Supplementary Material document for Bounded-memory adjusted scores estimation in generalized linear models with large data sets

Patrick Zietkiewicz*[1] and Ioannis Kosmidis†[1]

[1]Department of Statistics, University of Warwick
Gibbet Hill Road, Coventry, CV4 7AL, UK

July 14, 2023

Supplementary material

The Supplementary Material provides R code to reproduce all numerical results and figures in the main text and in the current document. The code is organized in the two directories `diverted-flights` and `high-dim-logistic`, for the case study of Section 4 and the computer experiment of Section 5, respectively. The `README` file in each directory provides specific instructions about have the results can be reproduced, along with the specific versions of the contributed R packages that have been used to produce the results. The `biglm` directory has a port of the `biglm` R package (Lumley, 2020), which implements the one- and two-pass IWLS variants for solving the bias-reducing adjusted score equations (Firth, 1993) and for maximum Jeffreys'-penalized likelihood estimation (Kosmidis and Firth, 2021). The Supplementary Material is also available at `https://github.com/ikosmidis/bigbr-supplementary-material`.

S1 Estimates and estimated standard errors for the parameters of model (12)

Table S1: Estimates and estimated standard errors for the parameters of model (12) through ML, mBR and mJPL with one- and two-pass IWLS implementations. The table also reports the elapsed time, number of iterations, and average time per iteration for each method and implementation. The parameters corresponding to reference categories are set to 0.

		ML		mBR		mJPL	
		1-pass	1-pass	1-pass	2-pass	1-pass	2-pass
α		−5.52	−6.38	−4.10	−4.10	−4.12	−4.12
		(4.89)	(53.18)	(0.35)	(0.35)	(0.36)	(0.36)
β_1	January	0	0	0	0	0	0
β_2	February	−0.04	−0.04	−0.04	−0.04	−0.04	−0.04
		(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
β_3	March	−0.10	−0.10	−0.10	−0.10	−0.10	−0.10

*patrick.zietkiewicz@warwick.ac.uk
†ioannis.kosmidis@warwick.ac.uk

		(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
β_4	April	-0.08	-0.08	-0.08	-0.08	-0.08	-0.08
		(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
β_5	May	0.01	0.01	0.01	0.01	0.01	0.01
		(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
β_6	June	0.06	0.06	0.06	0.06	0.06	0.06
		(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
β_7	July	-0.01	-0.01	-0.01	-0.01	-0.01	-0.01
		(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
β_8	August	-0.00	-0.00	-0.00	-0.00	-0.00	-0.00
		(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
β_9	September	-0.09	-0.09	-0.09	-0.09	-0.09	-0.09
		(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
β_{10}	October	-0.13	-0.13	-0.13	-0.13	-0.13	-0.13
		(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
β_{11}	November	-0.15	-0.15	-0.15	-0.15	-0.15	-0.15
		(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
β_{12}	December	0.03	0.03	0.03	0.03	0.03	0.03
		(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
γ_1	Monday	0	0	0	0	0	0
γ_2	Tuesday	-0.01	-0.01	-0.01	-0.01	-0.01	-0.01
		(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
γ_3	Wednesday	-0.02	-0.02	-0.02	-0.02	-0.02	-0.02
		(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
γ_4	Thursday	0.05	0.05	0.05	0.05	0.05	0.05
		(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
γ_5	Friday	0.09	0.09	0.09	0.09	0.09	0.09
		(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
γ_6	Saturday	-0.02	-0.02	-0.02	-0.02	-0.02	-0.02
		(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
γ_7	Sunday	0.07	0.07	0.07	0.07	0.07	0.07
		(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
δ_1	AA	2.49	3.36	1.08	1.08	1.09	1.09
		(4.89)	(53.18)	(0.35)	(0.35)	(0.36)	(0.36)
δ_2	AQ	0	0	0	0	0	0
δ_3	AS	2.59	3.45	1.17	1.17	1.18	1.18
		(4.89)	(53.18)	(0.35)	(0.35)	(0.36)	(0.36)
δ_4	CO	2.33	3.19	0.91	0.91	0.92	0.92
		(4.89)	(53.18)	(0.35)	(0.35)	(0.36)	(0.36)
δ_5	DL	2.36	3.22	0.94	0.94	0.95	0.95
		(4.89)	(53.18)	(0.35)	(0.35)	(0.36)	(0.36)
δ_6	HP	2.18	3.05	0.76	0.76	0.78	0.78
		(4.89)	(53.18)	(0.35)	(0.35)	(0.36)	(0.36)
δ_7	NW	2.44	3.30	1.02	1.02	1.04	1.04
		(4.89)	(53.18)	(0.35)	(0.35)	(0.36)	(0.36)
δ_8	TW	2.36	3.23	0.94	0.94	0.96	0.96
		(4.89)	(53.18)	(0.35)	(0.35)	(0.36)	(0.36)
δ_9	UA	2.35	3.22	0.94	0.94	0.95	0.95
		(4.89)	(53.18)	(0.35)	(0.35)	(0.36)	(0.36)
δ_{10}	US	2.54	3.41	1.12	1.12	1.14	1.14

