}_{\mathsf{QS},+}(J_{+,n})}{\mathsf{IMSE}^*_{\mathsf{QS},+}}$
26	1.032	11	1.157	19	1.047	13	1.086
27	1.019	12	1.088	20	1.026	14	1.045
28	1.010	13	1.043	21	1.012	15	1.018
29	1.004	14	1.017	22	1.004	16	1.004
30	1.001	15	1.003	23	1.000	17	0.998
31	1.000	16	1.000	24	1.000	18	1.000
32	1.001	17	1.004	25	1.004	19	1.007
33	1.004	18	1.015	26	1.010	20	1.019
34	1.009	19	1.030	27	1.019	21	1.035
35	1.015	20	1.050	28	1.029	22	1.054
36	1.022	21	1.072	29	1.042	23	1.075
$\hat{J}_{\mathsf{ES}\text{-}\mu,-,n}$	1.086	$\hat{J}_{\mathsf{ES}\text{-}\mu,+,n}$	0.9009	$\hat{J}_{\mathsf{QS}\text{-}\mu,-,n}$	0.9271	$\hat{J}_{\mathsf{QS}\text{-}\mu,+,n}$	0.9399
$\check{J}_{\mathsf{ES}\text{-}\mu,-,n}$	1.088	$\check{J}_{\mathsf{ES}\text{-}\mu,+,n}$	0.9005	$\check{J}_{\mathsf{QS}\text{-}\mu,-,n}$	0.9292	$\check{J}_{\mathsf{QS}\text{-}\mu,+,n}$	0.9394

Panel B: Summary Statistics for the Estimated Number of Bins

Pop. Par.		Min.	1st Qu.	Median	Mean	3rd Qu.	Max.	Std. Dev.
$J_{\mathsf{ES}\text{-}\mu,-,n} = 31$	$\hat{J}_{\mathsf{ES}\text{-}\mu,-,n}$	30	33	34	34.13	35	39	1.09
	$\check{J}_{\mathsf{ES}\text{-}\mu,-,n}$	31	33	34	34.08	35	38	1.01
$J_{\mathsf{ES}\text{-}\vartheta,-,n} = 114$	$\hat{J}_{\mathsf{ES}\text{-}\vartheta,-,n}$	98	112	115	115.1	118.2	134	5.18
	$\check{J}_{\mathsf{ES}\text{-}\vartheta,-,n}$	104	112	114	114.5	117	126	3.05
$J_{\mathsf{ES}\text{-}\mu,+,n} = 16$	$\hat{J}_{\mathsf{ES}\text{-}\mu,+,n}$	13	14	15	14.84	15	18	0.72
	$\check{J}_{\mathsf{ES}\text{-}\mu,+,n}$	13	14	15	14.83	15	17	0.70
$J_{\mathsf{ES}\text{-}\vartheta,+,n} = 118$	$\hat{J}_{\mathsf{ES}\text{-}\vartheta,+,n}$	102	116	120	120.3	124	145	5.63
	$\check{J}_{\mathsf{ES}\text{-}\vartheta,+,n}$	110	118	120	120.2	122	133	3.22
$J_{\mathsf{QS}\text{-}\mu,-,n} = 24$	$\hat{J}_{\mathsf{QS}\text{-}\mu,-,n}$	21	22	22	22.24	23	24	0.53
	$\check{J}_{\mathsf{QS}\text{-}\mu,-,n}$	21	22	22	22.2	22	24	0.50
$J_{\mathsf{QS}\text{-}\vartheta,-,n} = 114$	$\hat{J}_{\mathsf{QS}\text{-}\vartheta,-,n}$	104	112	115	114.8	117	128	3.46
	$\check{J}_{\mathsf{QS}\text{-}\vartheta,-,n}$	106	113	114	114.4	116	124	2.56
$J_{\mathsf{QS}\text{-}\mu,+,n} = 18$	$\hat{J}_{\mathsf{QS}\text{-}\mu,+,n}$	15	16	17	16.71	17	20	0.65
	$\check{J}_{\mathsf{QS}\text{-}\mu,+,n}$	15	16	17	16.72	17	20	0.65
$J_{\mathsf{QS}\text{-}\vartheta,+,n} = 118$	$\hat{J}_{\mathsf{QS}\text{-}\vartheta,+,n}$	108	117	120	119.9	122	134	3.66
	$\check{J}_{\mathsf{QS}\text{-}\vartheta,+,n}$	109	118	120	119.9	122	132	2.81

Notes:

(i) Population quantities:
$J_{\mathsf{ES}\text{-}\mu,\cdot,n}$ = IMSE-optimal partition size for ES RD Plot.
$J_{\mathsf{ES}\text{-}\vartheta,\cdot,n}$ = Mimicking variance partition size for ES RD Plot.
$J_{\mathsf{QS}\text{-}\mu,\cdot,n}$ = IMSE-optimal partition size for QS RD Plot.
$J_{\mathsf{QS}\text{-}\vartheta,\cdot,n}$ = Mimicking variance partition size for QS RD Plot.
$\mathsf{IMSE}^*_{\mathsf{ES},\cdot} = \mathsf{IMSE}_{\mathsf{ES},\cdot}(J_{\mathsf{ES}\text{-}\mu,\cdot,n})$ = ES IMSE function evaluated at optimal choice.
$\mathsf{IMSE}^*_{\mathsf{QS},\cdot} = \mathsf{IMSE}_{\mathsf{QS},\cdot}(J_{\mathsf{QS}\text{-}\mu,\cdot,n})$ = QS IMSE function evaluated at optimal choice.

(ii) Estimators:
$\hat{J}_{\mathsf{ES}\text{-}\mu,\cdot,n}$ = spacings estimator of $J_{\mathsf{ES}\text{-}\mu,\cdot,n}$; $\check{J}_{\mathsf{ES}\text{-}\mu,\cdot,n}$ = polynomial estimator of $J_{\mathsf{ES}\text{-}\mu,\cdot,n}$.
$\hat{J}_{\mathsf{ES}\text{-}\vartheta,\cdot,n}$ = spacings estimator of $J_{\mathsf{ES}\text{-}\vartheta,\cdot,n}$; $\check{J}_{\mathsf{ES}\text{-}\vartheta,\cdot,n}$ = polynomial estimator of $J_{\mathsf{ES}\text{-}\vartheta,\cdot,n}$.
$\hat{J}_{\mathsf{QS}\text{-}\mu,\cdot,n}$ = spacings estimator of $J_{\mathsf{QS}\text{-}\mu,\cdot,n}$; $\check{J}_{\mathsf{QS}\text{-}\mu,\cdot,n}$ = polynomial estimator of $J_{\mathsf{QS}\text{-}\mu,\cdot,n}$.
$\hat{J}_{\mathsf{QS}\text{-}\vartheta,\cdot,n}$ = spacings estimator of $J_{\mathsf{QS}\text{-}\vartheta,\cdot,n}$; $\check{J}_{\mathsf{QS}\text{-}\vartheta,\cdot,n}$ = polynomial estimator of $J_{\mathsf{QS}\text{-}\vartheta,\cdot,n}$.

Table SA-4: Simulations Results for Model 3

Panel A: IMSE for Grid of Number of Bins and Estimated Choices

$J_{-,n}$	$\frac{\mathsf{IMSE}_{\mathsf{ES},-}(J_{-,n})}{\mathsf{IMSE}^*_{\mathsf{ES},-}}$	$J_{+,n}$	$\frac{\mathsf{IMSE}_{\mathsf{ES},+}(J_{+,n})}{\mathsf{IMSE}^*_{\mathsf{ES},+}}$	$J_{-,n}$	$\frac{\mathsf{IMSE}_{\mathsf{QS},-}(J_{-,n})}{\mathsf{IMSE}^*_{\mathsf{QS},-}}$	$J_{+,n}$	$\frac{\mathsf{IMSE}_{\mathsf{QS},+}(J_{+,n})}{\mathsf{IMSE}^*_{\mathsf{QS},+}}$
49	1.008	8	1.279	40	1.010	8	1.265
50	1.005	9	1.149	41	1.006	9	1.139
51	1.002	10	1.071	42	1.002	10	1.064
52	1.001	11	1.027	43	1.000	11	1.023
53	1.000	12	1.005	44	1.000	12	1.003
54	1.000	13	1.000	45	1.000	13	1.000
55	1.001	14	1.007	46	1.001	14	1.008
56	1.002	15	1.022	47	1.003	15	1.025
57	1.004	16	1.044	48	1.006	16	1.048
58	1.006	17	1.071	49	1.010	17	1.076
59	1.009	18	1.102	50	1.014	18	1.108
$\hat{J}_{\mathsf{ES}\text{-}\mu,-,n}$	1.09	$\hat{J}_{\mathsf{ES}\text{-}\mu,+,n}$	0.9534	$\hat{J}_{\mathsf{QS}\text{-}\mu,-,n}$	0.869	$\hat{J}_{\mathsf{QS}\text{-}\mu,+,n}$	0.9628
$\check{J}_{\mathsf{ES}\text{-}\mu,-,n}$	1.097	$\check{J}_{\mathsf{ES}\text{-}\mu,+,n}$	0.9504	$\check{J}_{\mathsf{QS}\text{-}\mu,-,n}$	0.872	$\check{J}_{\mathsf{QS}\text{-}\mu,+,n}$	0.9609

Panel B: Summary Statistics for the Estimated Number of Bins

Pop. Par.		Min.	1st Qu.	Median	Mean	3rd Qu.	Max.	Std. Dev.
$J_{\mathsf{ES}\text{-}\mu,-,n}=54$	$\hat{J}_{\mathsf{ES}\text{-}\mu,-,n}$	54	58	59	59.05	60	65	1.59
	$\check{J}_{\mathsf{ES}\text{-}\mu,-,n}$	54	58	59	58.85	60	64	1.28
$J_{\mathsf{ES}\text{-}\vartheta,-,n}=112$	$\hat{J}_{\mathsf{ES}\text{-}\vartheta,-,n}$	90	108	112	112.1	116	138	6.65
	$\check{J}_{\mathsf{ES}\text{-}\vartheta,-,n}$	99	108	111	110.9	114	127	4.08
$J_{\mathsf{ES}\text{-}\mu,+,n}=13$	$\hat{J}_{\mathsf{ES}\text{-}\mu,+,n}$	11	12	13	12.79	13	16	0.73
	$\check{J}_{\mathsf{ES}\text{-}\mu,+,n}$	11	12	13	12.8	13	16	0.68
$J_{\mathsf{ES}\text{-}\vartheta,+,n}=149$	$\hat{J}_{\mathsf{ES}\text{-}\vartheta,+,n}$	111	140	147	147.6	155	193	10.94
	$\check{J}_{\mathsf{ES}\text{-}\vartheta,+,n}$	125	143	148	147.8	152	174	6.47
$J_{\mathsf{QS}\text{-}\mu,-,n}=45$	$\hat{J}_{\mathsf{QS}\text{-}\mu,-,n}$	36	38	39	38.8	39	42	0.82
	$\check{J}_{\mathsf{QS}\text{-}\mu,-,n}$	36	38	39	38.72	39	42	0.78
$J_{\mathsf{QS}\text{-}\vartheta,-,n}=155$	$\hat{J}_{\mathsf{QS}\text{-}\vartheta,-,n}$	140	151	154	154.2	157	168	4.07
	$\check{J}_{\mathsf{QS}\text{-}\vartheta,-,n}$	142	151	153	153.3	155	165	3.12
$J_{\mathsf{QS}\text{-}\mu,+,n}=13$	$\hat{J}_{\mathsf{QS}\text{-}\mu,+,n}$	11	12	13	12.74	13	15	0.61
	$\check{J}_{\mathsf{QS}\text{-}\mu,+,n}$	11	12	13	12.76	13	15	0.59
$J_{\mathsf{QS}\text{-}\vartheta,+,n}=149$	$\hat{J}_{\mathsf{QS}\text{-}\vartheta,+,n}$	119	142	147	147.5	153	182	8.29
	$\check{J}_{\mathsf{QS}\text{-}\vartheta,+,n}$	125	143	147	147.8	152	174	6.47

Notes:

(i) Population quantities:
$J_{\mathsf{ES}\text{-}\mu,\cdot,n}$ = IMSE-optimal partition size for ES RD Plot.
$J_{\mathsf{ES}\text{-}\vartheta,\cdot,n}$ = Mimicking variance partition size for ES RD Plot.
$J_{\mathsf{QS}\text{-}\mu,\cdot,n}$ = IMSE-optimal partition size for QS RD Plot.
$J_{\mathsf{QS}\text{-}\vartheta,\cdot,n}$ = Mimicking variance partition size for QS RD Plot.
$\mathsf{IMSE}^*_{\mathsf{ES},\cdot} = \mathsf{IMSE}_{\mathsf{ES},\cdot}(J_{\mathsf{ES}\text{-}\mu,\cdot,n})$ = ES IMSE function evaluated at optimal choice.
$\mathsf{IMSE}^*_{\mathsf{QS},\cdot} = \mathsf{IMSE}_{\mathsf{QS},\cdot}(J_{\mathsf{QS}\text{-}\mu,\cdot,n})$ = QS IMSE function evaluated at optimal choice.

(ii) Estimators:
$\hat{J}_{\mathsf{ES}\text{-}\mu,\cdot,n}$ = spacings estimator of $J_{\mathsf{ES}\text{-}\mu,\cdot,n}$; $\check{J}_{\mathsf{ES}\text{-}\mu,\cdot,n}$ = polynomial estimator of $J_{\mathsf{ES}\text{-}\mu,\cdot,n}$.
$\hat{J}_{\mathsf{ES}\text{-}\vartheta,\cdot,n}$ = spacings estimator of $J_{\mathsf{ES}\text{-}\vartheta,\cdot,n}$; $\check{J}_{\mathsf{ES}\text{-}\vartheta,\cdot,n}$ = polynomial estimator of $J_{\mathsf{ES}\text{-}\vartheta,\cdot,n}$.
$\hat{J}_{\mathsf{QS}\text{-}\mu,\cdot,n}$ = spacings estimator of $J_{\mathsf{QS}\text{-}\mu,\cdot,n}$; $\check{J}_{\mathsf{QS}\text{-}\mu,\cdot,n}$ = polynomial estimator of $J_{\mathsf{QS}\text{-}\mu,\cdot,n}$.
$\hat{J}_{\mathsf{QS}\text{-}\vartheta,\cdot,n}$ = spacings estimator of $J_{\mathsf{QS}\text{-}\vartheta,\cdot,n}$; $\check{J}_{\mathsf{QS}\text{-}\vartheta,\cdot,n}$ = polynomial estimator of $J_{\mathsf{QS}\text{-}\vartheta,\cdot,n}$.

