Online Appendix for *Intercept Estimation in Nonlinear Selection Models*

Wiji Arulampalam* Warwick University
Valentina Corradi † Surrey University
Daniel Gutknecht‡ Goethe University Frankfurt

March 12, 2023

Abstract This Online Appendix provides additional tables and figures on the Monte Carlo simulation exercise from Section 5. In particular, Section 1 contains the simulation results for the RMSE in table and graphic format (it contains the same results as in the main text, but also lists the results for two additional specifications of the estimator by Andrews and Schafgans (1998)). Section 2 on the other hand provides the corresponding results for the Mean and the Median Bias.

*Department of Economics, University of Warwick, Coventry CV4 7AL, UK. Email: Wiji.Arulampalam@warwick.ac.uk

†Corresponding Author: Department of Economics, University of Surrey, School of Economics, Guildford GU2 7XH, UK. Email: V.Corradi@surrey.ac.uk

‡Department of Business and Economics, Goethe University Frankfurt, Theodor-W.-Adorno-Platz 4, 60629 Frankfurt am Main, Germany. Email: Gutknecht@wiwi.uni-frankfurt.de

1 RMSE

Table 1: Additive Error Model (CASE I) - $\rho = 0$, RMSE

		Panel A: Oracle Index (w_i) - var(w_i) = 0.5 < var(v_i) = 1													
n		OLS	$h = 0.15$	$h = 0.10$	$h = 0.05$	$\widehat{h}$	HSZ(0.85)	HSZ(0.90)	HSZ(0.95)	AS(0.85, .5)	AS(0.85, 1)	AS(0.90, .5)	AS(0.90, 1)	AS(0.95, .5)	AS(0.95, 1)
600	RMSE	0.040	0.111	0.136	0.189	0.122	0.104	0.127	0.182	0.122	0.136	0.154	0.171	0.225	0.248
1000	RMSE	0.031	0.090	0.108	0.149	0.099	0.082	0.101	0.144	0.097	0.107	0.121	0.132	0.172	0.191
		Panel B: Oracle Index (w_i) - var(w_i) = var(v_i) = 1													
n		OLS	$h = 0.15$	$h = 0.10$	$h = 0.05$	$\widehat{h}$	HSZ(0.85)	HSZ(0.90)	HSZ(0.95)	AS(0.85, .5)	AS(0.85, 1)	AS(0.90, .5)	AS(0.90, 1)	AS(0.95, .5)	AS(0.95, 1)
600	RMSE	0.041	0.109	0.132	0.184	0.120	0.105	0.128	0.184	0.119	0.131	0.148	0.164	0.217	0.239
1000	RMSE	0.031	0.088	0.105	0.145	0.099	0.082	0.102	0.143	0.094	0.103	0.117	0.127	0.165	0.183
		Panel C: Oracle Index (w_i) - var(w_i) = 1.25 > var(v_i) = 1													
n		OLS	$h = 0.15$	$h = 0.10$	$h = 0.05$	$\widehat{h}$	HSZ(0.85)	HSZ(0.90)	HSZ(0.95)	AS(0.85, .5)	AS(0.85, 1)	AS(0.90, .5)	AS(0.90, 1)	AS(0.95, .5)	AS(0.95, 1)
600	RMSE	0.040	0.108	0.131	0.184	0.122	0.106	0.128	0.184	0.118	0.129	0.145	0.161	0.214	0.234
1000	RMSE	0.031	0.087	0.104	0.144	0.095	0.082	0.102	0.143	0.092	0.101	0.116	0.125	0.163	0.179
		Panel D: Oracle Index (w_i) - var(w_i) = 1.5 > var(v_i) = 1													
n		OLS	$h = 0.15$	$h = 0.10$	$h = 0.05$	$\widehat{h}$	HSZ(0.85)	HSZ(0.90)	HSZ(0.95)	AS(0.85, .5)	AS(0.85, 1)	AS(0.90, .5)	AS(0.90, 1)	AS(0.95, .5)	AS(0.95, 1)
600	RMSE	0.041	0.107	0.130	0.183	0.125	0.105	0.129	0.183	0.117	0.127	0.144	0.158	0.211	0.231
1000	RMSE	0.031	0.086	0.104	0.144	0.095	0.082	0.101	0.144	0.092	0.100	0.115	0.124	0.162	0.177
		Panel E: Klein Spady Index ($\widehat{w}_i$) - var(w_i) = var(v_i) = 1													
n		OLS	$h = 0.15$	$h = 0.10$	$h = 0.05$	$\widehat{h}$	HSZ(0.85)	HSZ(0.90)	HSZ(0.95)	AS(0.85, .5)	AS(0.85, 1)	AS(0.90, .5)	AS(0.90, 1)	AS(0.95, .5)	AS(0.95, 1)
600	RMSE	0.040	0.113	0.136	0.189	0.126	0.108	0.131	0.188	0.123	0.135	0.153	0.167	0.215	0.234
1000	RMSE	0.031	0.087	0.107	0.148	0.101	0.084	0.104	0.144	0.096	0.105	0.118	0.128	0.167	0.184

Notes: (1) Number of Monte Carlo replications: 1500; (2) columns $h = 0.15, 0.10, 0.05$ correspond to the estimator $\widehat{\theta}_A$ with a fixed bandwidth size, while $\widehat{h}$ denotes the same estimator with a data-driven bandwidth; (3) HSZ$(\cdot)$ corresponds to the estimator (18), with δ_n set to the 85%, 90%, and 95% (unconditional) quantile of $z_i'\widehat{\gamma}$; (4) AS$(\cdot, \cdot)$ corresponds to the estimator in (19), with δ_n again set as in (3) and $b \in \{0.5, 1\}$.