		(4.89)	(53.18)	(0.35)	(0.35)	(0.36)	(0.36)
δ_{11}	WN	2.41	3.28	0.99	0.99	1.01	1.01
		(4.89)	(53.18)	(0.35)	(0.35)	(0.36)	(0.36)
$\zeta_{(d)}$	Departure time	0.00	0.00	0.00	0.00	0.00	0.00
		(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
$\zeta_{(a)}$	Arrival Time	0.00	0.00	0.00	0.00	0.00	0.00
		(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
ρ	Distance	0.00	0.00	0.00	0.00	0.00	0.00
		(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
$\psi_{(d),1}$	Origin x	0.03	0.03	0.03	0.03	0.03	0.03
		(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
$\psi_{(d),1}$	Origin y	0.05	0.05	0.05	0.05	0.05	0.05
		(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
$\psi_{(d),1}$	Origin z	-0.02	-0.02	-0.02	-0.02	-0.02	-0.02
		(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
$\psi_{(a),1}$	Destination x	-0.00	-0.00	-0.00	-0.00	-0.00	-0.00
		(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
$\psi_{(a),2}$	Destination y	-0.05	-0.05	-0.05	-0.05	-0.05	-0.05
		(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
$\psi_{(a),3}$	Destination z	0.02	0.02	0.02	0.02	0.02	0.02
		(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Time (sec)		216.78	276.82	220.52	380.75	217.59	378.03
Iterations		15	20	12	12	12	12
Time/Iteration (sec)		14.45	13.84	18.38	31.73	18.13	31.50

S2 Simulation experiment in Section 5: Setting a

Figures S1-S18 show the phase transition curves derived using Candès and Sur (2020, Theorem 2.1) for $\rho^2 \in \{0, 0.25, 0.5, 0.75, 0.8, 0.9\}$ for the computer experiment in Section 5, overlaid with scatterplots of estimates of β from 5 independent realizations of y and X vs the true β, at 30 (κ, γ) points, for $n \in \{1000, 2000, 3000\}$, and with β^* set to an equi-spaced grid of length $p = \lceil n\kappa \rceil$ between -10 and 10. The ML estimates asymptotically exist in the white regions, and do not asymptotically exist in the grey regions. The top figure shows the ML estimator $\hat{\beta}$ (red) and mJPL estimator $\tilde{\beta}$ (orange), and the bottom figure shows $\hat{\beta}$ (red) and the estimator $\beta^\dagger = q(\kappa, \gamma, \rho)\tilde{\beta}$ (orange) with $q(\kappa, \gamma, \rho)$ as in (13). The black lines are intercept zero and slope one reference lines, and the green lines are estimated from a simple linear regression of the estimates on the truth.

Figure S1: Setting a with $n = 1000$ and $\rho^2 = 0$.