Table SA-5: Simulations Results for Model 4

Panel A: IMSE for Grid of Number of Bins and Estimated Choices

$J_{-,n}$	$\frac{\mathsf{IMSE}_{\mathtt{ES},-}(J_{-,n})}{\mathsf{IMSE}^*_{\mathtt{ES},-}}$	$J_{+,n}$	$\frac{\mathsf{IMSE}_{\mathtt{ES},+}(J_{+,n})}{\mathsf{IMSE}^*_{\mathtt{ES},+}}$	$J_{-,n}$	$\frac{\mathsf{IMSE}_{\mathtt{QS},-}(J_{-,n})}{\mathsf{IMSE}^*_{\mathtt{QS},-}}$	$J_{+,n}$	$\frac{\mathsf{IMSE}_{\mathtt{QS},+}(J_{+,n})}{\mathsf{IMSE}^*_{\mathtt{QS},+}}$
16	1.080	19	1.059	15	1.072	30	1.025
17	1.047	20	1.035	16	1.039	31	1.015
18	1.024	21	1.018	17	1.017	32	1.008
19	1.010	22	1.008	18	1.005	33	1.003
20	1.002	23	1.002	19	1.000	34	1.001
21	1.000	24	1.000	20	1.000	35	1.000
22	1.002	25	1.002	21	1.005	36	1.001
23	1.009	26	1.006	22	1.014	37	1.003
24	1.018	27	1.014	23	1.027	38	1.007
25	1.030	28	1.023	24	1.042	39	1.011
26	1.044	29	1.034	25	1.059	40	1.017
$\hat{J}_{\mathtt{ES}\text{-}\mu,-,n}$	1.065	$\hat{J}_{\mathtt{ES}\text{-}\mu,+,n}$	0.8511	$\hat{J}_{\mathtt{QS}\text{-}\mu,-,n}$	0.9663	$\hat{J}_{\mathtt{QS}\text{-}\mu,+,n}$	0.9004
$\check{J}_{\mathtt{ES}\text{-}\mu,-,n}$	1.067	$\check{J}_{\mathtt{ES}\text{-}\mu,+,n}$	0.8504	$\check{J}_{\mathtt{QS}\text{-}\mu,-,n}$	0.9679	$\check{J}_{\mathtt{QS}\text{-}\mu,+,n}$	0.9003

Panel B: Summary Statistics for the Estimated Number of Bins

Pop. Par.		Min.	1st Qu.	Median	Mean	3rd Qu.	Max.	Std. Dev.
$J_{\mathtt{ES}\text{-}\mu,-,n} = 21$	$\hat{J}_{\mathtt{ES}\text{-}\mu,-,n}$	19	22	23	22.86	24	28	1.04
	$\check{J}_{\mathtt{ES}\text{-}\mu,-,n}$	19	22	23	22.83	23	26	0.91
$J_{\mathtt{ES}\text{-}\vartheta,-,n} = 145$	$\hat{J}_{\mathtt{ES}\text{-}\vartheta,-,n}$	106	141	148	148.3	156	201	11.48
	$\check{J}_{\mathtt{ES}\text{-}\vartheta,-,n}$	125	143	147	147.6	152	179	6.59
$J_{\mathtt{ES}\text{-}\mu,+,n} = 24$	$\hat{J}_{\mathtt{ES}\text{-}\mu,+,n}$	17	20	21	20.91	22	27	1.33
	$\check{J}_{\mathtt{ES}\text{-}\mu,+,n}$	17	20	21	20.91	22	27	1.30
$J_{\mathtt{ES}\text{-}\vartheta,+,n} = 102$	$\hat{J}_{\mathtt{ES}\text{-}\vartheta,+,n}$	82	99	103	103.6	108	130	6.29
	$\check{J}_{\mathtt{ES}\text{-}\vartheta,+,n}$	90	101	103	103.5	106	119	3.95
$J_{\mathtt{QS}\text{-}\mu,-,n} = 20$	$\hat{J}_{\mathtt{QS}\text{-}\mu,-,n}$	17	19	19	19.44	20	23	0.74
	$\check{J}_{\mathtt{QS}\text{-}\mu,-,n}$	17	19	19	19.43	20	22	0.70
$J_{\mathtt{QS}\text{-}\vartheta,-,n} = 145$	$\hat{J}_{\mathtt{QS}\text{-}\vartheta,-,n}$	120	144	149	149.6	155	187	8.59
	$\check{J}_{\mathtt{QS}\text{-}\vartheta,-,n}$	126	145	149	149.1	153	181	6.60
$J_{\mathtt{QS}\text{-}\mu,+,n} = 35$	$\hat{J}_{\mathtt{QS}\text{-}\mu,+,n}$	28	31	32	31.91	33	40	1.61
	$\check{J}_{\mathtt{QS}\text{-}\mu,+,n}$	28	31	32	31.92	33	40	1.61
$J_{\mathtt{QS}\text{-}\vartheta,+,n} = 141$	$\hat{J}_{\mathtt{QS}\text{-}\vartheta,+,n}$	130	140	143	142.9	146	159	3.97
	$\check{J}_{\mathtt{QS}\text{-}\vartheta,+,n}$	131	141	143	142.9	145	155	3.25

Notes:

(i) Population quantities:
$J_{\mathtt{ES}\text{-}\mu,\cdot,n}$ = IMSE-optimal partition size for ES RD Plot.
$J_{\mathtt{ES}\text{-}\vartheta,\cdot,n}$ = Mimicking variance partition size for ES RD Plot.
$J_{\mathtt{QS}\text{-}\mu,\cdot,n}$ = IMSE-optimal partition size for QS RD Plot.
$J_{\mathtt{QS}\text{-}\vartheta,\cdot,n}$ = Mimicking variance partition size for QS RD Plot.
$\mathsf{IMSE}^*_{\mathtt{ES},\cdot} = \mathsf{IMSE}_{\mathtt{ES},\cdot}(J_{\mathtt{ES}\text{-}\mu,\cdot,n})$ = ES IMSE function evaluated at optimal choice.
$\mathsf{IMSE}^*_{\mathtt{QS},\cdot} = \mathsf{IMSE}_{\mathtt{QS},\cdot}(J_{\mathtt{QS}\text{-}\mu,\cdot,n})$ = QS IMSE function evaluated at optimal choice.

(ii) Estimators:
$\hat{J}_{\mathtt{ES}\text{-}\mu,\cdot,n}$ = spacings estimator of $J_{\mathtt{ES}\text{-}\mu,\cdot,n}$; $\check{J}_{\mathtt{ES}\text{-}\mu,\cdot,n}$ = polynomial estimator of $J_{\mathtt{ES}\text{-}\mu,\cdot,n}$.
$\hat{J}_{\mathtt{ES}\text{-}\vartheta,\cdot,n}$ = spacings estimator of $J_{\mathtt{ES}\text{-}\vartheta,\cdot,n}$; $\check{J}_{\mathtt{ES}\text{-}\vartheta,\cdot,n}$ = polynomial estimator of $J_{\mathtt{ES}\text{-}\vartheta,\cdot,n}$.
$\hat{J}_{\mathtt{QS}\text{-}\mu,\cdot,n}$ = spacings estimator of $J_{\mathtt{QS}\text{-}\mu,\cdot,n}$; $\check{J}_{\mathtt{QS}\text{-}\mu,\cdot,n}$ = polynomial estimator of $J_{\mathtt{QS}\text{-}\mu,\cdot,n}$.
$\hat{J}_{\mathtt{QS}\text{-}\vartheta,\cdot,n}$ = spacings estimator of $J_{\mathtt{QS}\text{-}\vartheta,\cdot,n}$; $\check{J}_{\mathtt{QS}\text{-}\vartheta,\cdot,n}$ = polynomial estimator of $J_{\mathtt{QS}\text{-}\vartheta,\cdot,n}$.

Table SA-6: Simulations Results for Model 5

Panel A: IMSE for Grid of Number of Bins and Estimated Choices

$J_{-,n}$	$\frac{\mathsf{IMSE}_{\mathsf{ES},-}(J_{-,n})}{\mathsf{IMSE}^*_{\mathsf{ES},-}}$	$J_{+,n}$	$\frac{\mathsf{IMSE}_{\mathsf{ES},+}(J_{+,n})}{\mathsf{IMSE}^*_{\mathsf{ES},+}}$	$J_{-,n}$	$\frac{\mathsf{IMSE}_{\mathsf{QS},-}(J_{-,n})}{\mathsf{IMSE}^*_{\mathsf{QS},-}}$	$J_{+,n}$	$\frac{\mathsf{IMSE}_{\mathsf{QS},+}(J_{+,n})}{\mathsf{IMSE}^*_{\mathsf{QS},+}}$
41	1.013	7	1.247	30	1.016	6	1.472
42	1.008	8	1.113	31	1.008	7	1.240
43	1.004	9	1.039	32	1.003	8	1.110
44	1.002	10	1.004	33	1.000	9	1.041
45	1.000	11	0.994	34	0.999	10	1.008
46	1.000	12	1.000	35	1.000	11	1.000
47	1.001	13	1.018	36	1.002	12	1.008
48	1.002	14	1.045	37	1.006	13	1.028
49	1.004	15	1.078	38	1.011	14	1.057
50	1.007	16	1.116	39	1.017	15	1.092
51	1.011	17	1.158	40	1.024	16	1.131
$\hat{J}_{\mathsf{ES}\text{-}\mu,-,n}$	1.095	$\hat{J}_{\mathsf{ES}\text{-}\mu,+,n}$	0.9544	$\hat{J}_{\mathsf{QS}\text{-}\mu,-,n}$	0.8966	$\hat{J}_{\mathsf{QS}\text{-}\mu,+,n}$	0.9651
$\check{J}_{\mathsf{ES}\text{-}\mu,-,n}$	1.099	$\check{J}_{\mathsf{ES}\text{-}\mu,+,n}$	0.9521	$\check{J}_{\mathsf{QS}\text{-}\mu,-,n}$	0.8977	$\check{J}_{\mathsf{QS}\text{-}\mu,+,n}$	0.9629

Panel B: Summary Statistics for the Estimated Number of Bins

Pop. Par.		Min.	1st Qu.	Median	Mean	3rd Qu.	Max.	Std. Dev.
$J_{\mathsf{ES}\text{-}\mu,-,n} = 46$	$\hat{J}_{\mathsf{ES}\text{-}\mu,-,n}$	44	50	51	50.93	52	58	1.83
	$\check{J}_{\mathsf{ES}\text{-}\mu,-,n}$	45	50	51	50.82	52	57	1.61
$J_{\mathsf{ES}\text{-}\vartheta,-,n} = 109$	$\hat{J}_{\mathsf{ES}\text{-}\vartheta,-,n}$	77	104	110	109.9	115	139	8.11
	$\check{J}_{\mathsf{ES}\text{-}\vartheta,-,n}$	92	105	109	109.1	113	130	5.60
$J_{\mathsf{ES}\text{-}\mu,+,n} = 12$	$\hat{J}_{\mathsf{ES}\text{-}\mu,+,n}$	9	11	11	11.17	12	15	0.74
	$\check{J}_{\mathsf{ES}\text{-}\mu,+,n}$	9	11	11	11.17	12	14	0.69
$J_{\mathsf{ES}\text{-}\vartheta,+,n} = 119$	$\hat{J}_{\mathsf{ES}\text{-}\vartheta,+,n}$	82	113	120	120.4	127	161	10.36
	$\check{J}_{\mathsf{ES}\text{-}\vartheta,+,n}$	102	116	120	120.4	124	141	5.89
$J_{\mathsf{QS}\text{-}\mu,-,n} = 35$	$\hat{J}_{\mathsf{QS}\text{-}\mu,-,n}$	28	30	31	31.02	32	35	0.86
	$\check{J}_{\mathsf{QS}\text{-}\mu,-,n}$	28	30	31	31	31	35	0.84
$J_{\mathsf{QS}\text{-}\vartheta,-,n} = 109$	$\hat{J}_{\mathsf{QS}\text{-}\vartheta,-,n}$	99	107	109	109.4	111	120	2.75
	$\check{J}_{\mathsf{QS}\text{-}\vartheta,-,n}$	101	108	109	109.1	111	117	2.17
$J_{\mathsf{QS}\text{-}\mu,+,n} = 11$	$\hat{J}_{\mathsf{QS}\text{-}\mu,+,n}$	9	11	11	11.06	11	13	0.59
	$\check{J}_{\mathsf{QS}\text{-}\mu,+,n}$	9	11	11	11.06	11	13	0.57
$J_{\mathsf{QS}\text{-}\vartheta,+,n} = 119$	$\hat{J}_{\mathsf{QS}\text{-}\vartheta,+,n}$	99	115	119	119.8	124	149	6.86
	$\check{J}_{\mathsf{QS}\text{-}\vartheta,+,n}$	101	116	120	120.1	124	140	5.55

Notes:

(i) Population quantities:
$J_{\mathsf{ES}\text{-}\mu,\cdot,n}$ = IMSE-optimal partition size for ES RD Plot.
$J_{\mathsf{ES}\text{-}\vartheta,\cdot,n}$ = Mimicking variance partition size for ES RD Plot.
$J_{\mathsf{QS}\text{-}\mu,\cdot,n}$ = IMSE-optimal partition size for QS RD Plot.
$J_{\mathsf{QS}\text{-}\vartheta,\cdot,n}$ = Mimicking variance partition size for QS RD Plot.
$\mathsf{IMSE}^*_{\mathsf{ES},\cdot} = \mathsf{IMSE}_{\mathsf{ES},\cdot}(J_{\mathsf{ES}\text{-}\mu,\cdot,n})$ = ES IMSE function evaluated at optimal choice.
$\mathsf{IMSE}^*_{\mathsf{QS},\cdot} = \mathsf{IMSE}_{\mathsf{QS},\cdot}(J_{\mathsf{QS}\text{-}\mu,\cdot,n})$ = QS IMSE function evaluated at optimal choice.

(ii) Estimators:
$\hat{J}_{\mathsf{ES}\text{-}\mu,\cdot,n}$ = spacings estimator of $J_{\mathsf{ES}\text{-}\mu,\cdot,n}$; $\check{J}_{\mathsf{ES}\text{-}\mu,\cdot,n}$ = polynomial estimator of $J_{\mathsf{ES}\text{-}\mu,\cdot,n}$.
$\hat{J}_{\mathsf{ES}\text{-}\vartheta,\cdot,n}$ = spacings estimator of $J_{\mathsf{ES}\text{-}\vartheta,\cdot,n}$; $\check{J}_{\mathsf{ES}\text{-}\vartheta,\cdot,n}$ = polynomial estimator of $J_{\mathsf{ES}\text{-}\vartheta,\cdot,n}$.
$\hat{J}_{\mathsf{QS}\text{-}\mu,\cdot,n}$ = spacings estimator of $J_{\mathsf{QS}\text{-}\mu,\cdot,n}$; $\check{J}_{\mathsf{QS}\text{-}\mu,\cdot,n}$ = polynomial estimator of $J_{\mathsf{QS}\text{-}\mu,\cdot,n}$.
$\hat{J}_{\mathsf{QS}\text{-}\vartheta,\cdot,n}$ = spacings estimator of $J_{\mathsf{QS}\text{-}\vartheta,\cdot,n}$; $\check{J}_{\mathsf{QS}\text{-}\vartheta,\cdot,n}$ = polynomial estimator of $J_{\mathsf{QS}\text{-}\vartheta,\cdot,n}$.