Table 2: Additive Error Model (CASE I) - $\rho = +0.5$, RMSE

		Panel A: Oracle Index (w_i) - var(w_i) = 0.5 < var(v_i) = 1													
n		OLS	$h = 0.15$	$h = 0.10$	$h = 0.05$	$\widehat{h}$	HSZ(0.85)	HSZ(0.90)	HSZ(0.95)	AS(0.85, .5)	AS(0.85, 1)	AS(0.90, .5)	AS(0.90, 1)	AS(0.95, .5)	AS(0.95, 1)
600	RMSE	0.234	0.118	0.138	0.187	0.136	0.121	0.135	0.181	0.131	0.142	0.156	0.171	0.221	0.244
1000	RMSE	0.232	0.098	0.111	0.146	0.108	0.106	0.111	0.145	0.108	0.114	0.125	0.134	0.171	0.188
		Panel B: Oracle Index (w_i) - var(w_i) = var(v_i) = 1													
n		OLS	$h = 0.15$	$h = 0.10$	$h = 0.05$	$\widehat{h}$	HSZ(0.85)	HSZ(0.90)	HSZ(0.95)	AS(0.85, .5)	AS(0.85, 1)	AS(0.90, .5)	AS(0.90, 1)	AS(0.95, .5)	AS(0.95, 1)
600	RMSE	0.203	0.107	0.130	0.184	0.126	0.108	0.127	0.182	0.120	0.131	0.148	0.162	0.213	0.235
1000	RMSE	0.202	0.083	0.101	0.141	0.102	0.086	0.099	0.142	0.092	0.101	0.114	0.125	0.164	0.179
		Panel C: Oracle Index (w_i) - var(w_i) = 1.25 > var(v_i) = 1													
n		OLS	$h = 0.15$	$h = 0.10$	$h = 0.05$	$\widehat{h}$	HSZ(0.85)	HSZ(0.90)	HSZ(0.95)	AS(0.85, .5)	AS(0.85, 1)	AS(0.90, .5)	AS(0.90, 1)	AS(0.95, .5)	AS(0.95, 1)
600	RMSE	0.192	0.106	0.129	0.184	0.132	0.105	0.127	0.182	0.118	0.129	0.146	0.159	0.210	0.231
1000	RMSE	0.191	0.082	0.100	0.140	0.103	0.084	0.099	0.143	0.090	0.099	0.112	0.123	0.162	0.177
		Panel D: Oracle Index (w_i) - var(w_i) = 1.5 > var(v_i) = 1													
n		OLS	$h = 0.15$	$h = 0.10$	$h = 0.05$	$\widehat{h}$	HSZ(0.85)	HSZ(0.90)	HSZ(0.95)	AS(0.85, .5)	AS(0.85, 1)	AS(0.90, .5)	AS(0.90, 1)	AS(0.95, .5)	AS(0.95, 1)
600	RMSE	0.183	0.107	0.129	0.184	0.136	0.105	0.127	0.183	0.117	0.127	0.145	0.158	0.209	0.228
1000	RMSE	0.181	0.083	0.100	0.141	0.103	0.083	0.098	0.143	0.089	0.097	0.111	0.122	0.161	0.175
		Panel E: Klein Spady Index ($\widehat{w}_i$) - var(w_i) = var(v_i) = 1													
n		OLS	$h = 0.15$	$h = 0.10$	$h = 0.05$	$\widehat{h}$	HSZ(0.85)	HSZ(0.90)	HSZ(0.95)	AS(0.85, .5)	AS(0.85, 1)	AS(0.90, .5)	AS(0.90, 1)	AS(0.95, .5)	AS(0.95, 1)
600	RMSE	0.204	0.114	0.135	0.185	0.128	0.111	0.132	0.182	0.123	0.134	0.151	0.165	0.213	0.233
1000	RMSE	0.202	0.083	0.103	0.142	0.097	0.087	0.099	0.143	0.094	0.102	0.114	0.125	0.161	0.180

Notes: (1) Number of Monte Carlo replications: 1500; (2) columns $h = 0.15, 0.10, 0.05$ correspond to the estimator $\widehat{\theta}_A$ with a fixed bandwidth size, while $\widehat{h}$ denotes the same estimator with a data-driven bandwidth; (3) HSZ($\cdot$) corresponds to the estimator (18), with δ_n set to the 85%, 90%, and 95% (unconditional) quantile of $z_i'\widehat{\gamma}$; (4) AS($\cdot, \cdot$) corresponds to the estimator in (19), with δ_n again set as in (3) and $b \in \{0.5, 1\}$.

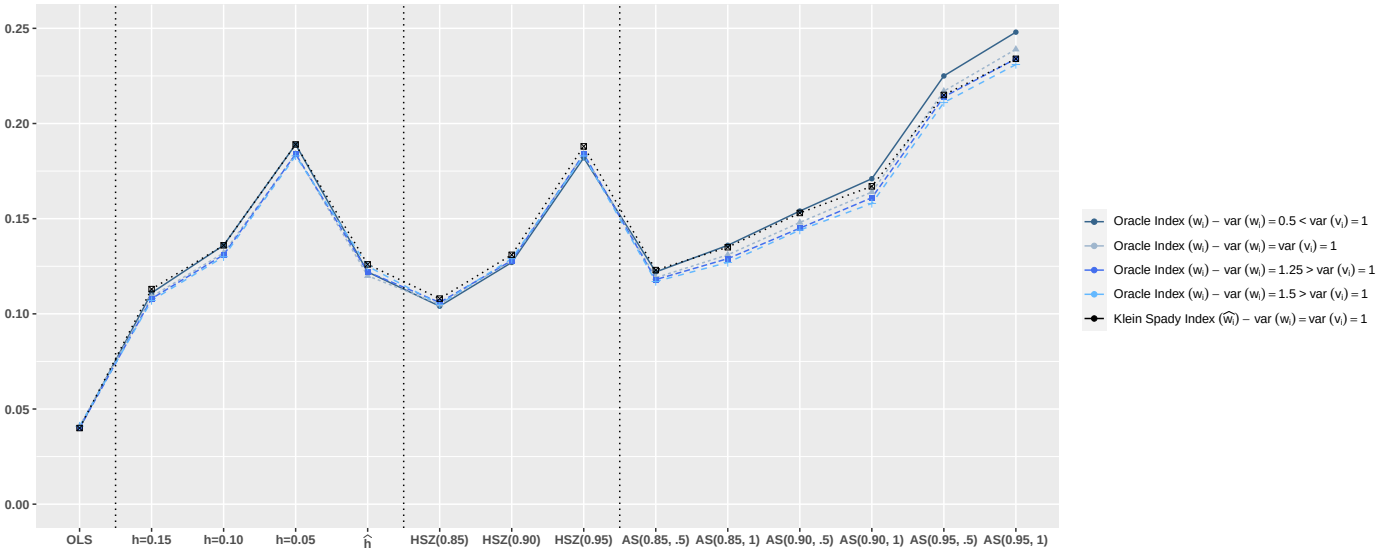
Table 3: Additive Error Model (CASE I) - $\rho = -0.5$, RMSE