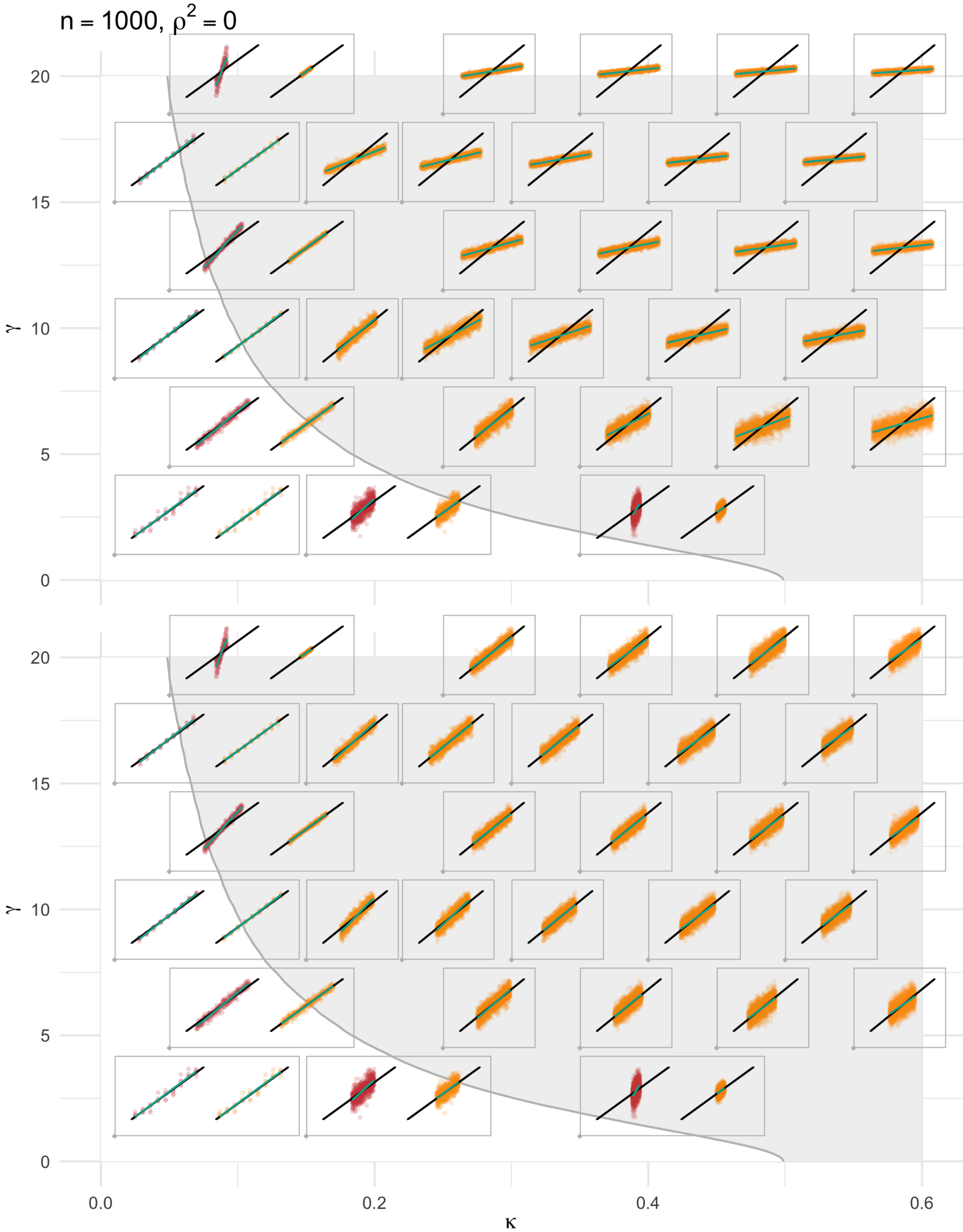


Figure S2: Setting a with $n = 1000$ and $\rho^2 = 0.25$.

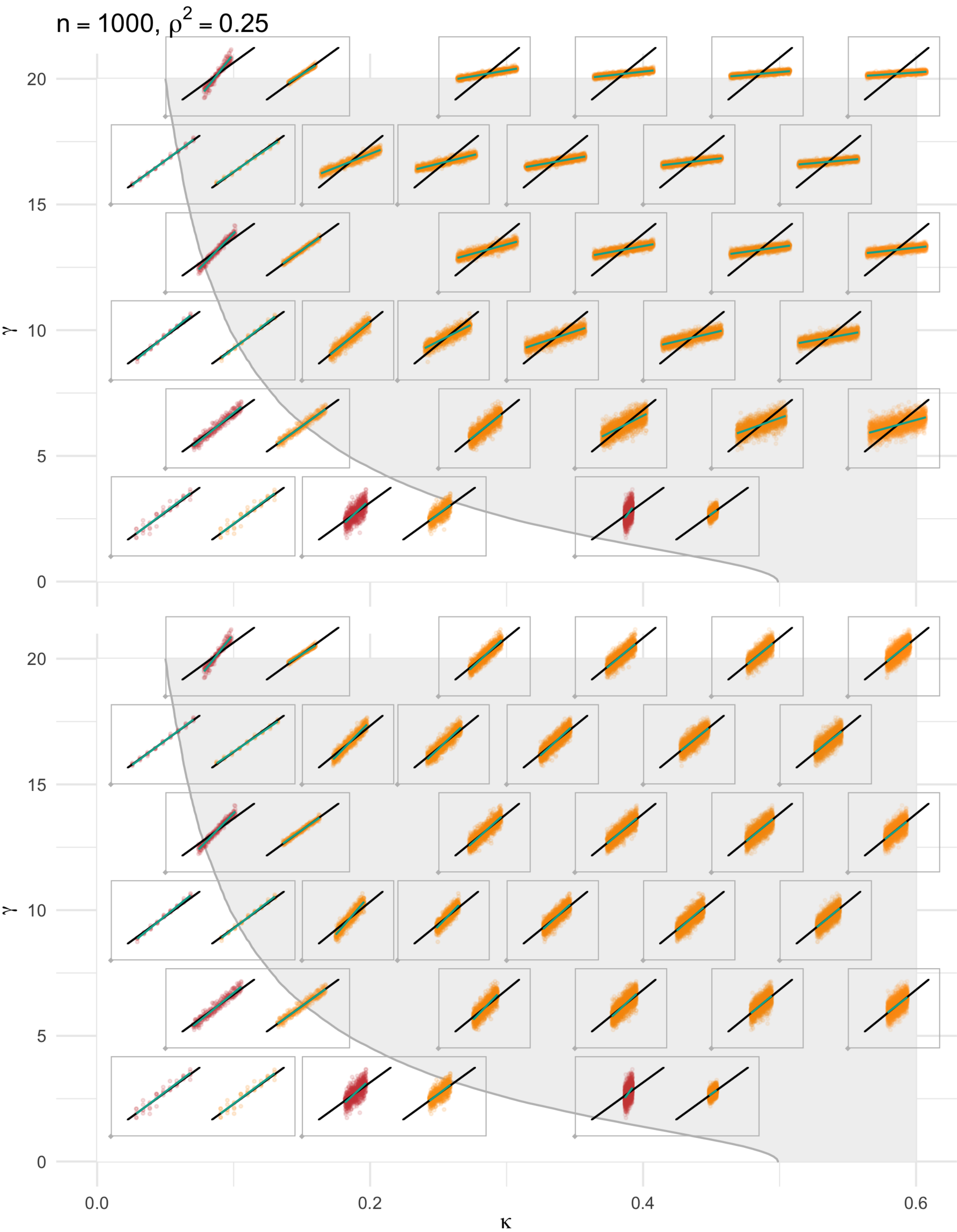


Figure S3: Setting a with $n = 1000$ and $\rho^2 = 0.5$.