Table SA-7: Simulations Results for Model 6

Panel A: IMSE for Grid of Number of Bins and Estimated Choices

$J_{-,n}$	$\frac{\mathsf{IMSE}_{\mathtt{ES},-}(J_{-,n})}{\mathsf{IMSE}^*_{\mathtt{ES},-}}$	$J_{+,n}$	$\frac{\mathsf{IMSE}_{\mathtt{ES},+}(J_{+,n})}{\mathsf{IMSE}^*_{\mathtt{ES},+}}$	$J_{-,n}$	$\frac{\mathsf{IMSE}_{\mathtt{QS},-}(J_{-,n})}{\mathsf{IMSE}^*_{\mathtt{QS},-}}$	$J_{+,n}$	$\frac{\mathsf{IMSE}_{\mathtt{QS},+}(J_{+,n})}{\mathsf{IMSE}^*_{\mathtt{QS},+}}$
13	1.119	16	1.069	12	1.121	22	1.044
14	1.068	17	1.039	13	1.066	23	1.026
15	1.035	18	1.018	14	1.031	24	1.014
16	1.014	19	1.006	15	1.011	25	1.005
17	1.003	20	1.000	16	1.001	26	1.001
18	1.000	21	1.000	17	1.000	27	1.000
19	1.003	22	1.004	18	1.006	28	1.002
20	1.011	23	1.012	19	1.017	29	1.005
21	1.023	24	1.022	20	1.032	30	1.011
22	1.039	25	1.035	21	1.050	31	1.019
23	1.057	26	1.051	22	1.072	32	1.029
$\hat{J}_{\mathtt{ES}\text{-}\mu,-,n}$	1.065	$\hat{J}_{\mathtt{ES}\text{-}\mu,+,n}$	0.8495	$\hat{J}_{\mathtt{QS}\text{-}\mu,-,n}$	1.008	$\hat{J}_{\mathtt{QS}\text{-}\mu,+,n}$	0.9261
$\check{J}_{\mathtt{ES}\text{-}\mu,-,n}$	1.065	$\check{J}_{\mathtt{ES}\text{-}\mu,+,n}$	0.8493	$\check{J}_{\mathtt{QS}\text{-}\mu,-,n}$	1.008	$\check{J}_{\mathtt{QS}\text{-}\mu,+,n}$	0.9264

Panel B: Summary Statistics for the Estimated Number of Bins

Pop. Par.		Min.	1st Qu.	Median	Mean	3rd Qu.	Max.	Std. Dev.
$J_{\mathtt{ES}\text{-}\mu,-,n} = 18$	$\hat{J}_{\mathtt{ES}\text{-}\mu,-,n}$	16	19	20	19.71	20	24	1.23
	$\check{J}_{\mathtt{ES}\text{-}\mu,-,n}$	16	19	20	19.69	20	24	1.17
$J_{\mathtt{ES}\text{-}\vartheta,-,n} = 119$	$\hat{J}_{\mathtt{ES}\text{-}\vartheta,-,n}$	87	113	120	120.2	127	165	9.88
	$\check{J}_{\mathtt{ES}\text{-}\vartheta,-,n}$	102	116	120	119.7	124	145	5.92
$J_{\mathtt{ES}\text{-}\mu,+,n} = 21$	$\hat{J}_{\mathtt{ES}\text{-}\mu,+,n}$	13	17	18	18.14	19	25	1.71
	$\check{J}_{\mathtt{ES}\text{-}\mu,+,n}$	14	17	18	18.13	19	26	1.69
$J_{\mathtt{ES}\text{-}\vartheta,+,n} = 102$	$\hat{J}_{\mathtt{ES}\text{-}\vartheta,+,n}$	75	97	102	102.4	108	137	7.90
	$\check{J}_{\mathtt{ES}\text{-}\vartheta,+,n}$	82	98	102	102.2	106	124	5.77
$J_{\mathtt{QS}\text{-}\mu,-,n} = 17$	$\hat{J}_{\mathtt{QS}\text{-}\mu,-,n}$	15	17	17	17.31	18	20	0.94
	$\check{J}_{\mathtt{QS}\text{-}\mu,-,n}$	15	17	17	17.31	18	20	0.92
$J_{\mathtt{QS}\text{-}\vartheta,-,n} = 119$	$\hat{J}_{\mathtt{QS}\text{-}\vartheta,-,n}$	97	115	120	119.8	124	146	6.81
	$\check{J}_{\mathtt{QS}\text{-}\vartheta,-,n}$	104	116	119	119.6	123	142	5.43
$J_{\mathtt{QS}\text{-}\mu,+,n} = 27$	$\hat{J}_{\mathtt{QS}\text{-}\mu,+,n}$	22	25	25	25.42	26	31	1.32
	$\check{J}_{\mathtt{QS}\text{-}\mu,+,n}$	22	25	25	25.42	26	31	1.31
$J_{\mathtt{QS}\text{-}\vartheta,+,n} = 102$	$\hat{J}_{\mathtt{QS}\text{-}\vartheta,+,n}$	94	100	101	101.3	103	109	2.43
	$\check{J}_{\mathtt{QS}\text{-}\vartheta,+,n}$	96	100	101	101.2	102	109	1.85

Notes:

(i) Population quantities:
$J_{\mathtt{ES}\text{-}\mu,\cdot,n}$ = IMSE-optimal partition size for ES RD Plot.
$J_{\mathtt{ES}\text{-}\vartheta,\cdot,n}$ = Mimicking variance partition size for ES RD Plot.
$J_{\mathtt{QS}\text{-}\mu,\cdot,n}$ = IMSE-optimal partition size for QS RD Plot.
$J_{\mathtt{QS}\text{-}\vartheta,\cdot,n}$ = Mimicking variance partition size for QS RD Plot.
$\mathsf{IMSE}^*_{\mathtt{ES},\cdot} = \mathsf{IMSE}_{\mathtt{ES},\cdot}(J_{\mathtt{ES}\text{-}\mu,\cdot,n})$ = ES IMSE function evaluated at optimal choice.
$\mathsf{IMSE}^*_{\mathtt{QS},\cdot} = \mathsf{IMSE}_{\mathtt{QS},\cdot}(J_{\mathtt{QS}\text{-}\mu,\cdot,n})$ = QS IMSE function evaluated at optimal choice.

(ii) Estimators:
$\hat{J}_{\mathtt{ES}\text{-}\mu,\cdot,n}$ = spacings estimator of $J_{\mathtt{ES}\text{-}\mu,\cdot,n}$; $\check{J}_{\mathtt{ES}\text{-}\mu,\cdot,n}$ = polynomial estimator of $J_{\mathtt{ES}\text{-}\mu,\cdot,n}$.
$\hat{J}_{\mathtt{ES}\text{-}\vartheta,\cdot,n}$ = spacings estimator of $J_{\mathtt{ES}\text{-}\vartheta,\cdot,n}$; $\check{J}_{\mathtt{ES}\text{-}\vartheta,\cdot,n}$ = polynomial estimator of $J_{\mathtt{ES}\text{-}\vartheta,\cdot,n}$.
$\hat{J}_{\mathtt{QS}\text{-}\mu,\cdot,n}$ = spacings estimator of $J_{\mathtt{QS}\text{-}\mu,\cdot,n}$; $\check{J}_{\mathtt{QS}\text{-}\mu,\cdot,n}$ = polynomial estimator of $J_{\mathtt{QS}\text{-}\mu,\cdot,n}$.
$\hat{J}_{\mathtt{QS}\text{-}\vartheta,\cdot,n}$ = spacings estimator of $J_{\mathtt{QS}\text{-}\vartheta,\cdot,n}$; $\check{J}_{\mathtt{QS}\text{-}\vartheta,\cdot,n}$ = polynomial estimator of $J_{\mathtt{QS}\text{-}\vartheta,\cdot,n}$.

Table SA-8: Simulations Results for Model 7

Panel A: IMSE for Grid of Number of Bins and Estimated Choices

$J_{-,n}$	$\frac{\mathsf{IMSE}_{\mathsf{ES},-}(J_{-,n})}{\mathsf{IMSE}^*_{\mathsf{ES},-}}$	$J_{+,n}$	$\frac{\mathsf{IMSE}_{\mathsf{ES},+}(J_{+,n})}{\mathsf{IMSE}^*_{\mathsf{ES},+}}$	$J_{-,n}$	$\frac{\mathsf{IMSE}_{\mathsf{QS},-}(J_{-,n})}{\mathsf{IMSE}^*_{\mathsf{QS},-}}$	$J_{+,n}$	$\frac{\mathsf{IMSE}_{\mathsf{QS},+}(J_{+,n})}{\mathsf{IMSE}^*_{\mathsf{QS},+}}$
49	1.008	8	1.279	40	1.010	8	1.265
50	1.005	9	1.149	41	1.006	9	1.139
51	1.002	10	1.071	42	1.002	10	1.064
52	1.001	11	1.027	43	1.000	11	1.023
53	1.000	12	1.005	44	1.000	12	1.003
54	1.000	13	1.000	45	1.000	13	1.000
55	1.001	14	1.007	46	1.001	14	1.008
56	1.002	15	1.022	47	1.003	15	1.025
57	1.004	16	1.044	48	1.006	16	1.048
58	1.006	17	1.071	49	1.010	17	1.076
59	1.009	18	1.102	50	1.014	18	1.108
$\hat{J}_{\mathsf{ES}\text{-}\mu,-,n}$	1.097	$\hat{J}_{\mathsf{ES}\text{-}\mu,+,n}$	0.9335	$\hat{J}_{\mathsf{QS}\text{-}\mu,-,n}$	0.9043	$\hat{J}_{\mathsf{QS}\text{-}\mu,+,n}$	0.9649
$\check{J}_{\mathsf{ES}\text{-}\mu,-,n}$	1.104	$\check{J}_{\mathsf{ES}\text{-}\mu,+,n}$	0.9308	$\check{J}_{\mathsf{QS}\text{-}\mu,-,n}$	0.9079	$\check{J}_{\mathsf{QS}\text{-}\mu,+,n}$	0.9629

Panel B: Summary Statistics for the Estimated Number of Bins

Pop. Par.		Min.	1st Qu.	Median	Mean	3rd Qu.	Max.	Std. Dev.
$J_{\mathsf{ES}\text{-}\mu,-,n} = 54$	$\hat{J}_{\mathsf{ES}\text{-}\mu,-,n}$	53	58	59	59.38	61	66	1.98
	$\check{J}_{\mathsf{ES}\text{-}\mu,-,n}$	54	58	59	59.15	60	65	1.60
$J_{\mathsf{ES}\text{-}\vartheta,-,n} = 113$	$\hat{J}_{\mathsf{ES}\text{-}\vartheta,-,n}$	82	108	114	114.1	120	149	8.89
	$\check{J}_{\mathsf{ES}\text{-}\vartheta,-,n}$	94	108	113	112.7	117	137	6.08
$J_{\mathsf{ES}\text{-}\mu,+,n} = 13$	$\hat{J}_{\mathsf{ES}\text{-}\mu,+,n}$	10	12	13	12.57	13	17	0.82
	$\check{J}_{\mathsf{ES}\text{-}\mu,+,n}$	10	12	13	12.59	13	16	0.76
$J_{\mathsf{ES}\text{-}\vartheta,+,n} = 144$	$\hat{J}_{\mathsf{ES}\text{-}\vartheta,+,n}$	105	142	152	152.6	162	227	15.09
	$\check{J}_{\mathsf{ES}\text{-}\vartheta,+,n}$	117	146	152	152.5	159	188	9.37
$J_{\mathsf{QS}\text{-}\mu,-,n} = 45$	$\hat{J}_{\mathsf{QS}\text{-}\mu,-,n}$	38	40	40	40.33	41	44	0.84
	$\check{J}_{\mathsf{QS}\text{-}\mu,-,n}$	38	40	40	40.24	41	44	0.82
$J_{\mathsf{QS}\text{-}\vartheta,-,n} = 156$	$\hat{J}_{\mathsf{QS}\text{-}\vartheta,-,n}$	138	153	156	156.6	160	177	4.77
	$\check{J}_{\mathsf{QS}\text{-}\vartheta,-,n}$	142	153	156	155.6	158	170	3.95
$J_{\mathsf{QS}\text{-}\mu,+,n} = 13$	$\hat{J}_{\mathsf{QS}\text{-}\mu,+,n}$	11	12	13	12.7	13	16	0.69
	$\check{J}_{\mathsf{QS}\text{-}\mu,+,n}$	11	12	13	12.71	13	16	0.67
$J_{\mathsf{QS}\text{-}\vartheta,+,n} = 145$	$\hat{J}_{\mathsf{QS}\text{-}\vartheta,+,n}$	112	143	150	150.8	158	208	11.11
	$\check{J}_{\mathsf{QS}\text{-}\vartheta,+,n}$	115	144	151	151.1	157	188	9.56

Notes:

(i) Population quantities:
$J_{\mathsf{ES}\text{-}\mu,\cdot,n}$ = IMSE-optimal partition size for ES RD Plot.
$J_{\mathsf{ES}\text{-}\vartheta,\cdot,n}$ = Mimicking variance partition size for ES RD Plot.
$J_{\mathsf{QS}\text{-}\mu,\cdot,n}$ = IMSE-optimal partition size for QS RD Plot.
$J_{\mathsf{QS}\text{-}\vartheta,\cdot,n}$ = Mimicking variance partition size for QS RD Plot.
$\mathsf{IMSE}^*_{\mathsf{ES},\cdot} = \mathsf{IMSE}_{\mathsf{ES},\cdot}(J_{\mathsf{ES}\text{-}\mu,\cdot,n})$ = ES IMSE function evaluated at optimal choice.
$\mathsf{IMSE}^*_{\mathsf{QS},\cdot} = \mathsf{IMSE}_{\mathsf{QS},\cdot}(J_{\mathsf{QS}\text{-}\mu,\cdot,n})$ = QS IMSE function evaluated at optimal choice.

(ii) Estimators:
$\hat{J}_{\mathsf{ES}\text{-}\mu,\cdot,n}$ = spacings estimator of $J_{\mathsf{ES}\text{-}\mu,\cdot,n}$; $\check{J}_{\mathsf{ES}\text{-}\mu,\cdot,n}$ = polynomial estimator of $J_{\mathsf{ES}\text{-}\mu,\cdot,n}$.
$\hat{J}_{\mathsf{ES}\text{-}\vartheta,\cdot,n}$ = spacings estimator of $J_{\mathsf{ES}\text{-}\vartheta,\cdot,n}$; $\check{J}_{\mathsf{ES}\text{-}\vartheta,\cdot,n}$ = polynomial estimator of $J_{\mathsf{ES}\text{-}\vartheta,\cdot,n}$.
$\hat{J}_{\mathsf{QS}\text{-}\mu,\cdot,n}$ = spacings estimator of $J_{\mathsf{QS}\text{-}\mu,\cdot,n}$; $\check{J}_{\mathsf{QS}\text{-}\mu,\cdot,n}$ = polynomial estimator of $J_{\mathsf{QS}\text{-}\mu,\cdot,n}$.
$\hat{J}_{\mathsf{QS}\text{-}\vartheta,\cdot,n}$ = spacings estimator of $J_{\mathsf{QS}\text{-}\vartheta,\cdot,n}$; $\check{J}_{\mathsf{QS}\text{-}\vartheta,\cdot,n}$ = polynomial estimator of $J_{\mathsf{QS}\text{-}\vartheta,\cdot,n}$.