		Panel A: Oracle Index (w_i) - var(w_i) = 0.5 < var(v_i) = 1													
n		OLS	$h = 0.15$	$h = 0.10$	$h = 0.05$	$\widehat{h}$	HSZ(0.85)	HSZ(0.90)	HSZ(0.95)	AS(0.85, .5)	AS(0.85, 1)	AS(0.90, .5)	AS(0.90, 1)	AS(0.95, .5)	AS(0.95, 1)
600	RMSE	0.234	0.121	0.139	0.188	0.130	0.124	0.137	0.184	0.134	0.144	0.157	0.173	0.223	0.246
1000	RMSE	0.231	0.098	0.112	0.149	0.111	0.105	0.113	0.146	0.109	0.116	0.127	0.137	0.173	0.192
		Panel B: Oracle Index (w_i) - var(w_i) = var(v_i) = 1													
n		OLS	$h = 0.15$	$h = 0.10$	$h = 0.05$	$\widehat{h}$	HSZ(0.85)	HSZ(0.90)	HSZ(0.95)	AS(0.85, .5)	AS(0.85, 1)	AS(0.90, .5)	AS(0.90, 1)	AS(0.95, .5)	AS(0.95, 1)
600	RMSE	0.204	0.107	0.129	0.180	0.117	0.107	0.126	0.180	0.119	0.129	0.145	0.160	0.210	0.234
1000	RMSE	0.200	0.084	0.102	0.143	0.100	0.084	0.101	0.141	0.093	0.102	0.116	0.126	0.165	0.182
		Panel C: Oracle Index (w_i) - var(w_i) = 1.25 > var(v_i) = 1													
n		OLS	$h = 0.15$	$h = 0.10$	$h = 0.05$	$\widehat{h}$	HSZ(0.85)	HSZ(0.90)	HSZ(0.95)	AS(0.85, .5)	AS(0.85, 1)	AS(0.90, .5)	AS(0.90, 1)	AS(0.95, .5)	AS(0.95, 1)
600	RMSE	0.193	0.106	0.128	0.179	0.118	0.104	0.125	0.180	0.117	0.127	0.144	0.158	0.208	0.229
1000	RMSE	0.189	0.082	0.102	0.142	0.099	0.081	0.099	0.142	0.091	0.100	0.114	0.124	0.162	0.179
		Panel D: Oracle Index (w_i) - var(w_i) = 1.5 > var(v_i) = 1													
n		OLS	$h = 0.15$	$h = 0.10$	$h = 0.05$	$\widehat{h}$	HSZ(0.85)	HSZ(0.90)	HSZ(0.95)	AS(0.85, .5)	AS(0.85, 1)	AS(0.90, .5)	AS(0.90, 1)	AS(0.95, .5)	AS(0.95, 1)
600	RMSE	0.183	0.105	0.128	0.180	0.121	0.103	0.126	0.180	0.116	0.125	0.143	0.156	0.206	0.227
1000	RMSE	0.180	0.082	0.101	0.142	0.097	0.080	0.099	0.141	0.089	0.098	0.112	0.122	0.161	0.176
		Panel E: Klein Spady Index ($\widehat{w}_i$) - var(w_i) = var(v_i) = 1													
n		OLS	$h = 0.15$	$h = 0.10$	$h = 0.05$	$\widehat{h}$	HSZ(0.85)	HSZ(0.90)	HSZ(0.95)	AS(0.85, .5)	AS(0.85, 1)	AS(0.90, .5)	AS(0.90, 1)	AS(0.95, .5)	AS(0.95, 1)
600	RMSE	0.204	0.109	0.129	0.178	0.118	0.108	0.129	0.175	0.120	0.129	0.145	0.157	0.201	0.223
1000	RMSE	0.202	0.085	0.105	0.147	0.102	0.086	0.105	0.144	0.097	0.105	0.118	0.128	0.166	0.185

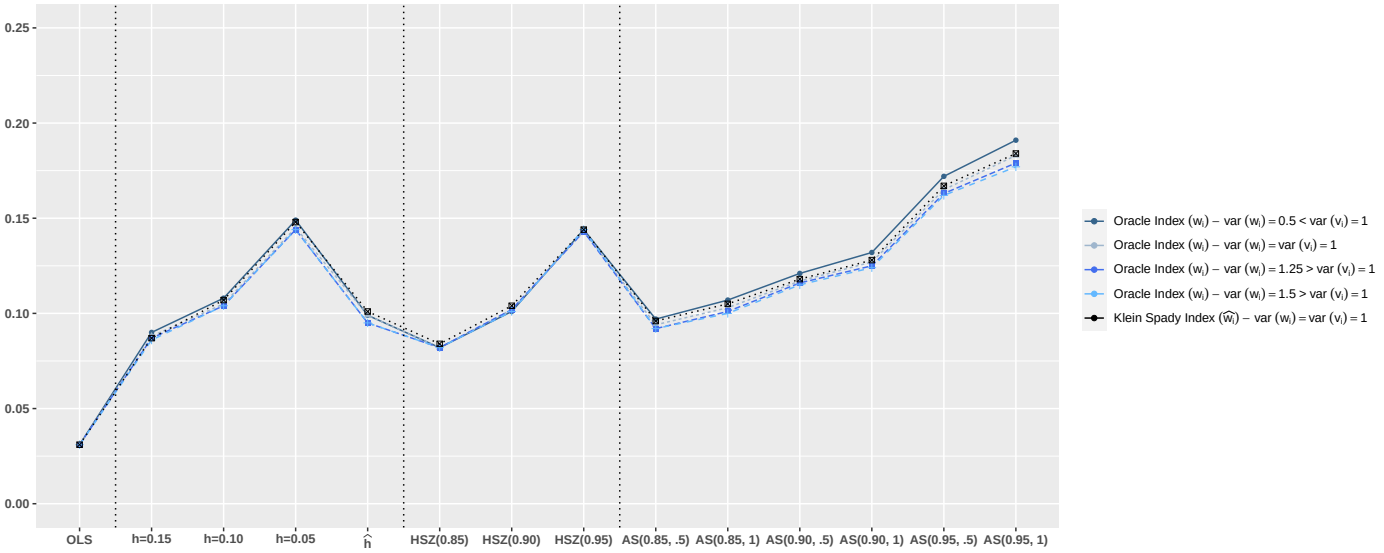
Notes: (1) Number of Monte Carlo replications: 1500; (2) columns $h = 0.15, 0.10, 0.05$ correspond to the estimator $\widehat{\theta}_A$ with a fixed bandwidth size, while $\widehat{h}$ denotes the same estimator with a data-driven bandwidth; (3) HSZ($\cdot$) corresponds to the estimator (18), with δ_n set to the 85%, 90%, and 95% (unconditional) quantile of $z_i'\widehat{\gamma}$; (4) AS($\cdot, \cdot$) corresponds to the estimator in (19), with δ_n again set as in (3) and $b \in \{0.5, 1\}$.