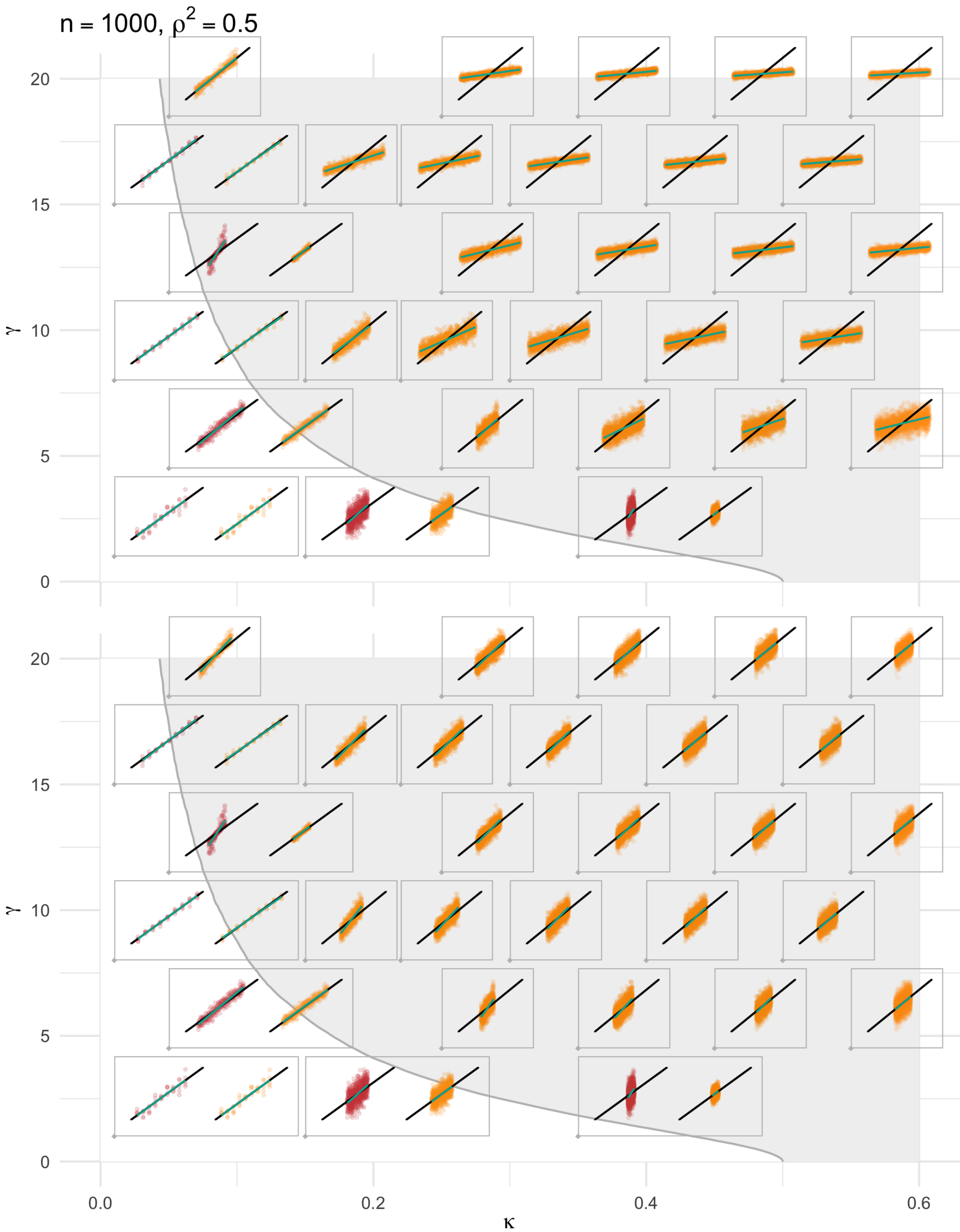


Figure S4: Setting a with $n = 1000$ and $\rho^2 = 0.75$.