Table SA-9: Simulations Results for Model 8

Panel A: IMSE for Grid of Number of Bins and Estimated Choices

$J_{-,n}$	$\frac{\mathsf{IMSE}_{\mathtt{ES},-}(J_{-,n})}{\mathsf{IMSE}^*_{\mathtt{ES},-}}$	$J_{+,n}$	$\frac{\mathsf{IMSE}_{\mathtt{ES},+}(J_{+,n})}{\mathsf{IMSE}^*_{\mathtt{ES},+}}$	$J_{-,n}$	$\frac{\mathsf{IMSE}_{\mathtt{QS},-}(J_{-,n})}{\mathsf{IMSE}^*_{\mathtt{QS},-}}$	$J_{+,n}$	$\frac{\mathsf{IMSE}_{\mathtt{QS},+}(J_{+,n})}{\mathsf{IMSE}^*_{\mathtt{QS},+}}$
16	1.080	19	1.059	15	1.072	30	1.025
17	1.047	20	1.035	16	1.039	31	1.015
18	1.024	21	1.018	17	1.017	32	1.008
19	1.010	22	1.008	18	1.005	33	1.003
20	1.002	23	1.002	19	1.000	34	1.001
21	1.000	24	1.000	20	1.000	35	1.000
22	1.002	25	1.002	21	1.005	36	1.001
23	1.009	26	1.006	22	1.014	37	1.003
24	1.018	27	1.014	23	1.027	38	1.007
25	1.030	28	1.023	24	1.042	39	1.011
26	1.044	29	1.034	25	1.059	40	1.017
$\hat{J}_{\mathtt{ES}\text{-}\mu,-,n}$	1.039	$\hat{J}_{\mathtt{ES}\text{-}\mu,+,n}$	0.8473	$\hat{J}_{\mathtt{QS}\text{-}\mu,-,n}$	1.019	$\hat{J}_{\mathtt{QS}\text{-}\mu,+,n}$	0.9442
$\check{J}_{\mathtt{ES}\text{-}\mu,-,n}$	1.042	$\check{J}_{\mathtt{ES}\text{-}\mu,+,n}$	0.8474	$\check{J}_{\mathtt{QS}\text{-}\mu,-,n}$	1.021	$\check{J}_{\mathtt{QS}\text{-}\mu,+,n}$	0.9443

Panel B: Summary Statistics for the Estimated Number of Bins

Pop. Par.		Min.	1st Qu.	Median	Mean	3rd Qu.	Max.	Std. Dev.
$J_{\mathtt{ES}\text{-}\mu,-,n} = 21$	$\hat{J}_{\mathtt{ES}\text{-}\mu,-,n}$	18	22	22	22.36	23	26	1.16
	$\check{J}_{\mathtt{ES}\text{-}\mu,-,n}$	18	22	22	22.32	23	26	1.04
$J_{\mathtt{ES}\text{-}\vartheta,-,n} = 150$	$\hat{J}_{\mathtt{ES}\text{-}\vartheta,-,n}$	103	139	147	147.9	156	207	12.86
	$\check{J}_{\mathtt{ES}\text{-}\vartheta,-,n}$	121	142	146	146.7	151.2	175	7.34
$J_{\mathtt{ES}\text{-}\mu,+,n} = 24$	$\hat{J}_{\mathtt{ES}\text{-}\mu,+,n}$	16	20	21	20.85	22	26	1.47
	$\check{J}_{\mathtt{ES}\text{-}\mu,+,n}$	16	20	21	20.84	22	26	1.40
$J_{\mathtt{ES}\text{-}\vartheta,+,n} = 102$	$\hat{J}_{\mathtt{ES}\text{-}\vartheta,+,n}$	68	94	100	100.5	107	140	9.47
	$\check{J}_{\mathtt{ES}\text{-}\vartheta,+,n}$	77	95	100	100.1	105	128	6.87
$J_{\mathtt{QS}\text{-}\mu,-,n} = 20$	$\hat{J}_{\mathtt{QS}\text{-}\mu,-,n}$	17	20	21	20.53	21	24	0.93
	$\check{J}_{\mathtt{QS}\text{-}\mu,-,n}$	17	20	20	20.5	21	24	0.89
$J_{\mathtt{QS}\text{-}\vartheta,-,n} = 151$	$\hat{J}_{\mathtt{QS}\text{-}\vartheta,-,n}$	119	143	149	149	155	191	9.17
	$\check{J}_{\mathtt{QS}\text{-}\vartheta,-,n}$	123	144	148	148.4	153	176	7.31
$J_{\mathtt{QS}\text{-}\mu,+,n} = 35$	$\hat{J}_{\mathtt{QS}\text{-}\mu,+,n}$	28	32	34	33.72	35	43	1.85
	$\check{J}_{\mathtt{QS}\text{-}\mu,+,n}$	29	32	34	33.72	35	43	1.84
$J_{\mathtt{QS}\text{-}\vartheta,+,n} = 142$	$\hat{J}_{\mathtt{QS}\text{-}\vartheta,+,n}$	122	136	139	139.2	142	157	4.79
	$\check{J}_{\mathtt{QS}\text{-}\vartheta,+,n}$	123	136	139	139.2	142	154	4.18

Notes:

(i) Population quantities:
$J_{\mathtt{ES}\text{-}\mu,\cdot,n}$ = IMSE-optimal partition size for ES RD Plot.
$J_{\mathtt{ES}\text{-}\vartheta,\cdot,n}$ = Mimicking variance partition size for ES RD Plot.
$J_{\mathtt{QS}\text{-}\mu,\cdot,n}$ = IMSE-optimal partition size for QS RD Plot.
$J_{\mathtt{QS}\text{-}\vartheta,\cdot,n}$ = Mimicking variance partition size for QS RD Plot.
$\mathsf{IMSE}^*_{\mathtt{ES},\cdot} = \mathsf{IMSE}_{\mathtt{ES},\cdot}(J_{\mathtt{ES}\text{-}\mu,\cdot,n})$ = ES IMSE function evaluated at optimal choice.
$\mathsf{IMSE}^*_{\mathtt{QS},\cdot} = \mathsf{IMSE}_{\mathtt{QS},\cdot}(J_{\mathtt{QS}\text{-}\mu,\cdot,n})$ = QS IMSE function evaluated at optimal choice.

(ii) Estimators:
$\hat{J}_{\mathtt{ES}\text{-}\mu,\cdot,n}$ = spacings estimator of $J_{\mathtt{ES}\text{-}\mu,\cdot,n}$; $\check{J}_{\mathtt{ES}\text{-}\mu,\cdot,n}$ = polynomial estimator of $J_{\mathtt{ES}\text{-}\mu,\cdot,n}$.
$\hat{J}_{\mathtt{ES}\text{-}\vartheta,\cdot,n}$ = spacings estimator of $J_{\mathtt{ES}\text{-}\vartheta,\cdot,n}$; $\check{J}_{\mathtt{ES}\text{-}\vartheta,\cdot,n}$ = polynomial estimator of $J_{\mathtt{ES}\text{-}\vartheta,\cdot,n}$.
$\hat{J}_{\mathtt{QS}\text{-}\mu,\cdot,n}$ = spacings estimator of $J_{\mathtt{QS}\text{-}\mu,\cdot,n}$; $\check{J}_{\mathtt{QS}\text{-}\mu,\cdot,n}$ = polynomial estimator of $J_{\mathtt{QS}\text{-}\mu,\cdot,n}$.
$\hat{J}_{\mathtt{QS}\text{-}\vartheta,\cdot,n}$ = spacings estimator of $J_{\mathtt{QS}\text{-}\vartheta,\cdot,n}$; $\check{J}_{\mathtt{QS}\text{-}\vartheta,\cdot,n}$ = polynomial estimator of $J_{\mathtt{QS}\text{-}\vartheta,\cdot,n}$.

Table SA-10: Simulations Results for Model 9

Panel A: IMSE for Grid of Number of Bins and Estimated Choices

$J_{-,n}$	$\frac{\mathsf{IMSE}_{\mathtt{ES},-}(J_{-,n})}{\mathsf{IMSE}^*_{\mathtt{ES},-}}$	$J_{+,n}$	$\frac{\mathsf{IMSE}_{\mathtt{ES},+}(J_{+,n})}{\mathsf{IMSE}^*_{\mathtt{ES},+}}$	$J_{-,n}$	$\frac{\mathsf{IMSE}_{\mathtt{QS},-}(J_{-,n})}{\mathsf{IMSE}^*_{\mathtt{QS},-}}$	$J_{+,n}$	$\frac{\mathsf{IMSE}_{\mathtt{QS},+}(J_{+,n})}{\mathsf{IMSE}^*_{\mathtt{QS},+}}$
15	1.088	12	1.133	61	1.006	23	1.028
16	1.051	13	1.075	62	1.004	24	1.015
17	1.026	14	1.037	63	1.002	25	1.006
18	1.010	15	1.014	64	1.001	26	1.001
19	1.002	16	1.003	65	1.000	27	0.999
20	1.000	17	1.000	66	1.000	28	1.000
21	1.003	18	1.004	67	1.000	29	1.003
22	1.010	19	1.014	68	1.001	30	1.009
23	1.020	20	1.027	69	1.002	31	1.016
24	1.034	21	1.045	70	1.004	32	1.025
25	1.049	22	1.065	71	1.006	33	1.035
$\hat{J}_{\mathtt{ES}\text{-}\mu,-,n}$	0.9429	$\hat{J}_{\mathtt{ES}\text{-}\mu,+,n}$	0.9666	$\hat{J}_{\mathtt{QS}\text{-}\mu,-,n}$	1.026	$\hat{J}_{\mathtt{QS}\text{-}\mu,+,n}$	0.71
$\check{J}_{\mathtt{ES}\text{-}\mu,-,n}$	0.9447	$\check{J}_{\mathtt{ES}\text{-}\mu,+,n}$	0.9633	$\check{J}_{\mathtt{QS}\text{-}\mu,-,n}$	1.027	$\check{J}_{\mathtt{QS}\text{-}\mu,+,n}$	0.7095

Panel B: Summary Statistics for the Estimated Number of Bins

Pop. Par.		Min.	1st Qu.	Median	Mean	3rd Qu.	Max.	Std. Dev.
$J_{\mathtt{ES}\text{-}\mu,-,n} = 20$	$\hat{J}_{\mathtt{ES}\text{-}\mu,-,n}$	16	19	19	19.22	20	23	0.94
	$\check{J}_{\mathtt{ES}\text{-}\mu,-,n}$	17	19	19	19.19	20	23	0.86
$J_{\mathtt{ES}\text{-}\vartheta,-,n} = 103$	$\hat{J}_{\mathtt{ES}\text{-}\vartheta,-,n}$	71	97	103	103.1	109	132	8.83
	$\check{J}_{\mathtt{ES}\text{-}\vartheta,-,n}$	83	99	102	102.4	106	123	5.77
$J_{\mathtt{ES}\text{-}\mu,+,n} = 17$	$\hat{J}_{\mathtt{ES}\text{-}\mu,+,n}$	14	16	17	16.81	17	20	0.82
	$\check{J}_{\mathtt{ES}\text{-}\mu,+,n}$	15	16	17	16.83	17	20	0.77
$J_{\mathtt{ES}\text{-}\vartheta,+,n} = 96$	$\hat{J}_{\mathtt{ES}\text{-}\vartheta,+,n}$	69	92	96	96.25	101	120	7.06
	$\check{J}_{\mathtt{ES}\text{-}\vartheta,+,n}$	77	93	97	96.48	100	114	4.64
$J_{\mathtt{QS}\text{-}\mu,-,n} = 66$	$\hat{J}_{\mathtt{QS}\text{-}\mu,-,n}$	45	64	68	68.02	72	89	6.29
	$\check{J}_{\mathtt{QS}\text{-}\mu,-,n}$	45	64	68	68.01	72	89	6.26
$J_{\mathtt{QS}\text{-}\vartheta,-,n} = 103$	$\hat{J}_{\mathtt{QS}\text{-}\vartheta,-,n}$	93	101	103	102.7	105	114	3.02
	$\check{J}_{\mathtt{QS}\text{-}\vartheta,-,n}$	95	101	103	102.6	104	112	2.18
$J_{\mathtt{QS}\text{-}\mu,+,n} = 28$	$\hat{J}_{\mathtt{QS}\text{-}\mu,+,n}$	14	18	19	19.77	21	41	3.08
	$\check{J}_{\mathtt{QS}\text{-}\mu,+,n}$	14	18	19	19.77	21	41	3.08
$J_{\mathtt{QS}\text{-}\vartheta,+,n} = 96$	$\hat{J}_{\mathtt{QS}\text{-}\vartheta,+,n}$	86	94	96	95.83	98	107	2.66
	$\check{J}_{\mathtt{QS}\text{-}\vartheta,+,n}$	89	95	96	95.91	97	103	1.86

Notes:

(i) Population quantities:
$J_{\mathtt{ES}\text{-}\mu,\cdot,n}$ = IMSE-optimal partition size for ES RD Plot.
$J_{\mathtt{ES}\text{-}\vartheta,\cdot,n}$ = Mimicking variance partition size for ES RD Plot.
$J_{\mathtt{QS}\text{-}\mu,\cdot,n}$ = IMSE-optimal partition size for QS RD Plot.
$J_{\mathtt{QS}\text{-}\vartheta,\cdot,n}$ = Mimicking variance partition size for QS RD Plot.
$\mathsf{IMSE}^*_{\mathtt{ES},\cdot} = \mathsf{IMSE}_{\mathtt{ES},\cdot}(J_{\mathtt{ES}\text{-}\mu,\cdot,n})$ = ES IMSE function evaluated at optimal choice.
$\mathsf{IMSE}^*_{\mathtt{QS},\cdot} = \mathsf{IMSE}_{\mathtt{QS},\cdot}(J_{\mathtt{QS}\text{-}\mu,\cdot,n})$ = QS IMSE function evaluated at optimal choice.

(ii) Estimators:
$\hat{J}_{\mathtt{ES}\text{-}\mu,\cdot,n}$ = spacings estimator of $J_{\mathtt{ES}\text{-}\mu,\cdot,n}$; $\check{J}_{\mathtt{ES}\text{-}\mu,\cdot,n}$ = polynomial estimator of $J_{\mathtt{ES}\text{-}\mu,\cdot,n}$.
$\hat{J}_{\mathtt{ES}\text{-}\vartheta,\cdot,n}$ = spacings estimator of $J_{\mathtt{ES}\text{-}\vartheta,\cdot,n}$; $\check{J}_{\mathtt{ES}\text{-}\vartheta,\cdot,n}$ = polynomial estimator of $J_{\mathtt{ES}\text{-}\vartheta,\cdot,n}$.
$\hat{J}_{\mathtt{QS}\text{-}\mu,\cdot,n}$ = spacings estimator of $J_{\mathtt{QS}\text{-}\mu,\cdot,n}$; $\check{J}_{\mathtt{QS}\text{-}\mu,\cdot,n}$ = polynomial estimator of $J_{\mathtt{QS}\text{-}\mu,\cdot,n}$.
$\hat{J}_{\mathtt{QS}\text{-}\vartheta,\cdot,n}$ = spacings estimator of $J_{\mathtt{QS}\text{-}\vartheta,\cdot,n}$; $\check{J}_{\mathtt{QS}\text{-}\vartheta,\cdot,n}$ = polynomial estimator of $J_{\mathtt{QS}\text{-}\vartheta,\cdot,n}$.