Figure 1: RMSE Comparison - Additive Error Model (CASE I) - $\rho = 0$

(a) $n = 600$



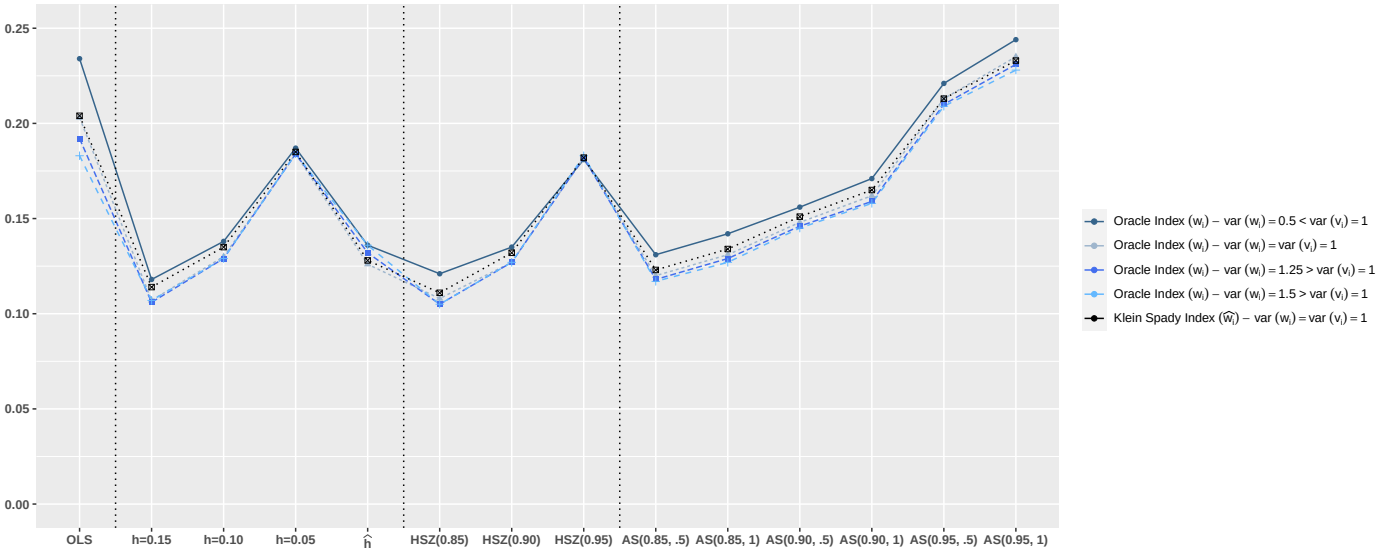
(b) $n = 1000$



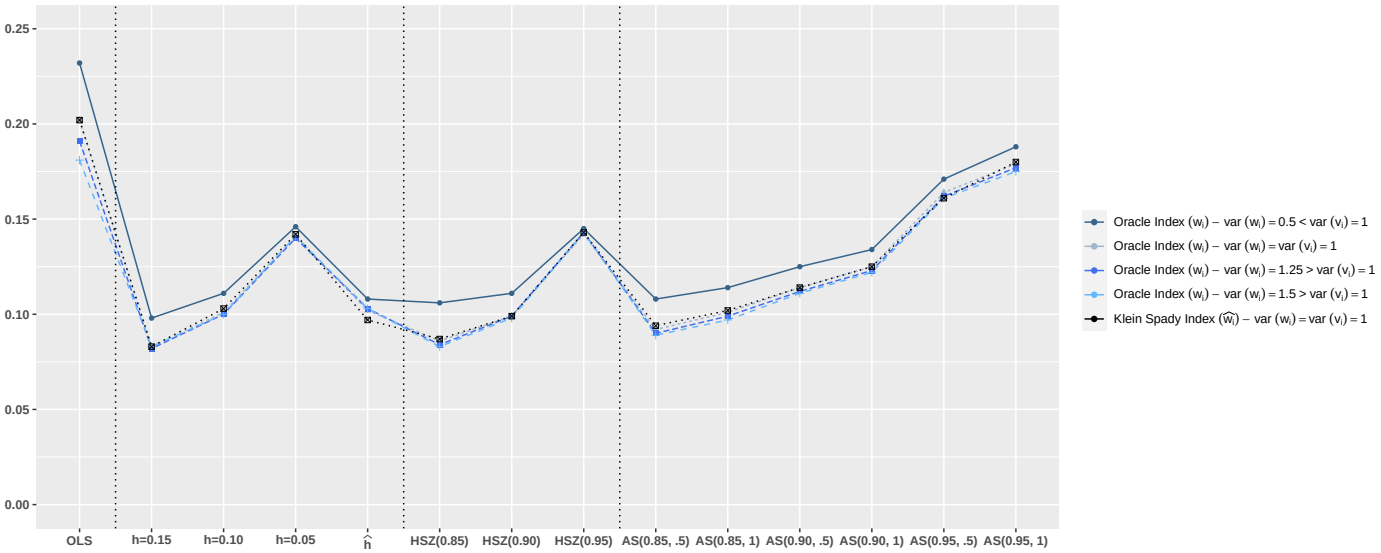
Notes: (1) Number of Monte Carlo replications: 1500; (2) $h = 0.15, 0.10, 0.05$ correspond to the estimator $\widehat{\theta}_A$ with a fixed bandwidth size, while $\widehat{h}$ denotes the same estimator with a data-driven bandwidth; (3) HSZ$(\cdot)$ corresponds to the estimator (18), with δ_n set to the 85%, 90%, and 95% (unconditional) quantile of $z_i'\widehat{\gamma}$; (4) AS$(\cdot,\cdot)$ corresponds to the estimator in (19), with δ_n again set as in (3) and $b \in \{0.5, 1\}$.