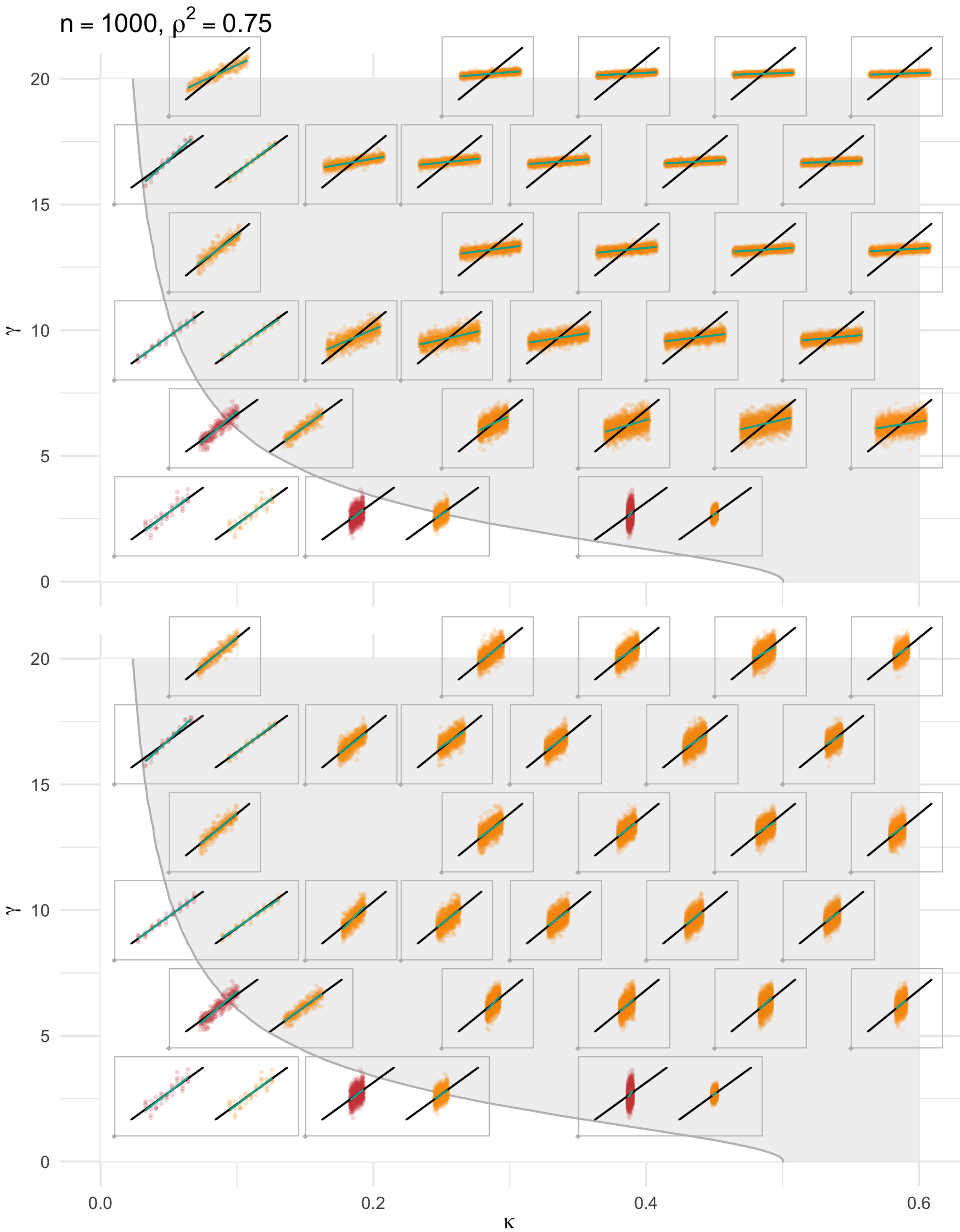


Figure S5: Setting a with $n = 1000$ and $\rho^2 = 0.8$.

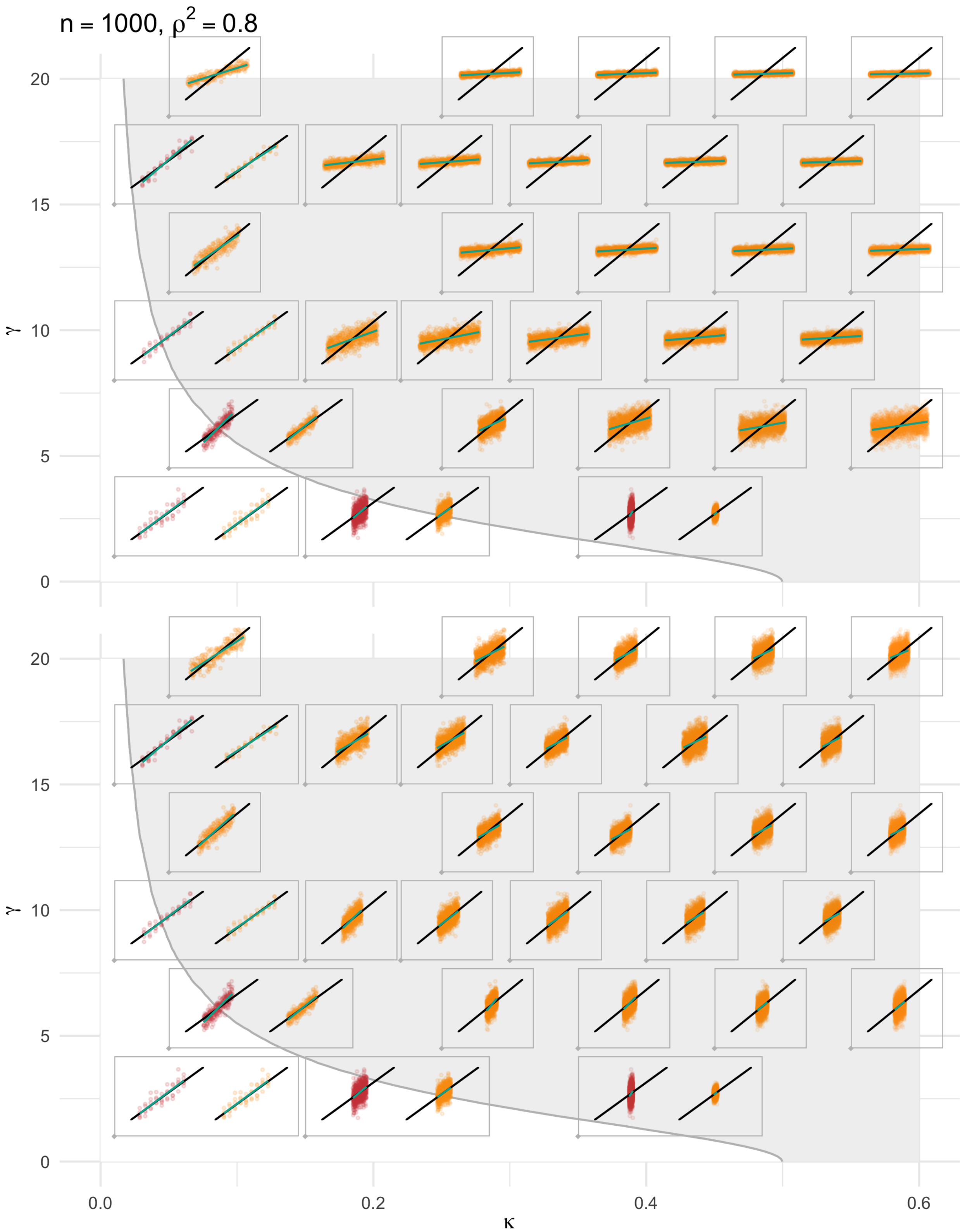


Figure S6: Setting a with $n = 1000$ and $\rho^2 = 0.9$.