Table SA-11: Simulations Results for Model 10

Panel A: IMSE for Grid of Number of Bins and Estimated Choices

$J_{-,n}$	$\frac{\mathsf{IMSE}_{\mathsf{ES},-}(J_{-,n})}{\mathsf{IMSE}^*_{\mathsf{ES},-}}$	$J_{+,n}$	$\frac{\mathsf{IMSE}_{\mathsf{ES},+}(J_{+,n})}{\mathsf{IMSE}^*_{\mathsf{ES},+}}$	$J_{-,n}$	$\frac{\mathsf{IMSE}_{\mathsf{QS},-}(J_{-,n})}{\mathsf{IMSE}^*_{\mathsf{QS},-}}$	$J_{+,n}$	$\frac{\mathsf{IMSE}_{\mathsf{QS},+}(J_{+,n})}{\mathsf{IMSE}^*_{\mathsf{QS},+}}$
17	1.064	11	1.129	28	1.018	13	1.110
18	1.036	12	1.068	29	1.009	14	1.062
19	1.018	13	1.030	30	1.003	15	1.030
20	1.006	14	1.008	31	1.000	16	1.011
21	1.001	15	0.999	32	0.999	17	1.002
22	1.000	16	1.000	33	1.000	18	1.000
23	1.003	17	1.008	34	1.003	19	1.004
24	1.010	18	1.021	35	1.007	20	1.013
25	1.020	19	1.039	36	1.012	21	1.027
26	1.032	20	1.061	37	1.019	22	1.043
27	1.046	21	1.086	38	1.027	23	1.062
$\hat{J}_{\mathsf{ES}\text{-}\mu,-,n}$	1.047	$\hat{J}_{\mathsf{ES}\text{-}\mu,+,n}$	0.9967	$\hat{J}_{\mathsf{QS}\text{-}\mu,-,n}$	1.044	$\hat{J}_{\mathsf{QS}\text{-}\mu,+,n}$	0.8817
$\check{J}_{\mathsf{ES}\text{-}\mu,-,n}$	1.049	$\check{J}_{\mathsf{ES}\text{-}\mu,+,n}$	0.995	$\check{J}_{\mathsf{QS}\text{-}\mu,-,n}$	1.045	$\check{J}_{\mathsf{QS}\text{-}\mu,+,n}$	0.8811

Panel B: Summary Statistics for the Estimated Number of Bins

Pop. Par.		Min.	1st Qu.	Median	Mean	3rd Qu.	Max.	Std. Dev.
$J_{\mathsf{ES}\text{-}\mu,-,n} = 22$	$\hat{J}_{\mathsf{ES}\text{-}\mu,-,n}$	20	22	23	22.89	23	26	0.81
	$\check{J}_{\mathsf{ES}\text{-}\mu,-,n}$	20	22	23	22.86	23	26	0.75
$J_{\mathsf{ES}\text{-}\vartheta,-,n} = 121$	$\hat{J}_{\mathsf{ES}\text{-}\vartheta,-,n}$	91	106	110	109.8	113	131	5.35
	$\check{J}_{\mathsf{ES}\text{-}\vartheta,-,n}$	99	107	109	109.3	111	120	2.97
$J_{\mathsf{ES}\text{-}\mu,+,n} = 16$	$\hat{J}_{\mathsf{ES}\text{-}\mu,+,n}$	14	15	16	15.66	16	18	0.54
	$\check{J}_{\mathsf{ES}\text{-}\mu,+,n}$	14	15	16	15.68	16	17	0.51
$J_{\mathsf{ES}\text{-}\vartheta,+,n} = 111$	$\hat{J}_{\mathsf{ES}\text{-}\vartheta,+,n}$	78	94	97	97.45	101	116	4.68
	$\check{J}_{\mathsf{ES}\text{-}\vartheta,+,n}$	89	96	98	97.57	99	107	2.62
$J_{\mathsf{QS}\text{-}\mu,-,n} = 33$	$\hat{J}_{\mathsf{QS}\text{-}\mu,-,n}$	27	32	33	33.45	35	41	1.69
	$\check{J}_{\mathsf{QS}\text{-}\mu,-,n}$	28	32	33	33.43	35	41	1.67
$J_{\mathsf{QS}\text{-}\vartheta,-,n} = 121$	$\hat{J}_{\mathsf{QS}\text{-}\vartheta,-,n}$	97	107	109	109.4	111	121	3.33
	$\check{J}_{\mathsf{QS}\text{-}\vartheta,-,n}$	101	107	109	109.2	111	120	2.46
$J_{\mathsf{QS}\text{-}\mu,+,n} = 18$	$\hat{J}_{\mathsf{QS}\text{-}\mu,+,n}$	13	15	16	15.93	17	22	1.21
	$\check{J}_{\mathsf{QS}\text{-}\mu,+,n}$	13	15	16	15.93	17	22	1.20
$J_{\mathsf{QS}\text{-}\vartheta,+,n} = 111$	$\hat{J}_{\mathsf{QS}\text{-}\vartheta,+,n}$	88	95	97	97.35	99	108	2.82
	$\check{J}_{\mathsf{QS}\text{-}\vartheta,+,n}$	90	96	97	97.4	99	105	1.98

Notes:

(i) Population quantities:
$J_{\mathsf{ES}\text{-}\mu,\cdot,n}$ = IMSE-optimal partition size for ES RD Plot.
$J_{\mathsf{ES}\text{-}\vartheta,\cdot,n}$ = Mimicking variance partition size for ES RD Plot.
$J_{\mathsf{QS}\text{-}\mu,\cdot,n}$ = IMSE-optimal partition size for QS RD Plot.
$J_{\mathsf{QS}\text{-}\vartheta,\cdot,n}$ = Mimicking variance partition size for QS RD Plot.
$\mathsf{IMSE}^*_{\mathsf{ES},\cdot} = \mathsf{IMSE}_{\mathsf{ES},\cdot}(J_{\mathsf{ES}\text{-}\mu,\cdot,n})$ = ES IMSE function evaluated at optimal choice.
$\mathsf{IMSE}^*_{\mathsf{QS},\cdot} = \mathsf{IMSE}_{\mathsf{QS},\cdot}(J_{\mathsf{QS}\text{-}\mu,\cdot,n})$ = QS IMSE function evaluated at optimal choice.

(ii) Estimators:
$\hat{J}_{\mathsf{ES}\text{-}\mu,\cdot,n}$ = spacings estimator of $J_{\mathsf{ES}\text{-}\mu,\cdot,n}$; $\check{J}_{\mathsf{ES}\text{-}\mu,\cdot,n}$ = polynomial estimator of $J_{\mathsf{ES}\text{-}\mu,\cdot,n}$.
$\hat{J}_{\mathsf{ES}\text{-}\vartheta,\cdot,n}$ = spacings estimator of $J_{\mathsf{ES}\text{-}\vartheta,\cdot,n}$; $\check{J}_{\mathsf{ES}\text{-}\vartheta,\cdot,n}$ = polynomial estimator of $J_{\mathsf{ES}\text{-}\vartheta,\cdot,n}$.
$\hat{J}_{\mathsf{QS}\text{-}\mu,\cdot,n}$ = spacings estimator of $J_{\mathsf{QS}\text{-}\mu,\cdot,n}$; $\check{J}_{\mathsf{QS}\text{-}\mu,\cdot,n}$ = polynomial estimator of $J_{\mathsf{QS}\text{-}\mu,\cdot,n}$.
$\hat{J}_{\mathsf{QS}\text{-}\vartheta,\cdot,n}$ = spacings estimator of $J_{\mathsf{QS}\text{-}\vartheta,\cdot,n}$; $\check{J}_{\mathsf{QS}\text{-}\vartheta,\cdot,n}$ = polynomial estimator of $J_{\mathsf{QS}\text{-}\vartheta,\cdot,n}$.

Table SA-12: Simulations Results for Model 11

Panel A: IMSE for Grid of Number of Bins and Estimated Choices

$J_{-,n}$	$\frac{\mathsf{IMSE}_{\mathsf{ES},-}(J_{-,n})}{\mathsf{IMSE}^*_{\mathsf{ES},-}}$	$J_{+,n}$	$\frac{\mathsf{IMSE}_{\mathsf{ES},+}(J_{+,n})}{\mathsf{IMSE}^*_{\mathsf{ES},+}}$	$J_{-,n}$	$\frac{\mathsf{IMSE}_{\mathsf{QS},-}(J_{-,n})}{\mathsf{IMSE}^*_{\mathsf{QS},-}}$	$J_{+,n}$	$\frac{\mathsf{IMSE}_{\mathsf{QS},+}(J_{+,n})}{\mathsf{IMSE}^*_{\mathsf{QS},+}}$
25	1.026	9	1.224	40	1.008	10	1.169
26	1.014	10	1.122	41	1.004	11	1.091
27	1.006	11	1.059	42	1.001	12	1.042
28	1.002	12	1.022	43	1.000	13	1.014
29	1.000	13	1.004	44	0.999	14	1.001
30	1.000	14	1.000	45	1.000	15	1.000
31	1.003	15	1.006	46	1.002	16	1.007
32	1.007	16	1.019	47	1.004	17	1.021
33	1.013	17	1.039	48	1.007	18	1.040
34	1.021	18	1.063	49	1.011	19	1.062
35	1.029	19	1.091	50	1.016	20	1.089
$\hat{J}_{\mathsf{ES}\text{-}\mu,-,n}$	1.036	$\hat{J}_{\mathsf{ES}\text{-}\mu,+,n}$	0.9962	$\hat{J}_{\mathsf{QS}\text{-}\mu,-,n}$	1.083	$\hat{J}_{\mathsf{QS}\text{-}\mu,+,n}$	0.9214
$\check{J}_{\mathsf{ES}\text{-}\mu,-,n}$	1.041	$\check{J}_{\mathsf{ES}\text{-}\mu,+,n}$	0.9944	$\check{J}_{\mathsf{QS}\text{-}\mu,-,n}$	1.085	$\check{J}_{\mathsf{QS}\text{-}\mu,+,n}$	0.9201

Panel B: Summary Statistics for the Estimated Number of Bins

Pop. Par.		Min.	1st Qu.	Median	Mean	3rd Qu.	Max.	Std. Dev.
$J_{\mathsf{ES}\text{-}\mu,-,n} = 30$	$\hat{J}_{\mathsf{ES}\text{-}\mu,-,n}$	28	30	31	30.57	31	33	0.73
	$\check{J}_{\mathsf{ES}\text{-}\mu,-,n}$	29	30	30	30.49	31	32	0.63
$J_{\mathsf{ES}\text{-}\vartheta,-,n} = 150$	$\hat{J}_{\mathsf{ES}\text{-}\vartheta,-,n}$	112	128	132	132	136	155	5.48
	$\check{J}_{\mathsf{ES}\text{-}\vartheta,-,n}$	119	129	131	130.9	133	144	3.28
$J_{\mathsf{ES}\text{-}\mu,+,n} = 14$	$\hat{J}_{\mathsf{ES}\text{-}\mu,+,n}$	12	14	14	14.1	14	17	0.68
	$\check{J}_{\mathsf{ES}\text{-}\mu,+,n}$	12	14	14	14.12	14	16	0.63
$J_{\mathsf{ES}\text{-}\vartheta,+,n} = 147$	$\hat{J}_{\mathsf{ES}\text{-}\vartheta,+,n}$	99	121	127	127	133	165	8.76
	$\check{J}_{\mathsf{ES}\text{-}\vartheta,+,n}$	108	124	127	127	130	148	5.06
$J_{\mathsf{QS}\text{-}\mu,-,n} = 45$	$\hat{J}_{\mathsf{QS}\text{-}\mu,-,n}$	42	46	47	47.28	48	52	1.45
	$\check{J}_{\mathsf{QS}\text{-}\mu,-,n}$	42	46	47	47.22	48	52	1.42
$J_{\mathsf{QS}\text{-}\vartheta,-,n} = 153$	$\hat{J}_{\mathsf{QS}\text{-}\vartheta,-,n}$	120	130	133	132.9	135	146	3.52
	$\check{J}_{\mathsf{QS}\text{-}\vartheta,-,n}$	123	130	132	132.3	134	143	2.72
$J_{\mathsf{QS}\text{-}\mu,+,n} = 15$	$\hat{J}_{\mathsf{QS}\text{-}\mu,+,n}$	11	13	14	13.75	14	18	0.93
	$\check{J}_{\mathsf{QS}\text{-}\mu,+,n}$	11	13	14	13.75	14	18	0.92
$J_{\mathsf{QS}\text{-}\vartheta,+,n} = 144$	$\hat{J}_{\mathsf{QS}\text{-}\vartheta,+,n}$	103	119	123	123.5	127	147	6.05
	$\check{J}_{\mathsf{QS}\text{-}\vartheta,+,n}$	106	120	123	123.7	127	144	4.64

Notes:

(i) Population quantities:
$J_{\mathsf{ES}\text{-}\mu,\cdot,n}$ = IMSE-optimal partition size for ES RD Plot.
$J_{\mathsf{ES}\text{-}\vartheta,\cdot,n}$ = Mimicking variance partition size for ES RD Plot.
$J_{\mathsf{QS}\text{-}\mu,\cdot,n}$ = IMSE-optimal partition size for QS RD Plot.
$J_{\mathsf{QS}\text{-}\vartheta,\cdot,n}$ = Mimicking variance partition size for QS RD Plot.
$\mathsf{IMSE}^*_{\mathsf{ES},\cdot} = \mathsf{IMSE}_{\mathsf{ES},\cdot}(J_{\mathsf{ES}\text{-}\mu,\cdot,n})$ = ES IMSE function evaluated at optimal choice.
$\mathsf{IMSE}^*_{\mathsf{QS},\cdot} = \mathsf{IMSE}_{\mathsf{QS},\cdot}(J_{\mathsf{QS}\text{-}\mu,\cdot,n})$ = QS IMSE function evaluated at optimal choice.

(ii) Estimators:
$\hat{J}_{\mathsf{ES}\text{-}\mu,\cdot,n}$ = spacings estimator of $J_{\mathsf{ES}\text{-}\mu,\cdot,n}$; $\check{J}_{\mathsf{ES}\text{-}\mu,\cdot,n}$ = polynomial estimator of $J_{\mathsf{ES}\text{-}\mu,\cdot,n}$.
$\hat{J}_{\mathsf{ES}\text{-}\vartheta,\cdot,n}$ = spacings estimator of $J_{\mathsf{ES}\text{-}\vartheta,\cdot,n}$; $\check{J}_{\mathsf{ES}\text{-}\vartheta,\cdot,n}$ = polynomial estimator of $J_{\mathsf{ES}\text{-}\vartheta,\cdot,n}$.
$\hat{J}_{\mathsf{QS}\text{-}\mu,\cdot,n}$ = spacings estimator of $J_{\mathsf{QS}\text{-}\mu,\cdot,n}$; $\check{J}_{\mathsf{QS}\text{-}\mu,\cdot,n}$ = polynomial estimator of $J_{\mathsf{QS}\text{-}\mu,\cdot,n}$.
$\hat{J}_{\mathsf{QS}\text{-}\vartheta,\cdot,n}$ = spacings estimator of $J_{\mathsf{QS}\text{-}\vartheta,\cdot,n}$; $\check{J}_{\mathsf{QS}\text{-}\vartheta,\cdot,n}$ = polynomial estimator of $J_{\mathsf{QS}\text{-}\vartheta,\cdot,n}$.