Figure 2: RMSE Comparison - Additive Error Model (CASE I) - $\rho = +0.5$

(a) $n = 600$



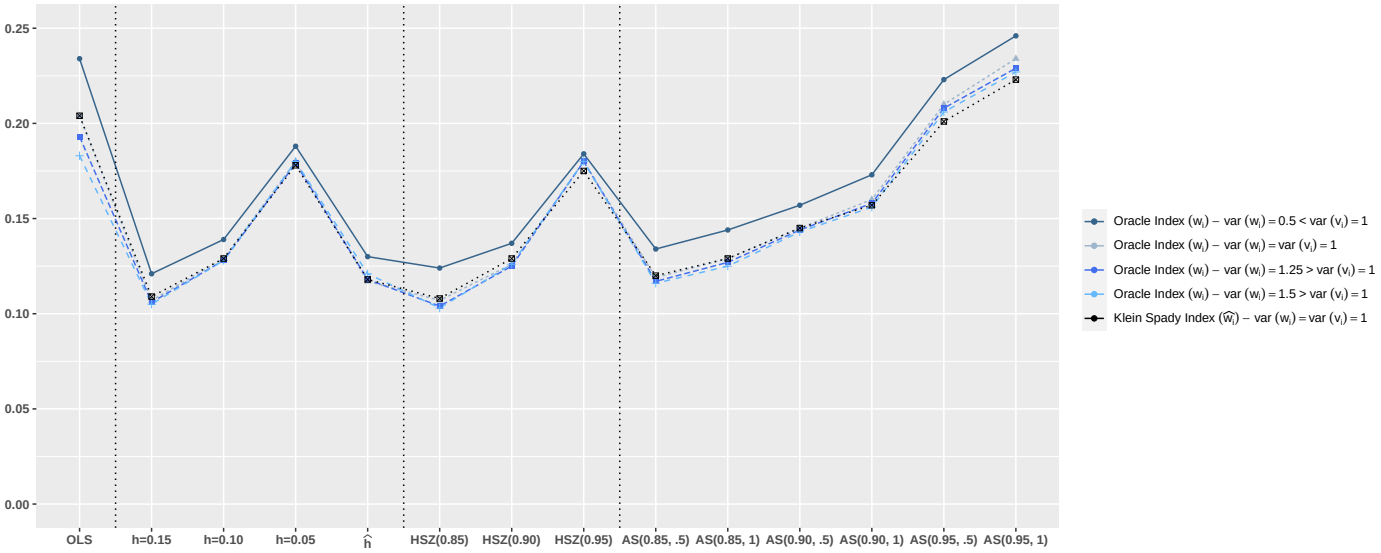
(b) $n = 1000$



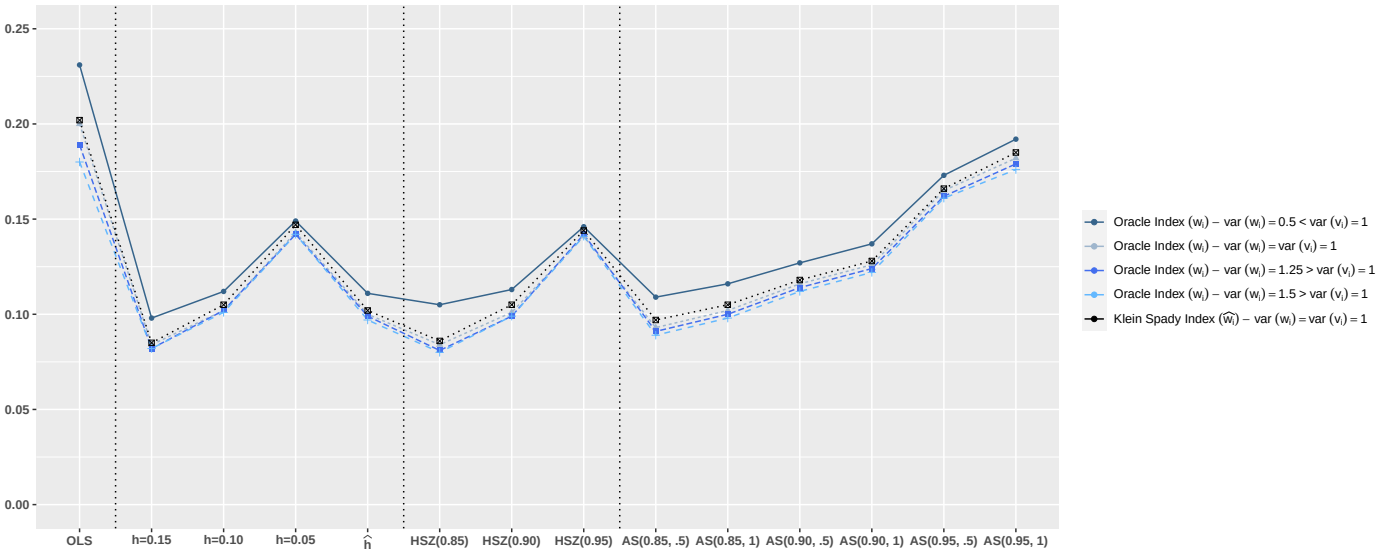
Notes: (1) Number of Monte Carlo replications: 1500; (2) $h = 0.15, 0.10, 0.05$ correspond to the estimator $\widehat{\theta}_A$ with a fixed bandwidth size, while $\widehat{h}$ denotes the same estimator with a data-driven bandwidth; (3) HSZ$(\cdot)$ corresponds to the estimator (18), with δ_n set to the 85%, 90%, and 95% (unconditional) quantile of $z_i'\widehat{\gamma}$; (4) AS$(\cdot,\cdot)$ corresponds to the estimator in (19), with δ_n again set as in (3) and $b \in \{0.5, 1\}$.

Figure 3: RMSE Comparison - Additive Error Model (CASE I) - $\rho = -0.5$

(a) $n = 600$



(b) $n = 1000$



Notes: (1) Number of Monte Carlo replications: 1500; (2) $h = 0.15, 0.10, 0.05$ correspond to the estimator $\widehat{\theta}_A$ with a fixed bandwidth size, while $\widehat{h}$ denotes the same estimator with a data-driven bandwidth; (3) HSZ$(\cdot)$ corresponds to the estimator (18), with δ_n set to the 85%, 90%, and 95% (unconditional) quantile of $z_i'\widehat{\gamma}$; (4) AS$(\cdot,\cdot)$ corresponds to the estimator in (19), with δ_n again set as in (3) and $b \in \{0.5, 1\}$.