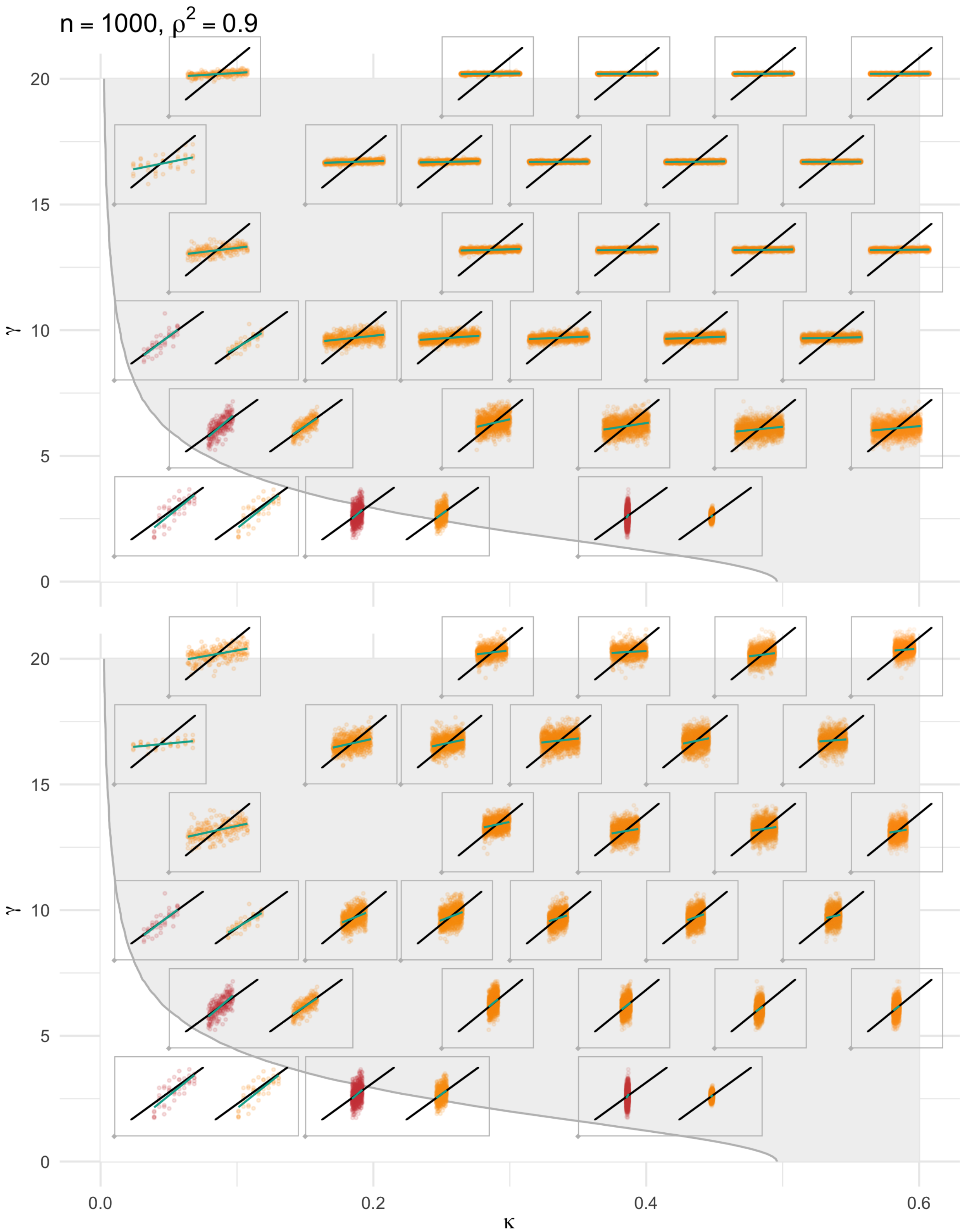


Figure S7: Setting a with $n = 2000$ and $\rho^2 = 0$.