Table SA-13: Simulations Results for Model 12

Panel A: IMSE for Grid of Number of Bins and Estimated Choices

$J_{-,n}$	$\frac{\mathsf{IMSE}_{\mathtt{ES},-}(J_{-,n})}{\mathsf{IMSE}^*_{\mathtt{ES},-}}$	$J_{+,n}$	$\frac{\mathsf{IMSE}_{\mathtt{ES},+}(J_{+,n})}{\mathsf{IMSE}^*_{\mathtt{ES},+}}$	$J_{-,n}$	$\frac{\mathsf{IMSE}_{\mathtt{QS},-}(J_{-,n})}{\mathsf{IMSE}^*_{\mathtt{QS},-}}$	$J_{+,n}$	$\frac{\mathsf{IMSE}_{\mathtt{QS},+}(J_{+,n})}{\mathsf{IMSE}^*_{\mathtt{QS},+}}$
15	1.061	16	1.075	24	1.024	21	1.034
16	1.030	17	1.043	25	1.012	22	1.018
17	1.011	18	1.021	26	1.004	23	1.007
18	1.001	19	1.008	27	1.000	24	1.001
19	0.998	20	1.001	28	0.999	25	0.999
20	1.000	21	1.000	29	1.000	26	1.000
21	1.007	22	1.003	30	1.003	27	1.004
22	1.017	23	1.010	31	1.009	28	1.010
23	1.031	24	1.020	32	1.016	29	1.018
24	1.047	25	1.033	33	1.025	30	1.028
25	1.066	26	1.047	34	1.034	31	1.040
$\hat{J}_{\mathtt{ES}\text{-}\mu,-,n}$	1.014	$\hat{J}_{\mathtt{ES}\text{-}\mu,+,n}$	0.9924	$\hat{J}_{\mathtt{QS}\text{-}\mu,-,n}$	1.097	$\hat{J}_{\mathtt{QS}\text{-}\mu,+,n}$	0.8544
$\check{J}_{\mathtt{ES}\text{-}\mu,-,n}$	1.015	$\check{J}_{\mathtt{ES}\text{-}\mu,+,n}$	0.9926	$\check{J}_{\mathtt{QS}\text{-}\mu,-,n}$	1.098	$\check{J}_{\mathtt{QS}\text{-}\mu,+,n}$	0.8545

Panel B: Summary Statistics for the Estimated Number of Bins

Pop. Par.		Min.	1st Qu.	Median	Mean	3rd Qu.	Max.	Std. Dev.
$J_{\mathtt{ES}\text{-}\mu,-,n} = 20$	$\hat{J}_{\mathtt{ES}\text{-}\mu,-,n}$	17	19	19	19.48	20	23	0.86
	$\check{J}_{\mathtt{ES}\text{-}\mu,-,n}$	17	19	19	19.46	20	22	0.77
$J_{\mathtt{ES}\text{-}\vartheta,-,n} = 157$	$\hat{J}_{\mathtt{ES}\text{-}\vartheta,-,n}$	108	136	143	143.3	150	189	10.43
	$\check{J}_{\mathtt{ES}\text{-}\vartheta,-,n}$	124	138	142	142.5	147	164	5.94
$J_{\mathtt{ES}\text{-}\mu,+,n} = 21$	$\hat{J}_{\mathtt{ES}\text{-}\mu,+,n}$	19	20	21	20.81	21	22	0.54
	$\check{J}_{\mathtt{ES}\text{-}\mu,+,n}$	20	21	21	20.81	21	22	0.47
$J_{\mathtt{ES}\text{-}\vartheta,+,n} = 134$	$\hat{J}_{\mathtt{ES}\text{-}\vartheta,+,n}$	94	108	111	111	114	130	4.85
	$\check{J}_{\mathtt{ES}\text{-}\vartheta,+,n}$	100	109	111	110.8	113	120	2.80
$J_{\mathtt{QS}\text{-}\mu,-,n} = 29$	$\hat{J}_{\mathtt{QS}\text{-}\mu,-,n}$	25	30	31	30.67	32	37	1.77
	$\check{J}_{\mathtt{QS}\text{-}\mu,-,n}$	25	30	31	30.65	32	37	1.73
$J_{\mathtt{QS}\text{-}\vartheta,-,n} = 153$	$\hat{J}_{\mathtt{QS}\text{-}\vartheta,-,n}$	118	135	140	139.9	144	169	7.19
	$\check{J}_{\mathtt{QS}\text{-}\vartheta,-,n}$	122	136	139	139.7	143	160	5.49
$J_{\mathtt{QS}\text{-}\mu,+,n} = 26$	$\hat{J}_{\mathtt{QS}\text{-}\mu,+,n}$	18	21	22	21.8	23	29	1.68
	$\check{J}_{\mathtt{QS}\text{-}\mu,+,n}$	18	21	22	21.8	23	29	1.66
$J_{\mathtt{QS}\text{-}\vartheta,+,n} = 135$	$\hat{J}_{\mathtt{QS}\text{-}\vartheta,+,n}$	103	111	113	113	115	125	2.90
	$\check{J}_{\mathtt{QS}\text{-}\vartheta,+,n}$	106	111	113	113	114	122	2.19

Notes:

(i) Population quantities:
$J_{\mathtt{ES}\text{-}\mu,\cdot,n}$ = IMSE-optimal partition size for ES RD Plot.
$J_{\mathtt{ES}\text{-}\vartheta,\cdot,n}$ = Mimicking variance partition size for ES RD Plot.
$J_{\mathtt{QS}\text{-}\mu,\cdot,n}$ = IMSE-optimal partition size for QS RD Plot.
$J_{\mathtt{QS}\text{-}\vartheta,\cdot,n}$ = Mimicking variance partition size for QS RD Plot.
$\mathsf{IMSE}^*_{\mathtt{ES},\cdot} = \mathsf{IMSE}_{\mathtt{ES},\cdot}(J_{\mathtt{ES}\text{-}\mu,\cdot,n})$ = ES IMSE function evaluated at optimal choice.
$\mathsf{IMSE}^*_{\mathtt{QS},\cdot} = \mathsf{IMSE}_{\mathtt{QS},\cdot}(J_{\mathtt{QS}\text{-}\mu,\cdot,n})$ = QS IMSE function evaluated at optimal choice.

(ii) Estimators:
$\hat{J}_{\mathtt{ES}\text{-}\mu,\cdot,n}$ = spacings estimator of $J_{\mathtt{ES}\text{-}\mu,\cdot,n}$; $\check{J}_{\mathtt{ES}\text{-}\mu,\cdot,n}$ = polynomial estimator of $J_{\mathtt{ES}\text{-}\mu,\cdot,n}$.
$\hat{J}_{\mathtt{ES}\text{-}\vartheta,\cdot,n}$ = spacings estimator of $J_{\mathtt{ES}\text{-}\vartheta,\cdot,n}$; $\check{J}_{\mathtt{ES}\text{-}\vartheta,\cdot,n}$ = polynomial estimator of $J_{\mathtt{ES}\text{-}\vartheta,\cdot,n}$.
$\hat{J}_{\mathtt{QS}\text{-}\mu,\cdot,n}$ = spacings estimator of $J_{\mathtt{QS}\text{-}\mu,\cdot,n}$; $\check{J}_{\mathtt{QS}\text{-}\mu,\cdot,n}$ = polynomial estimator of $J_{\mathtt{QS}\text{-}\mu,\cdot,n}$.
$\hat{J}_{\mathtt{QS}\text{-}\vartheta,\cdot,n}$ = spacings estimator of $J_{\mathtt{QS}\text{-}\vartheta,\cdot,n}$; $\check{J}_{\mathtt{QS}\text{-}\vartheta,\cdot,n}$ = polynomial estimator of $J_{\mathtt{QS}\text{-}\vartheta,\cdot,n}$.

Table SA-14: Simulations Results for Model 13

Panel A: IMSE for Grid of Number of Bins and Estimated Choices

$J_{-,n}$	$\frac{\mathsf{IMSE}_{\mathtt{ES},-}(J_{-,n})}{\mathsf{IMSE}^*_{\mathtt{ES},-}}$	$J_{+,n}$	$\frac{\mathsf{IMSE}_{\mathtt{ES},+}(J_{+,n})}{\mathsf{IMSE}^*_{\mathtt{ES},+}}$	$J_{-,n}$	$\frac{\mathsf{IMSE}_{\mathtt{QS},-}(J_{-,n})}{\mathsf{IMSE}^*_{\mathtt{QS},-}}$	$J_{+,n}$	$\frac{\mathsf{IMSE}_{\mathtt{QS},+}(J_{+,n})}{\mathsf{IMSE}^*_{\mathtt{QS},+}}$
15	1.088	12	1.133	61	1.006	23	1.028
16	1.051	13	1.075	62	1.003	24	1.014
17	1.026	14	1.037	63	1.002	25	1.006
18	1.010	15	1.014	64	1.001	26	1.001
19	1.002	16	1.003	65	1.000	27	0.999
20	1.000	17	1.000	66	1.000	28	1.000
21	1.003	18	1.004	67	1.000	29	1.003
22	1.010	19	1.014	68	1.001	30	1.009
23	1.020	20	1.028	69	1.002	31	1.016
24	1.034	21	1.045	70	1.004	32	1.025
25	1.049	22	1.066	71	1.006	33	1.035
$\hat{J}_{\mathtt{ES}\text{-}\mu,-,n}$	0.95	$\hat{J}_{\mathtt{ES}\text{-}\mu,+,n}$	0.9652	$\hat{J}_{\mathtt{QS}\text{-}\mu,-,n}$	1.092	$\hat{J}_{\mathtt{QS}\text{-}\mu,+,n}$	0.8257
$\check{J}_{\mathtt{ES}\text{-}\mu,-,n}$	0.9532	$\check{J}_{\mathtt{ES}\text{-}\mu,+,n}$	0.9578	$\check{J}_{\mathtt{QS}\text{-}\mu,-,n}$	1.093	$\check{J}_{\mathtt{QS}\text{-}\mu,+,n}$	0.8247

Panel B: Summary Statistics for the Estimated Number of Bins

Pop. Par.		Min.	1st Qu.	Median	Mean	3rd Qu.	Max.	Std. Dev.
$J_{\mathtt{ES}\text{-}\mu,-,n} = 20$	$\hat{J}_{\mathtt{ES}\text{-}\mu,-,n}$	15	19	19	19.4	20	24	1.02
	$\check{J}_{\mathtt{ES}\text{-}\mu,-,n}$	16	19	19	19.36	20	23	0.92
$J_{\mathtt{ES}\text{-}\vartheta,-,n} = 104$	$\hat{J}_{\mathtt{ES}\text{-}\vartheta,-,n}$	55	97	106	104.8	113	143	11.59
	$\check{J}_{\mathtt{ES}\text{-}\vartheta,-,n}$	64	98	104	103.9	110	135	8.93
$J_{\mathtt{ES}\text{-}\mu,+,n} = 17$	$\hat{J}_{\mathtt{ES}\text{-}\mu,+,n}$	13	16	17	16.84	17	21	1.00
	$\check{J}_{\mathtt{ES}\text{-}\mu,+,n}$	14	16	17	16.87	17	20	0.88
$J_{\mathtt{ES}\text{-}\vartheta,+,n} = 96$	$\hat{J}_{\mathtt{ES}\text{-}\vartheta,+,n}$	46	88	96	94.64	102	126	10.72
	$\check{J}_{\mathtt{ES}\text{-}\vartheta,+,n}$	57	90	96	95.13	100	117	7.54
$J_{\mathtt{QS}\text{-}\mu,-,n} = 66$	$\hat{J}_{\mathtt{QS}\text{-}\mu,-,n}$	49	68	72	72.34	77	104	6.56
	$\check{J}_{\mathtt{QS}\text{-}\mu,-,n}$	49	68	72	72.33	77	105	6.53
$J_{\mathtt{QS}\text{-}\vartheta,-,n} = 104$	$\hat{J}_{\mathtt{QS}\text{-}\vartheta,-,n}$	93	102	104	103.8	106	118	3.22
	$\check{J}_{\mathtt{QS}\text{-}\vartheta,-,n}$	96	102	104	103.7	105	114	2.47
$J_{\mathtt{QS}\text{-}\mu,+,n} = 28$	$\hat{J}_{\mathtt{QS}\text{-}\mu,+,n}$	13	19	22	22.89	25	51	4.86
	$\check{J}_{\mathtt{QS}\text{-}\mu,+,n}$	13	19	22	22.9	25	51	4.86
$J_{\mathtt{QS}\text{-}\vartheta,+,n} = 96$	$\hat{J}_{\mathtt{QS}\text{-}\vartheta,+,n}$	85	92	93	93.37	95	103	2.67
	$\check{J}_{\mathtt{QS}\text{-}\vartheta,+,n}$	88	92	93	93.49	95	102	1.96

Notes:

(i) Population quantities:
$J_{\mathtt{ES}\text{-}\mu,\cdot,n}$ = IMSE-optimal partition size for ES RD Plot.
$J_{\mathtt{ES}\text{-}\vartheta,\cdot,n}$ = Mimicking variance partition size for ES RD Plot.
$J_{\mathtt{QS}\text{-}\mu,\cdot,n}$ = IMSE-optimal partition size for QS RD Plot.
$J_{\mathtt{QS}\text{-}\vartheta,\cdot,n}$ = Mimicking variance partition size for QS RD Plot.
$\mathsf{IMSE}^*_{\mathtt{ES},\cdot} = \mathsf{IMSE}_{\mathtt{ES},\cdot}(J_{\mathtt{ES}\text{-}\mu,\cdot,n})$ = ES IMSE function evaluated at optimal choice.
$\mathsf{IMSE}^*_{\mathtt{QS},\cdot} = \mathsf{IMSE}_{\mathtt{QS},\cdot}(J_{\mathtt{QS}\text{-}\mu,\cdot,n})$ = QS IMSE function evaluated at optimal choice.

(ii) Estimators:
$\hat{J}_{\mathtt{ES}\text{-}\mu,\cdot,n}$ = spacings estimator of $J_{\mathtt{ES}\text{-}\mu,\cdot,n}$; $\check{J}_{\mathtt{ES}\text{-}\mu,\cdot,n}$ = polynomial estimator of $J_{\mathtt{ES}\text{-}\mu,\cdot,n}$.
$\hat{J}_{\mathtt{ES}\text{-}\vartheta,\cdot,n}$ = spacings estimator of $J_{\mathtt{ES}\text{-}\vartheta,\cdot,n}$; $\check{J}_{\mathtt{ES}\text{-}\vartheta,\cdot,n}$ = polynomial estimator of $J_{\mathtt{ES}\text{-}\vartheta,\cdot,n}$.
$\hat{J}_{\mathtt{QS}\text{-}\mu,\cdot,n}$ = spacings estimator of $J_{\mathtt{QS}\text{-}\mu,\cdot,n}$; $\check{J}_{\mathtt{QS}\text{-}\mu,\cdot,n}$ = polynomial estimator of $J_{\mathtt{QS}\text{-}\mu,\cdot,n}$.
$\hat{J}_{\mathtt{QS}\text{-}\vartheta,\cdot,n}$ = spacings estimator of $J_{\mathtt{QS}\text{-}\vartheta,\cdot,n}$; $\check{J}_{\mathtt{QS}\text{-}\vartheta,\cdot,n}$ = polynomial estimator of $J_{\mathtt{QS}\text{-}\vartheta,\cdot,n}$.