2 Mean & Median Bias

Table 4: Additive Error Model (CASE I) - $\rho = 0$, MBIAS and MDBIAS

		Panel A: Oracle Index (w_i) - var(w_i) = 0.5 < var(v_i) = 1													
n		OLS	$h = 0.15$	$h = 0.10$	$h = 0.05$	$\widehat{h}$	HSZ(0.85)	HSZ(0.90)	HSZ(0.95)	AS(0.85, .5)	AS(0.85, 1)	AS(0.90, .5)	AS(0.90, 1)	AS(0.95, .5)	AS(0.95, 1)
600	MBIAS	−0.001	0.001	0.001	0.003	0.002	0.002	0.003	0.003	0.002	0.002	0.002	0.000	−0.002	0.000
	MDBIAS	0.001	−0.001	−0.004	0.000	0.004	−0.002	0.000	0.003	0.000	0.000	0.002	−0.002	0.001	0.004
1000	MBIAS	−0.001	0.000	−0.001	0.002	0.003	−0.001	0.000	0.002	0.000	0.001	0.000	0.002	0.002	0.004
	MDBIAS	−0.001	0.001	−0.001	0.007	0.002	0.001	0.002	0.005	0.002	0.003	0.002	0.005	0.004	−0.001
		Panel B: Oracle Index (w_i) - var(w_i) = var(v_i) = 1													
n		OLS	$h = 0.15$	$h = 0.10$	$h = 0.05$	$\widehat{h}$	HSZ(0.85)	HSZ(0.90)	HSZ(0.95)	AS(0.85, .5)	AS(0.85, 1)	AS(0.90, .5)	AS(0.90, 1)	AS(0.95, .5)	AS(0.95, 1)
600	MBIAS	0.000	0.002	0.002	0.004	0.003	0.001	0.003	0.004	0.003	0.003	0.004	0.002	−0.001	0.000
	MDBIAS	−0.002	0.000	0.001	0.002	0.003	−0.001	0.001	0.006	0.001	0.002	0.002	0.002	0.001	0.003
1000	MBIAS	−0.001	0.000	−0.002	0.002	0.002	−0.002	0.000	0.002	0.000	0.000	0.000	0.000	0.001	0.003
	MDBIAS	−0.002	0.000	0.001	0.003	0.004	−0.001	0.002	0.005	0.000	0.003	0.001	0.002	0.004	0.000
		Panel C: Oracle Index (w_i) - var(w_i) = 1.25 > var(v_i) = 1													
n		OLS	$h = 0.15$	$h = 0.10$	$h = 0.05$	$\widehat{h}$	HSZ(0.85)	HSZ(0.90)	HSZ(0.95)	AS(0.85, .5)	AS(0.85, 1)	AS(0.90, .5)	AS(0.90, 1)	AS(0.95, .5)	AS(0.95, 1)
600	MBIAS	0.000	0.002	0.002	0.004	0.001	0.002	0.004	0.004	0.003	0.004	0.004	0.002	0.000	−0.001
	MDBIAS	−0.002	0.002	0.003	0.002	0.001	0.002	0.001	0.005	0.002	0.004	0.002	0.005	0.003	0.004
1000	MBIAS	−0.001	−0.001	−0.002	0.002	0.002	−0.002	0.000	0.001	0.000	0.000	0.000	0.000	0.000	0.002
	MDBIAS	−0.001	−0.001	−0.001	0.002	0.004	−0.001	0.000	0.002	0.000	0.002	0.001	0.001	0.002	0.001
		Panel D: Oracle Index (w_i) - var(w_i) = 1.5 > var(v_i) = 1													
n		OLS	$h = 0.15$	$h = 0.10$	$h = 0.05$	$\widehat{h}$	HSZ(0.85)	HSZ(0.90)	HSZ(0.95)	AS(0.85, .5)	AS(0.85, 1)	AS(0.90, .5)	AS(0.90, 1)	AS(0.95, .5)	AS(0.95, 1)
600	MBIAS	0.000	0.002	0.002	0.004	0.004	0.002	0.004	0.004	0.004	0.004	0.004	0.003	0.001	−0.001
	MDBIAS	−0.001	0.000	0.003	0.002	0.006	0.002	0.000	0.005	0.004	0.005	0.003	0.004	0.003	0.005
1000	MBIAS	−0.001	0.000	−0.002	0.002	−0.001	−0.003	0.000	0.001	0.000	0.000	0.000	0.000	0.000	0.002
	MDBIAS	−0.002	−0.001	−0.001	0.002	0.004	0.000	0.002	0.001	−0.001	0.001	0.001	0.000	0.002	0.001
		Panel E: Klein Spady Index ($\widehat{w}_i$) - var(w_i) = var(v_i) = 1													
n		OLS	$h = 0.15$	$h = 0.10$	$h = 0.05$	$\widehat{h}$	HSZ(0.85)	HSZ(0.90)	HSZ(0.95)	AS(0.85, .5)	AS(0.85, 1)	AS(0.90, .5)	AS(0.90, 1)	AS(0.95, .5)	AS(0.95, 1)
600	MBIAS	0.000	0.001	0.004	0.011	0.004	0.003	0.007	0.009	0.006	0.007	0.009	0.007	0.005	0.005
	MDBIAS	−0.002	0.001	0.005	0.011	0.000	0.003	0.007	0.009	0.004	0.005	0.008	0.005	0.001	0.005
1000	MBIAS	0.000	0.003	0.002	0.003	0.001	0.002	0.001	0.003	0.002	0.002	0.002	0.002	0.000	0.002
	MDBIAS	0.000	0.008	0.002	0.002	0.002	0.003	−0.001	0.001	0.002	0.003	0.001	0.001	−0.003	0.002

Notes: (1) Number of Monte Carlo replications: 1500; (2) columns $h = 0.15, 0.10, 0.05$ correspond to the estimator $\widehat{\theta}_A$ with a fixed bandwidth size, while $\widehat{h}$ denotes the same estimator with a data-driven bandwidth; (3) HSZ$(\cdot)$ corresponds to the estimator (18), with δ_n set to the 85%, 90%, and 95% (unconditional) quantile of $z_i'\widehat{\gamma}$; (4) AS$(\cdot,\cdot)$ corresponds to the estimator in (19), with δ_n again set as in (3) and $b \in \{0.5, 1\}$.