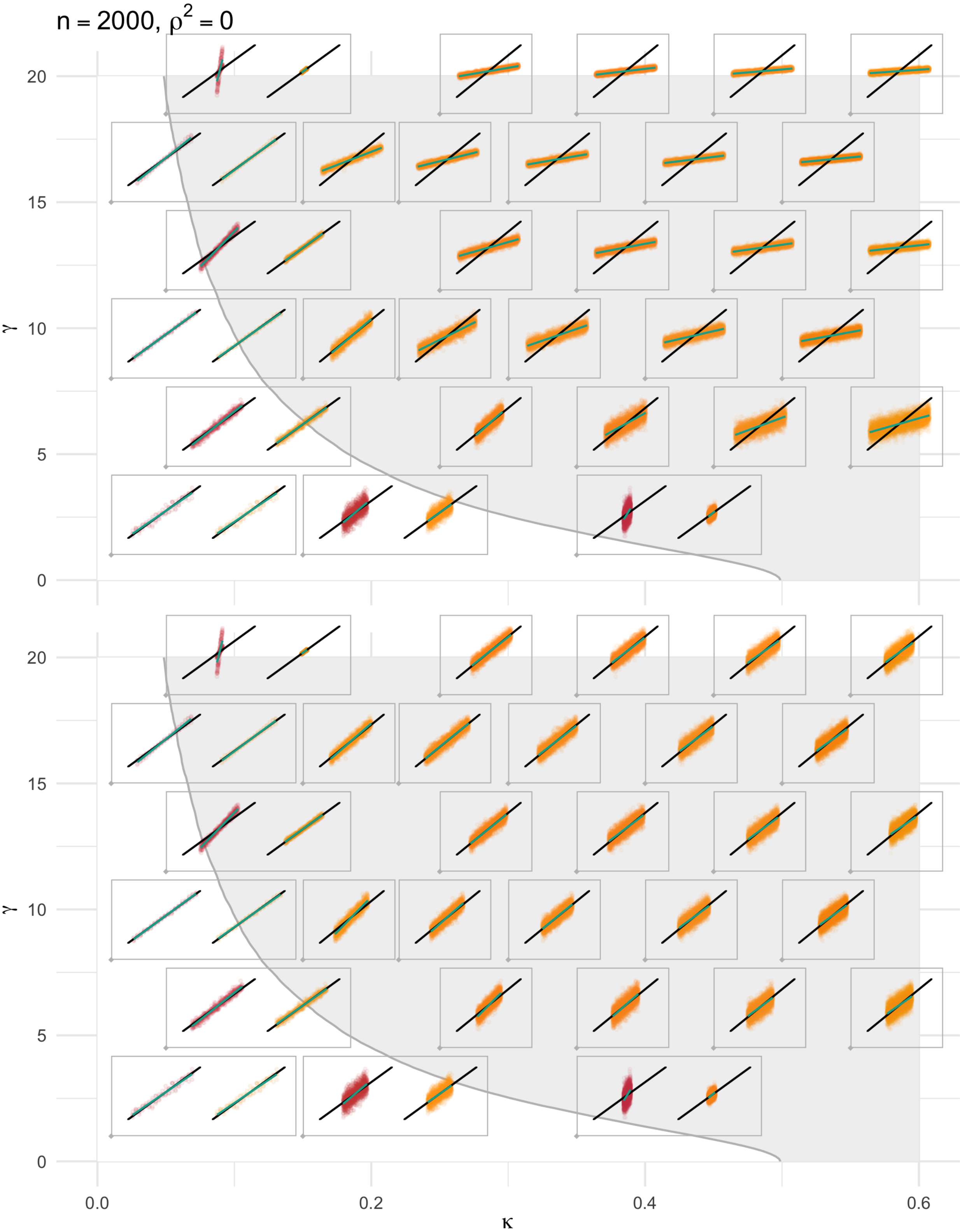


Figure S8: Setting a with $n = 2000$ and $\rho^2 = 0.25$.