Table SA-15: Simulations Results for Model 14

Panel A: IMSE for Grid of Number of Bins and Estimated Choices

$J_{-,n}$	$\frac{\mathsf{IMSE}_{\mathsf{ES},-}(J_{-,n})}{\mathsf{IMSE}^*_{\mathsf{ES},-}}$	$J_{+,n}$	$\frac{\mathsf{IMSE}_{\mathsf{ES},+}(J_{+,n})}{\mathsf{IMSE}^*_{\mathsf{ES},+}}$	$J_{-,n}$	$\frac{\mathsf{IMSE}_{\mathsf{QS},-}(J_{-,n})}{\mathsf{IMSE}^*_{\mathsf{QS},-}}$	$J_{+,n}$	$\frac{\mathsf{IMSE}_{\mathsf{QS},+}(J_{+,n})}{\mathsf{IMSE}^*_{\mathsf{QS},+}}$
17	1.064	11	1.129	28	1.018	13	1.110
18	1.036	12	1.068	29	1.009	14	1.061
19	1.017	13	1.030	30	1.003	15	1.030
20	1.006	14	1.008	31	1.000	16	1.011
21	1.001	15	0.999	32	0.999	17	1.002
22	1.000	16	1.000	33	1.000	18	1.000
23	1.003	17	1.008	34	1.003	19	1.004
24	1.010	18	1.021	35	1.007	20	1.014
25	1.020	19	1.040	36	1.012	21	1.027
26	1.032	20	1.061	37	1.019	22	1.043
27	1.046	21	1.086	38	1.027	23	1.062
$\hat{J}_{\mathsf{ES}\text{-}\mu,-,n}$	1.059	$\hat{J}_{\mathsf{ES}\text{-}\mu,+,n}$	0.9816	$\hat{J}_{\mathsf{QS}\text{-}\mu,-,n}$	1.172	$\hat{J}_{\mathsf{QS}\text{-}\mu,+,n}$	0.8606
$\check{J}_{\mathsf{ES}\text{-}\mu,-,n}$	1.061	$\check{J}_{\mathsf{ES}\text{-}\mu,+,n}$	0.98	$\check{J}_{\mathsf{QS}\text{-}\mu,-,n}$	1.173	$\check{J}_{\mathsf{QS}\text{-}\mu,+,n}$	0.8602

Panel B: Summary Statistics for the Estimated Number of Bins

Pop. Par.		Min.	1st Qu.	Median	Mean	3rd Qu.	Max.	Std. Dev.
$J_{\mathsf{ES}\text{-}\mu,-,n} = 22$	$\hat{J}_{\mathsf{ES}\text{-}\mu,-,n}$	20	23	23	23.06	24	26	0.86
	$\check{J}_{\mathsf{ES}\text{-}\mu,-,n}$	20	23	23	23.03	24	26	0.77
$J_{\mathsf{ES}\text{-}\vartheta,-,n} = 121$	$\hat{J}_{\mathsf{ES}\text{-}\vartheta,-,n}$	77	105	109	109	113	131	6.57
	$\check{J}_{\mathsf{ES}\text{-}\vartheta,-,n}$	92	106	108	108.5	111	125	3.58
$J_{\mathsf{ES}\text{-}\mu,+,n} = 16$	$\hat{J}_{\mathsf{ES}\text{-}\mu,+,n}$	14	15	15	15.43	16	17	0.59
	$\check{J}_{\mathsf{ES}\text{-}\mu,+,n}$	14	15	15	15.43	16	17	0.54
$J_{\mathsf{ES}\text{-}\vartheta,+,n} = 111$	$\hat{J}_{\mathsf{ES}\text{-}\vartheta,+,n}$	75	95	99	98.67	102	119	5.67
	$\check{J}_{\mathsf{ES}\text{-}\vartheta,+,n}$	85	97	99	98.73	101	110	3.31
$J_{\mathsf{QS}\text{-}\mu,-,n} = 33$	$\hat{J}_{\mathsf{QS}\text{-}\mu,-,n}$	30	36	37	37.42	39	45	1.94
	$\check{J}_{\mathsf{QS}\text{-}\mu,-,n}$	30	36	37	37.4	39	44	1.92
$J_{\mathsf{QS}\text{-}\vartheta,-,n} = 121$	$\hat{J}_{\mathsf{QS}\text{-}\vartheta,-,n}$	97	106	109	108.8	111	121	3.57
	$\check{J}_{\mathsf{QS}\text{-}\vartheta,-,n}$	98	107	109	108.6	110	119	2.77
$J_{\mathsf{QS}\text{-}\mu,+,n} = 18$	$\hat{J}_{\mathsf{QS}\text{-}\mu,+,n}$	13	15	15	15.56	16	21	1.14
	$\check{J}_{\mathsf{QS}\text{-}\mu,+,n}$	13	15	15	15.56	16	21	1.13
$J_{\mathsf{QS}\text{-}\vartheta,+,n} = 111$	$\hat{J}_{\mathsf{QS}\text{-}\vartheta,+,n}$	88	96	98	98.56	101	111	3.05
	$\check{J}_{\mathsf{QS}\text{-}\vartheta,+,n}$	91	97	98	98.59	100	108	2.30

Notes:

(i) Population quantities:
$J_{\mathsf{ES}\text{-}\mu,\cdot,n}$ = IMSE-optimal partition size for ES RD Plot.
$J_{\mathsf{ES}\text{-}\vartheta,\cdot,n}$ = Mimicking variance partition size for ES RD Plot.
$J_{\mathsf{QS}\text{-}\mu,\cdot,n}$ = IMSE-optimal partition size for QS RD Plot.
$J_{\mathsf{QS}\text{-}\vartheta,\cdot,n}$ = Mimicking variance partition size for QS RD Plot.
$\mathsf{IMSE}^*_{\mathsf{ES},\cdot} = \mathsf{IMSE}_{\mathsf{ES},\cdot}(J_{\mathsf{ES}\text{-}\mu,\cdot,n})$ = ES IMSE function evaluated at optimal choice.
$\mathsf{IMSE}^*_{\mathsf{QS},\cdot} = \mathsf{IMSE}_{\mathsf{QS},\cdot}(J_{\mathsf{QS}\text{-}\mu,\cdot,n})$ = QS IMSE function evaluated at optimal choice.

(ii) Estimators:
$\hat{J}_{\mathsf{ES}\text{-}\mu,\cdot,n}$ = spacings estimator of $J_{\mathsf{ES}\text{-}\mu,\cdot,n}$; $\check{J}_{\mathsf{ES}\text{-}\mu,\cdot,n}$ = polynomial estimator of $J_{\mathsf{ES}\text{-}\mu,\cdot,n}$.
$\hat{J}_{\mathsf{ES}\text{-}\vartheta,\cdot,n}$ = spacings estimator of $J_{\mathsf{ES}\text{-}\vartheta,\cdot,n}$; $\check{J}_{\mathsf{ES}\text{-}\vartheta,\cdot,n}$ = polynomial estimator of $J_{\mathsf{ES}\text{-}\vartheta,\cdot,n}$.
$\hat{J}_{\mathsf{QS}\text{-}\mu,\cdot,n}$ = spacings estimator of $J_{\mathsf{QS}\text{-}\mu,\cdot,n}$; $\check{J}_{\mathsf{QS}\text{-}\mu,\cdot,n}$ = polynomial estimator of $J_{\mathsf{QS}\text{-}\mu,\cdot,n}$.
$\hat{J}_{\mathsf{QS}\text{-}\vartheta,\cdot,n}$ = spacings estimator of $J_{\mathsf{QS}\text{-}\vartheta,\cdot,n}$; $\check{J}_{\mathsf{QS}\text{-}\vartheta,\cdot,n}$ = polynomial estimator of $J_{\mathsf{QS}\text{-}\vartheta,\cdot,n}$.

Table SA-16: Simulations Results for Model 15

Panel A: IMSE for Grid of Number of Bins and Estimated Choices

$J_{-,n}$	$\frac{\mathsf{IMSE}_{\mathtt{ES},-}(J_{-,n})}{\mathsf{IMSE}^*_{\mathtt{ES},-}}$	$J_{+,n}$	$\frac{\mathsf{IMSE}_{\mathtt{ES},+}(J_{+,n})}{\mathsf{IMSE}^*_{\mathtt{ES},+}}$	$J_{-,n}$	$\frac{\mathsf{IMSE}_{\mathtt{QS},-}(J_{-,n})}{\mathsf{IMSE}^*_{\mathtt{QS},-}}$	$J_{+,n}$	$\frac{\mathsf{IMSE}_{\mathtt{QS},+}(J_{+,n})}{\mathsf{IMSE}^*_{\mathtt{QS},+}}$
25	1.026	9	1.223	40	1.008	10	1.168
26	1.014	10	1.121	41	1.004	11	1.090
27	1.006	11	1.058	42	1.001	12	1.041
28	1.001	12	1.022	43	1.000	13	1.014
29	0.999	13	1.004	44	0.999	14	1.001
30	1.000	14	1.000	45	1.000	15	1.000
31	1.003	15	1.006	46	1.002	16	1.007
32	1.007	16	1.020	47	1.004	17	1.021
33	1.013	17	1.039	48	1.007	18	1.040
34	1.021	18	1.063	49	1.011	19	1.063
35	1.030	19	1.091	50	1.016	20	1.089
$\hat{J}_{\mathtt{ES}\text{-}\mu,-,n}$	1.041	$\hat{J}_{\mathtt{ES}\text{-}\mu,+,n}$	0.9839	$\hat{J}_{\mathtt{QS}\text{-}\mu,-,n}$	1.158	$\hat{J}_{\mathtt{QS}\text{-}\mu,+,n}$	0.9187
$\check{J}_{\mathtt{ES}\text{-}\mu,-,n}$	1.046	$\check{J}_{\mathtt{ES}\text{-}\mu,+,n}$	0.9816	$\check{J}_{\mathtt{QS}\text{-}\mu,-,n}$	1.161	$\check{J}_{\mathtt{QS}\text{-}\mu,+,n}$	0.9176

Panel B: Summary Statistics for the Estimated Number of Bins

Pop. Par.		Min.	1st Qu.	Median	Mean	3rd Qu.	Max.	Std. Dev.
$J_{\mathtt{ES}\text{-}\mu,-,n} = 30$	$\hat{J}_{\mathtt{ES}\text{-}\mu,-,n}$	28	30	31	30.51	31	33	0.78
	$\check{J}_{\mathtt{ES}\text{-}\mu,-,n}$	28	30	30	30.42	31	33	0.67
$J_{\mathtt{ES}\text{-}\vartheta,-,n} = 149$	$\hat{J}_{\mathtt{ES}\text{-}\vartheta,-,n}$	102	126	130	130	134	158	6.56
	$\check{J}_{\mathtt{ES}\text{-}\vartheta,-,n}$	111	127	129	129	132	141	3.77
$J_{\mathtt{ES}\text{-}\mu,+,n} = 14$	$\hat{J}_{\mathtt{ES}\text{-}\mu,+,n}$	12	13	14	13.93	14	17	0.75
	$\check{J}_{\mathtt{ES}\text{-}\mu,+,n}$	12	14	14	13.94	14	16	0.68
$J_{\mathtt{ES}\text{-}\vartheta,+,n} = 140$	$\hat{J}_{\mathtt{ES}\text{-}\vartheta,+,n}$	87	117	124	124.4	131	168	11.00
	$\check{J}_{\mathtt{ES}\text{-}\vartheta,+,n}$	101	119	124	124.3	129	155	6.83
$J_{\mathtt{QS}\text{-}\mu,-,n} = 45$	$\hat{J}_{\mathtt{QS}\text{-}\mu,-,n}$	44	49	50	50.34	51	56	1.63
	$\check{J}_{\mathtt{QS}\text{-}\mu,-,n}$	44	49	50	50.28	51	56	1.59
$J_{\mathtt{QS}\text{-}\vartheta,-,n} = 151$	$\hat{J}_{\mathtt{QS}\text{-}\vartheta,-,n}$	120	129	131	131.3	134	144	3.61
	$\check{J}_{\mathtt{QS}\text{-}\vartheta,-,n}$	120	129	131	130.8	133	142	2.84
$J_{\mathtt{QS}\text{-}\mu,+,n} = 15$	$\hat{J}_{\mathtt{QS}\text{-}\mu,+,n}$	11	13	14	13.65	14	19	1.10
	$\check{J}_{\mathtt{QS}\text{-}\mu,+,n}$	11	13	14	13.66	14	19	1.11
$J_{\mathtt{QS}\text{-}\vartheta,+,n} = 137$	$\hat{J}_{\mathtt{QS}\text{-}\vartheta,+,n}$	98	115	120	120.4	125	152	7.05
	$\check{J}_{\mathtt{QS}\text{-}\vartheta,+,n}$	103	116	120	120.5	124	143	5.95

Notes:

(i) Population quantities:
$J_{\mathtt{ES}\text{-}\mu,\cdot,n}$ = IMSE-optimal partition size for ES RD Plot.
$J_{\mathtt{ES}\text{-}\vartheta,\cdot,n}$ = Mimicking variance partition size for ES RD Plot.
$J_{\mathtt{QS}\text{-}\mu,\cdot,n}$ = IMSE-optimal partition size for QS RD Plot.
$J_{\mathtt{QS}\text{-}\vartheta,\cdot,n}$ = Mimicking variance partition size for QS RD Plot.
$\mathsf{IMSE}^*_{\mathtt{ES},\cdot} = \mathsf{IMSE}_{\mathtt{ES},\cdot}(J_{\mathtt{ES}\text{-}\mu,\cdot,n})$ = ES IMSE function evaluated at optimal choice.
$\mathsf{IMSE}^*_{\mathtt{QS},\cdot} = \mathsf{IMSE}_{\mathtt{QS},\cdot}(J_{\mathtt{QS}\text{-}\mu,\cdot,n})$ = QS IMSE function evaluated at optimal choice.

(ii) Estimators:
$\hat{J}_{\mathtt{ES}\text{-}\mu,\cdot,n}$ = spacings estimator of $J_{\mathtt{ES}\text{-}\mu,\cdot,n}$; $\check{J}_{\mathtt{ES}\text{-}\mu,\cdot,n}$ = polynomial estimator of $J_{\mathtt{ES}\text{-}\mu,\cdot,n}$.
$\hat{J}_{\mathtt{ES}\text{-}\vartheta,\cdot,n}$ = spacings estimator of $J_{\mathtt{ES}\text{-}\vartheta,\cdot,n}$; $\check{J}_{\mathtt{ES}\text{-}\vartheta,\cdot,n}$ = polynomial estimator of $J_{\mathtt{ES}\text{-}\vartheta,\cdot,n}$.
$\hat{J}_{\mathtt{QS}\text{-}\mu,\cdot,n}$ = spacings estimator of $J_{\mathtt{QS}\text{-}\mu,\cdot,n}$; $\check{J}_{\mathtt{QS}\text{-}\mu,\cdot,n}$ = polynomial estimator of $J_{\mathtt{QS}\text{-}\mu,\cdot,n}$.
$\hat{J}_{\mathtt{QS}\text{-}\vartheta,\cdot,n}$ = spacings estimator of $J_{\mathtt{QS}\text{-}\vartheta,\cdot,n}$; $\check{J}_{\mathtt{QS}\text{-}\vartheta,\cdot,n}$ = polynomial estimator of $J_{\mathtt{QS}\text{-}\vartheta,\cdot,n}$.