Table 5: Additive Error Model (CASE I) - $\rho = +0.5$, MBIAS and MDBIAS

		Panel A: Oracle Index (w_i) - var(w_i) = 0.5 < var(v_i) = 1												
n		OLS	$h = 0.15$	$h = 0.10$	$h = 0.05$	$\widehat{h}$	HSZ(0.85)	HSZ(0.90)	HSZ(0.95)	AS(0.85, .5)	AS(0.85, 1)	AS(0.90, .5)	AS(0.90, 1)	AS(0.95, .5)	AS(0.95, 1)
600	MBIAS	−0.231	−0.048	−0.042	−0.030	−0.046	−0.069	−0.056	−0.041	−0.054	−0.049	−0.044	−0.041	−0.036	−0.031
	MDBIAS	−0.229	−0.047	−0.039	−0.025	−0.054	−0.071	−0.055	−0.035	−0.055	−0.044	−0.042	−0.036	−0.037	−0.031
1000	MBIAS	−0.230	−0.049	−0.043	−0.028	−0.047	−0.072	−0.057	−0.038	−0.056	−0.049	−0.045	−0.038	−0.029	−0.023
	MDBIAS	−0.232	−0.045	−0.040	−0.023	−0.053	−0.070	−0.054	−0.034	−0.054	−0.048	−0.040	−0.041	−0.029	−0.024
		Panel B: Oracle Index (w_i) - var(w_i) = var(v_i) = 1												
n		OLS	$h = 0.15$	$h = 0.10$	$h = 0.05$	$\widehat{h}$	HSZ(0.85)	HSZ(0.90)	HSZ(0.95)	AS(0.85, .5)	AS(0.85, 1)	AS(0.90, .5)	AS(0.90, 1)	AS(0.95, .5)	AS(0.95, 1)
600	MBIAS	−0.199	−0.007	−0.006	−0.002	−0.002	−0.030	−0.018	−0.010	−0.020	−0.017	−0.012	−0.012	−0.011	−0.010
	MDBIAS	−0.199	−0.008	−0.006	−0.001	−0.003	−0.031	−0.020	−0.010	−0.023	−0.016	−0.012	−0.009	−0.016	−0.010
1000	MBIAS	−0.200	−0.010	−0.009	−0.002	−0.003	−0.033	−0.022	−0.010	−0.023	−0.019	−0.015	−0.012	−0.008	−0.004
	MDBIAS	−0.201	−0.008	−0.009	−0.002	−0.004	−0.032	−0.021	−0.005	−0.022	−0.019	−0.014	−0.012	−0.008	−0.006
		Panel C: Oracle Index (w_i) - var(w_i) = 1.25 > var(v_i) = 1												
n		OLS	$h = 0.15$	$h = 0.10$	$h = 0.05$	$\widehat{h}$	HSZ(0.85)	HSZ(0.90)	HSZ(0.95)	AS(0.85, .5)	AS(0.85, 1)	AS(0.90, .5)	AS(0.90, 1)	AS(0.95, .5)	AS(0.95, 1)
600	MBIAS	−0.188	0.003	0.002	0.004	0.010	−0.020	−0.009	−0.004	−0.012	−0.009	−0.005	−0.006	−0.006	−0.006
	MDBIAS	−0.188	0.002	0.000	0.005	0.007	−0.023	−0.012	−0.005	−0.017	−0.011	−0.008	−0.006	−0.012	−0.009
1000	MBIAS	−0.188	0.000	−0.001	0.003	0.008	−0.023	−0.013	−0.004	−0.015	−0.012	−0.008	−0.007	−0.004	−0.001
	MDBIAS	−0.190	0.002	−0.002	0.003	0.008	−0.023	−0.010	0.001	−0.015	−0.012	−0.009	−0.007	−0.004	0.000
		Panel D: Oracle Index (w_i) - var(w_i) = 1.5 > var(v_i) = 1												
n		OLS	$h = 0.15$	$h = 0.10$	$h = 0.05$	$\widehat{h}$	HSZ(0.85)	HSZ(0.90)	HSZ(0.95)	AS(0.85, .5)	AS(0.85, 1)	AS(0.90, .5)	AS(0.90, 1)	AS(0.95, .5)	AS(0.95, 1)
600	MBIAS	−0.178	0.009	0.006	0.006	0.017	−0.014	−0.004	0.000	−0.007	−0.005	−0.001	−0.002	−0.003	−0.004
	MDBIAS	−0.178	0.007	0.006	0.007	0.018	−0.014	−0.006	−0.002	−0.011	−0.006	−0.002	−0.001	−0.007	−0.007
1000	MBIAS	−0.179	0.006	0.002	0.004	0.012	−0.017	−0.009	−0.001	−0.010	−0.008	−0.006	−0.004	−0.002	0.000
	MDBIAS	−0.180	0.007	0.002	0.006	0.017	−0.017	−0.007	0.007	−0.009	−0.006	−0.005	−0.004	0.001	0.000
		Panel E: Klein Spady Index ($\widehat{w}_i$) - var(w_i) = var(v_i) = 1												
n		OLS	$h = 0.15$	$h = 0.10$	$h = 0.05$	$\widehat{h}$	HSZ(0.85)	HSZ(0.90)	HSZ(0.95)	AS(0.85, .5)	AS(0.85, 1)	AS(0.90, .5)	AS(0.90, 1)	AS(0.95, .5)	AS(0.95, 1)
600	MBIAS	−0.200	−0.011	−0.006	0.000	−0.004	−0.031	−0.021	−0.008	−0.021	−0.017	−0.012	−0.011	−0.008	−0.006
	MDBIAS	−0.202	−0.011	−0.005	0.002	−0.006	−0.030	−0.021	−0.006	−0.021	−0.013	−0.010	−0.008	−0.003	−0.012
1000	MBIAS	−0.200	−0.007	−0.004	0.000	−0.005	−0.028	−0.019	−0.008	−0.021	−0.017	−0.013	−0.009	−0.003	0.001
	MDBIAS	−0.201	−0.005	−0.001	0.001	−0.004	−0.029	−0.018	−0.006	−0.020	−0.015	−0.009	−0.009	−0.003	0.002

Notes: (1) Number of Monte Carlo replications: 1500; (2) columns $h = 0.15, 0.10, 0.05$ correspond to the estimator $\widehat{\theta}_A$ with a fixed bandwidth size, while $\widehat{h}$ denotes the same estimator with a data-driven bandwidth; (3) HSZ$(\cdot)$ corresponds to the estimator (18), with δ_n set to the 85%, 90%, and 95% (unconditional) quantile of $z_i'\widehat{\gamma}$; (4) AS$(\cdot,\cdot)$ corresponds to the estimator in (19), with δ_n again set as in (3) and $b \in \{0.5, 1\}$.