Table SA-17: Simulations Results for Model 16

Panel A: IMSE for Grid of Number of Bins and Estimated Choices

$J_{-,n}$	$\frac{\mathsf{IMSE}_{\mathsf{ES},-}(J_{-,n})}{\mathsf{IMSE}^*_{\mathsf{ES},-}}$	$J_{+,n}$	$\frac{\mathsf{IMSE}_{\mathsf{ES},+}(J_{+,n})}{\mathsf{IMSE}^*_{\mathsf{ES},+}}$	$J_{-,n}$	$\frac{\mathsf{IMSE}_{\mathsf{QS},-}(J_{-,n})}{\mathsf{IMSE}^*_{\mathsf{QS},-}}$	$J_{+,n}$	$\frac{\mathsf{IMSE}_{\mathsf{QS},+}(J_{+,n})}{\mathsf{IMSE}^*_{\mathsf{QS},+}}$
15	1.059	16	1.073	24	1.023	21	1.033
16	1.030	17	1.042	25	1.011	22	1.017
17	1.011	18	1.021	26	1.004	23	1.007
18	1.001	19	1.007	27	1.000	24	1.001
19	0.998	20	1.001	28	0.999	25	0.999
20	1.000	21	1.000	29	1.000	26	1.000
21	1.007	22	1.003	30	1.004	27	1.004
22	1.018	23	1.010	31	1.009	28	1.010
23	1.031	24	1.021	32	1.016	29	1.019
24	1.048	25	1.033	33	1.025	30	1.029
25	1.066	26	1.048	34	1.035	31	1.041
$\hat{J}_{\mathsf{ES}\text{-}\mu,-,n}$	1.048	$\hat{J}_{\mathsf{ES}\text{-}\mu,+,n}$	0.9941	$\hat{J}_{\mathsf{QS}\text{-}\mu,-,n}$	1.071	$\hat{J}_{\mathsf{QS}\text{-}\mu,+,n}$	0.8365
$\check{J}_{\mathsf{ES}\text{-}\mu,-,n}$	1.05	$\check{J}_{\mathsf{ES}\text{-}\mu,+,n}$	0.9938	$\check{J}_{\mathsf{QS}\text{-}\mu,-,n}$	1.072	$\check{J}_{\mathsf{QS}\text{-}\mu,+,n}$	0.8365

Panel B: Summary Statistics for the Estimated Number of Bins

Pop. Par.		Min.	1st Qu.	Median	Mean	3rd Qu.	Max.	Std. Dev.
$J_{\mathsf{ES}\text{-}\mu,-,n} = 20$	$\hat{J}_{\mathsf{ES}\text{-}\mu,-,n}$	17	20	20	20.09	21	24	0.92
	$\check{J}_{\mathsf{ES}\text{-}\mu,-,n}$	17	20	20	20.05	21	23	0.80
$J_{\mathsf{ES}\text{-}\vartheta,-,n} = 155$	$\hat{J}_{\mathsf{ES}\text{-}\vartheta,-,n}$	94	132	139	138.9	146	179	11.15
	$\check{J}_{\mathsf{ES}\text{-}\vartheta,-,n}$	116	134	138	138	142	164	6.18
$J_{\mathsf{ES}\text{-}\mu,+,n} = 21$	$\hat{J}_{\mathsf{ES}\text{-}\mu,+,n}$	19	20	21	20.74	21	23	0.66
	$\check{J}_{\mathsf{ES}\text{-}\mu,+,n}$	19	20	21	20.74	21	23	0.58
$J_{\mathsf{ES}\text{-}\vartheta,+,n} = 134$	$\hat{J}_{\mathsf{ES}\text{-}\vartheta,+,n}$	85	108	112	112	116	132	6.23
	$\check{J}_{\mathsf{ES}\text{-}\vartheta,+,n}$	98	109	112	111.9	114.2	127	4.09
$J_{\mathsf{QS}\text{-}\mu,-,n} = 29$	$\hat{J}_{\mathsf{QS}\text{-}\mu,-,n}$	24	29	30	30.13	31	37	1.73
	$\check{J}_{\mathsf{QS}\text{-}\mu,-,n}$	24	29	30	30.11	31	37	1.70
$J_{\mathsf{QS}\text{-}\vartheta,-,n} = 151$	$\hat{J}_{\mathsf{QS}\text{-}\vartheta,-,n}$	113	132	136	136.6	141	163	7.10
	$\check{J}_{\mathsf{QS}\text{-}\vartheta,-,n}$	117	132	136	136.3	140	161	5.74
$J_{\mathsf{QS}\text{-}\mu,+,n} = 26$	$\hat{J}_{\mathsf{QS}\text{-}\mu,+,n}$	17	20	21	21.08	22	28	1.42
	$\check{J}_{\mathsf{QS}\text{-}\mu,+,n}$	17	20	21	21.07	22	28	1.41
$J_{\mathsf{QS}\text{-}\vartheta,+,n} = 136$	$\hat{J}_{\mathsf{QS}\text{-}\vartheta,+,n}$	102	111	113	113	115	125	3.35
	$\check{J}_{\mathsf{QS}\text{-}\vartheta,+,n}$	104	111	113	113	115	125	2.74

Notes:

(i) Population quantities:
$J_{\mathsf{ES}\text{-}\mu,\cdot,n}$ = IMSE-optimal partition size for ES RD Plot.
$J_{\mathsf{ES}\text{-}\vartheta,\cdot,n}$ = Mimicking variance partition size for ES RD Plot.
$J_{\mathsf{QS}\text{-}\mu,\cdot,n}$ = IMSE-optimal partition size for QS RD Plot.
$J_{\mathsf{QS}\text{-}\vartheta,\cdot,n}$ = Mimicking variance partition size for QS RD Plot.
$\mathsf{IMSE}^*_{\mathsf{ES},\cdot} = \mathsf{IMSE}_{\mathsf{ES},\cdot}(J_{\mathsf{ES}\text{-}\mu,\cdot,n})$ = ES IMSE function evaluated at optimal choice.
$\mathsf{IMSE}^*_{\mathsf{QS},\cdot} = \mathsf{IMSE}_{\mathsf{QS},\cdot}(J_{\mathsf{QS}\text{-}\mu,\cdot,n})$ = QS IMSE function evaluated at optimal choice.

(ii) Estimators:
$\hat{J}_{\mathsf{ES}\text{-}\mu,\cdot,n}$ = spacings estimator of $J_{\mathsf{ES}\text{-}\mu,\cdot,n}$; $\check{J}_{\mathsf{ES}\text{-}\mu,\cdot,n}$ = polynomial estimator of $J_{\mathsf{ES}\text{-}\mu,\cdot,n}$.
$\hat{J}_{\mathsf{ES}\text{-}\vartheta,\cdot,n}$ = spacings estimator of $J_{\mathsf{ES}\text{-}\vartheta,\cdot,n}$; $\check{J}_{\mathsf{ES}\text{-}\vartheta,\cdot,n}$ = polynomial estimator of $J_{\mathsf{ES}\text{-}\vartheta,\cdot,n}$.
$\hat{J}_{\mathsf{QS}\text{-}\mu,\cdot,n}$ = spacings estimator of $J_{\mathsf{QS}\text{-}\mu,\cdot,n}$; $\check{J}_{\mathsf{QS}\text{-}\mu,\cdot,n}$ = polynomial estimator of $J_{\mathsf{QS}\text{-}\mu,\cdot,n}$.
$\hat{J}_{\mathsf{QS}\text{-}\vartheta,\cdot,n}$ = spacings estimator of $J_{\mathsf{QS}\text{-}\vartheta,\cdot,n}$; $\check{J}_{\mathsf{QS}\text{-}\vartheta,\cdot,n}$ = polynomial estimator of $J_{\mathsf{QS}\text{-}\vartheta,\cdot,n}$.

6 Numerical Comparison of Partitioning Schemes

We proposed two alternative ways of constructing RD plots, one employing ES partitioning and the other employing QS partitioning. While developing a general theory for optimal partitioning scheme selection is beyond the scope of this paper, we can employ our IMSE expansions to compare the two partitioning schemes theoretically in order to assess their relative IMSE-optimality properties.

Without loss of generality we focus on the IMSE for the treatment group ("+" subindex). Assuming the regularity conditions imposed in the paper hold, we obtain (up to the ceiling operator for selecting the optimal partition sizes):

$$\mathsf{IMSE}_{\mathtt{ES},+}(J_{\mathtt{ES},+,n}) = \frac{\sqrt[3]{3}}{4}\mathsf{C}_{\mathtt{ES},+}n^{-2/3}\{1+o_{\mathbb{P}}(1)\},$$

$$\mathsf{IMSE}_{\mathtt{QS},+}(J_{\mathtt{QS},+,n}) = \frac{\sqrt[3]{3}}{4}\mathsf{C}_{\mathtt{QS},+}n^{-2/3}\{1+o_{\mathbb{P}}(1)\},$$

where

$$\mathsf{C}_{\mathtt{ES},+} = \left(\int_{\bar{x}}^{x_u}\left(\mu_+^{(1)}(x)\right)^2 w(x)dx\right)^{1/3}\left(\int_{\bar{x}}^{x_u}\frac{\sigma_+^2(x)}{f(x)}w(x)dx\right)^{2/3},$$

$$\mathsf{C}_{\mathtt{QS},+} = \left(\int_{\bar{x}}^{x_u}\left(\frac{\mu_+^{(1)}(x)}{f(x)}\right)^2 w(x)dx\right)^{1/3}\left(\int_{\bar{x}}^{x_u}\sigma_+^2(x)w(x)dx\right)^{2/3}.$$

Thus, in order to compare the performance of the partition-size selectors for ES and QS RD plots we need to compare the two DGP constants $\mathsf{C}_{\mathtt{ES},+}$ and $\mathsf{C}_{\mathtt{QS},+}$. It follows that when $f(x) \propto \kappa$ (i.e., the running variable is uniformly distributed), then $\mathsf{C}_{\mathtt{ES},+} = \mathsf{C}_{\mathtt{QS},+}$ and therefore both partitioning schemes have equal (asymptotic) IMSE when the corresponding optimal partition size is used. Unfortunately, when the density $f(x)$ is not constant on the support $[x_l, x_u]$, it is not possible to obtain a unique ranking between $\mathsf{IMSE}_{\mathtt{ES},+}(J_{\mathtt{ES},+,n})$ and $\mathsf{IMSE}_{\mathtt{QS},+}(J_{\mathtt{QS},+,n})$. Heuristically, the QS RD plots should perform better in cases where the data is sparse because the estimated quantile spaced partition should adapt to this situation better, but we have been unable to provide a formal ranking along these lines.

Nonetheless, in Table SA-18 we explore the ranking between the two partitioning schemes using the 16 data generating processes discussed in our simulation study (Table SA-1). As expected, this

Table SA-18: Comparison of Partitioning Schemes

	$\frac{\mathscr{B}_{\mathtt{ES},-}}{\mathscr{B}_{\mathtt{QS},-}}$	$\frac{\mathscr{V}_{\mathtt{ES},-}}{\mathscr{V}_{\mathtt{QS},-}}$	$\frac{\mathsf{IMSE}_{\mathtt{ES},-}(J_{\mathtt{ES}\text{-}\mu,-,n})}{\mathsf{IMSE}_{\mathtt{QS},-}(J_{\mathtt{QS}\text{-}\mu,-,n})}$	$\frac{\mathscr{B}_{\mathtt{ES},+}}{\mathscr{B}_{\mathtt{QS},+}}$	$\frac{\mathscr{V}_{\mathtt{ES},+}}{\mathscr{V}_{\mathtt{QS},+}}$	$\frac{\mathsf{IMSE}_{\mathtt{ES},+}(J_{\mathtt{ES}\text{-}\mu,+,n})}{\mathsf{IMSE}_{\mathtt{QS},+}(J_{\mathtt{QS}\text{-}\mu,+,n})}$
Model 1	1.000	1.000	1.000	1.000	1.000	1.000
Model 2	2.290	1.000	1.319	0.784	1.000	0.925
Model 3	2.466	1.389	1.682	1.038	1.004	1.016
Model 4	1.258	1.004	1.084	0.447	1.389	0.953
Model 5	2.466	1.000	1.352	1.038	1.000	1.004
Model 6	1.258	1.000	1.081	0.447	1.000	0.765
Model 7	2.466	1.389	1.682	1.038	1.004	1.016
Model 8	1.258	1.004	1.084	0.447	1.389	0.953
Model 9	0.028	1.000	0.303	0.241	1.000	0.624
Model 10	0.309	1.000	0.677	0.655	1.000	0.867
Model 11	0.301	1.015	0.677	0.831	0.977	0.928
Model 12	0.309	0.977	0.666	0.570	1.015	0.839
Model 13	0.028	1.000	0.303	0.241	1.000	0.624
Model 14	0.309	1.000	0.677	0.655	1.000	0.867
Model 15	0.301	1.015	0.677	0.831	0.977	0.928
Model 16	0.309	0.977	0.666	0.570	1.015	0.839

table shows that when $f(x)$ is uniform both IMSE are equal, while when $f(x)$ is not uniform either IMSE may dominate the other. This depends on the shape of the regression function (different for control and treatment sides) and conditional heteroskedasticity in the underlying true data generating process.

References

Aras, G., Jammalamadaka, S. R., and Zhou, X. (1989), "Limit Distribution of Spacings Statistics when the Sample Size is Random," *Statistics and Probability Letters*, 8, 451–456.

Calonico, S., Cattaneo, M. D., and Titiunik, R. (2014a), "Robust Data-Driven Inference in the Regression-Discontinuity Design," *Stata Journal*, 14, 909–946.

——— (2014b), "Supplement to 'Robust Nonparametric Confidence Intervals for Regression-Discontinuity Designs'," *Econometrica Supplemental Material*, 82, http://dx.doi.org/10.3982/ECTA11757.

——— (2015), "`rdrobust`: An R Package for Robust Nonparametric Inference in Regression-Discontinuity Designs," *R Journal*, 7, 38–51.

Cattaneo, M. D., and Farrell, M. H. (2013), "Optimal Convergence Rates, Bahadur Representation, and Asymptotic Normality of Partitioning Estimators," *Journal of Econometrics*, 174, 127–143.

Cattaneo, M. D., Frandsen, B., and Titiunik, R. (2015), "Randomization Inference in the Regression Discontinuity Design: An Application to Party Advantages in the U.S. Senate," *Journal of Causal Inference*, 3, 1–24.

Lee, D. S. (2008), "Randomized Experiments from Non-random Selection in U.S. House Elections," *Journal of Econometrics*, 142, 675–697.

Ludwig, J., and Miller, D. L. (2007), "Does Head Start Improve Children's Life Chances? Evidence from a Regression Discontinuity Design," *Quarterly Journal of Economics*, 122, 159–208.

Mason, D. M. (1984), "A Strong Limit Theorem for the Oscillation Modulus of the Uniform Empirical Quantile Process," *Stochastic Processes and its Applications*, 17, 127–136.

Newey, W. K. (1997), "Convergence Rates and Asymptotic Normality for Series Estimators," *Journal of Econometrics*, 79, 147–168.

Shorack, G., and Wellner, J. (2009), *Empirical Process with Applications to Statistics*, Siam.