Table 6: Additive Error Model (CASE I) - $\rho = -0.5$, MBIAS and MDBIAS

		Panel A: Oracle Index (w_i) - $\text{var}(w_i) = 0.5 < \text{var}(v_i) = 1$													
n		OLS	$h = 0.15$	$h = 0.10$	$h = 0.05$	$\widehat{h}$	HSZ(0.85)	HSZ(0.90)	HSZ(0.95)	AS(0.85, .5)	AS(0.85, 1)	AS(0.90, .5)	AS(0.90, 1)	AS(0.95, .5)	AS(0.95, 1)
600	MBIAS	0.231	0.052	0.045	0.035	0.049	0.074	0.059	0.045	0.059	0.053	0.048	0.043	0.034	0.032
	MDBIAS	0.231	0.050	0.039	0.039	0.053	0.073	0.058	0.052	0.057	0.053	0.052	0.044	0.035	0.030
1000	MBIAS	0.229	0.049	0.041	0.034	0.049	0.070	0.058	0.043	0.057	0.051	0.046	0.042	0.033	0.031
	MDBIAS	0.229	0.048	0.040	0.034	0.055	0.070	0.061	0.045	0.057	0.051	0.046	0.044	0.031	0.037
		Panel B: Oracle Index (w_i) - $\text{var}(w_i) = \text{var}(v_i) = 1$													
n		OLS	$h = 0.15$	$h = 0.10$	$h = 0.05$	$\widehat{h}$	HSZ(0.85)	HSZ(0.90)	HSZ(0.95)	AS(0.85, .5)	AS(0.85, 1)	AS(0.90, .5)	AS(0.90, 1)	AS(0.95, .5)	AS(0.95, 1)
600	MBIAS	0.200	0.011	0.010	0.008	0.008	0.033	0.024	0.017	0.025	0.022	0.019	0.015	0.010	0.009
	MDBIAS	0.201	0.011	0.008	0.015	0.007	0.033	0.026	0.023	0.025	0.026	0.018	0.022	0.010	0.009
1000	MBIAS	0.198	0.009	0.006	0.005	0.006	0.029	0.021	0.011	0.022	0.019	0.015	0.013	0.007	0.007
	MDBIAS	0.197	0.007	0.006	0.003	0.004	0.030	0.017	0.009	0.020	0.014	0.010	0.011	0.001	0.003
		Panel C: Oracle Index (w_i) - $\text{var}(w_i) = 1.25 > \text{var}(v_i) = 1$													
n		OLS	$h = 0.15$	$h = 0.10$	$h = 0.05$	$\widehat{h}$	HSZ(0.85)	HSZ(0.90)	HSZ(0.95)	AS(0.85, .5)	AS(0.85, 1)	AS(0.90, .5)	AS(0.90, 1)	AS(0.95, .5)	AS(0.95, 1)
600	MBIAS	0.189	0.001	0.002	0.003	−0.002	0.023	0.015	0.011	0.017	0.015	0.012	0.009	0.006	0.005
	MDBIAS	0.189	0.001	0.000	0.011	−0.006	0.020	0.018	0.017	0.017	0.016	0.011	0.014	0.008	0.005
1000	MBIAS	0.187	0.000	−0.002	0.001	−0.005	0.019	0.013	0.006	0.015	0.012	0.009	0.007	0.003	0.004
	MDBIAS	0.187	−0.001	−0.004	0.000	−0.010	0.019	0.011	0.005	0.013	0.007	0.007	0.006	−0.003	0.000
		Panel D: Oracle Index (w_i) - $\text{var}(w_i) = 1.5 > \text{var}(v_i) = 1$													
n		OLS	$h = 0.15$	$h = 0.10$	$h = 0.05$	$\widehat{h}$	HSZ(0.85)	HSZ(0.90)	HSZ(0.95)	AS(0.85, .5)	AS(0.85, 1)	AS(0.90, .5)	AS(0.90, 1)	AS(0.95, .5)	AS(0.95, 1)
600	MBIAS	0.179	−0.005	−0.002	0.000	−0.012	0.016	0.010	0.008	0.012	0.010	0.008	0.006	0.004	0.002
	MDBIAS	0.180	−0.005	−0.005	0.004	−0.018	0.015	0.012	0.011	0.011	0.013	0.007	0.011	0.009	0.002
1000	MBIAS	0.177	−0.006	−0.006	−0.001	−0.011	0.013	0.009	0.003	0.010	0.008	0.006	0.005	0.001	0.002
	MDBIAS	0.177	−0.007	−0.008	0.000	−0.016	0.012	0.006	0.002	0.009	0.007	0.004	0.003	−0.005	−0.003
		Panel E: Klein Spady Index $(\widehat{w}_i)$ - $\text{var}(w_i) = \text{var}(v_i) = 1$													
n		OLS	$h = 0.15$	$h = 0.10$	$h = 0.05$	$\widehat{h}$	HSZ(0.85)	HSZ(0.90)	HSZ(0.95)	AS(0.85, .5)	AS(0.85, 1)	AS(0.90, .5)	AS(0.90, 1)	AS(0.95, .5)	AS(0.95, 1)
600	MBIAS	0.201	0.013	0.012	0.016	0.014	0.033	0.029	0.024	0.029	0.026	0.025	0.021	0.014	0.011
	MDBIAS	0.201	0.014	0.013	0.017	0.012	0.030	0.028	0.023	0.027	0.023	0.022	0.018	0.019	0.015
1000	MBIAS	0.199	0.010	0.006	0.003	0.008	0.030	0.021	0.009	0.023	0.018	0.015	0.011	0.004	0.003
	MDBIAS	0.199	0.011	0.005	0.004	0.005	0.030	0.022	0.012	0.024	0.018	0.014	0.012	0.000	−0.001

Notes: (1) Number of Monte Carlo replications: 1500; (2) columns $h = 0.15, 0.10, 0.05$ correspond to the estimator $\widehat{\theta}_A$ with a fixed bandwidth size, while $\widehat{h}$ denotes the same estimator with a data-driven bandwidth; (3) HSZ$(\cdot)$ corresponds to the estimator (18), with δ_n set to the 85%, 90%, and 95% (unconditional) quantile of $z_i'\widehat{\gamma}$; (4) AS$(\cdot,\cdot)$ corresponds to the estimator in (19), with δ_n again set as in (3) and $b \in \{0.5, 1\}$.

References

Andrews, D. and M. Schafgans (1998). Semiparametric estimation of the intercept of a sample selection model. *Review of Economic Studies 65*, 497–517.