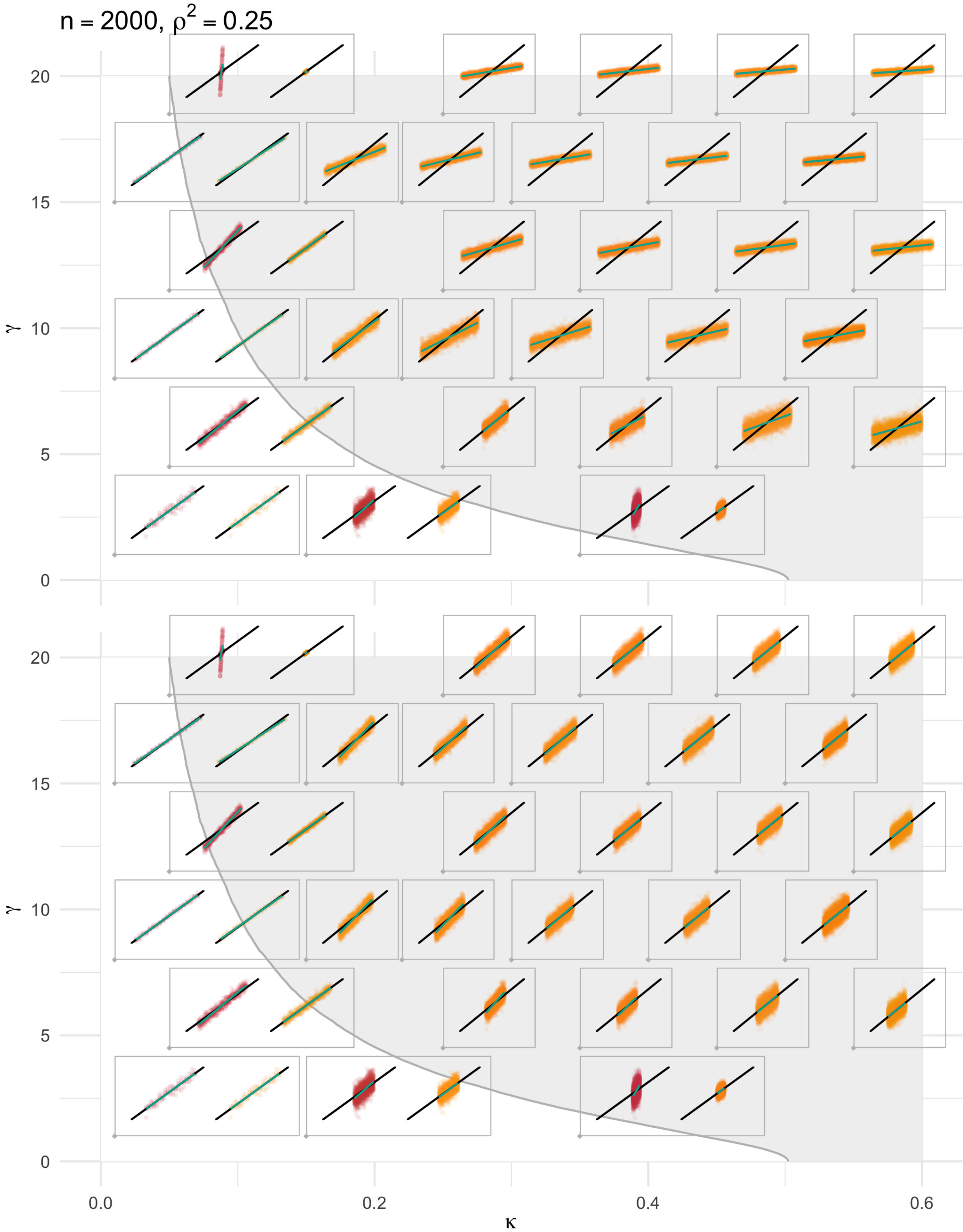


Figure S9: Setting a with $n = 2000$ and $\rho^2 = 0.5$.

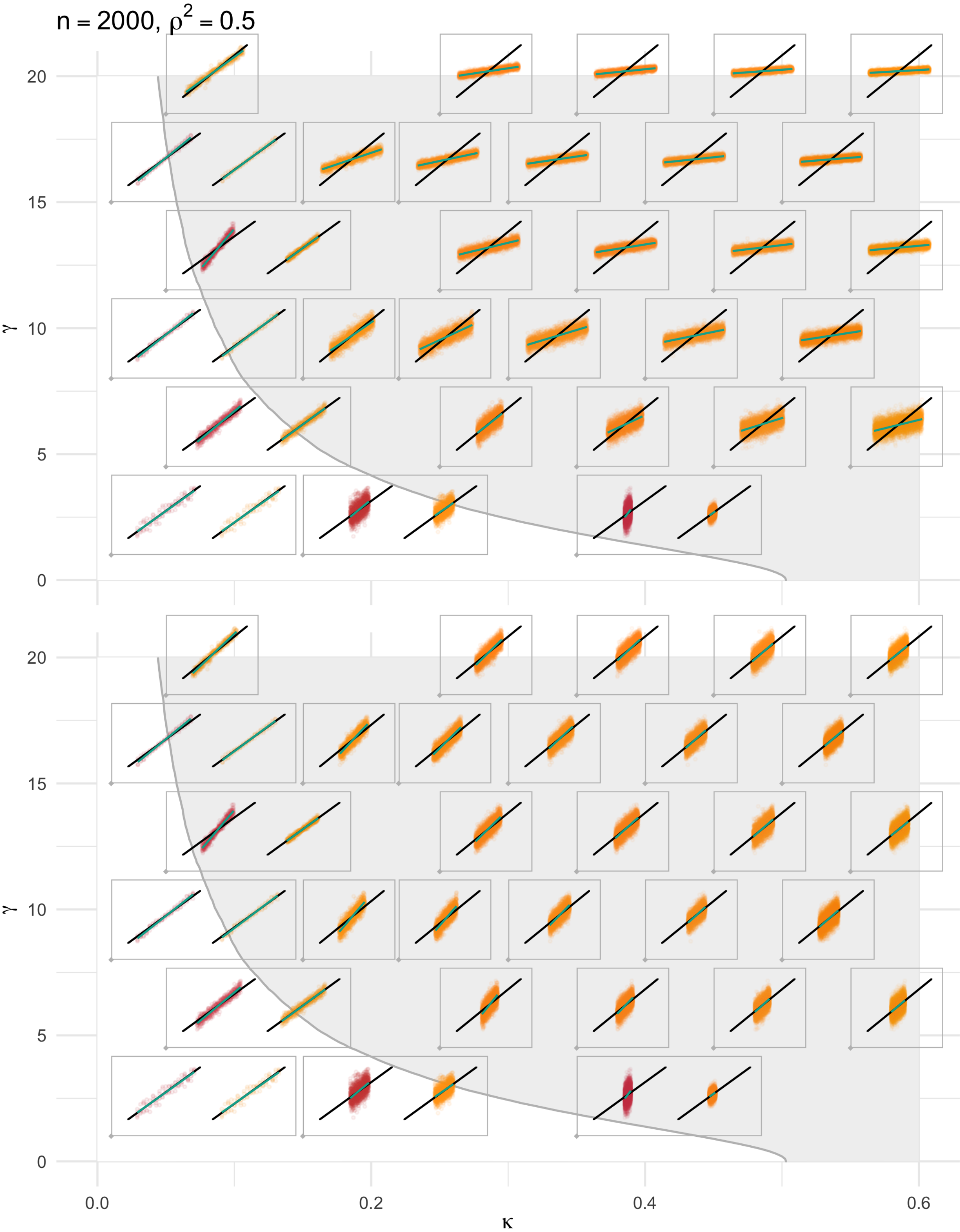


Figure S10: Setting a with $n = 2000$ and $\rho^2 = 0.75$.

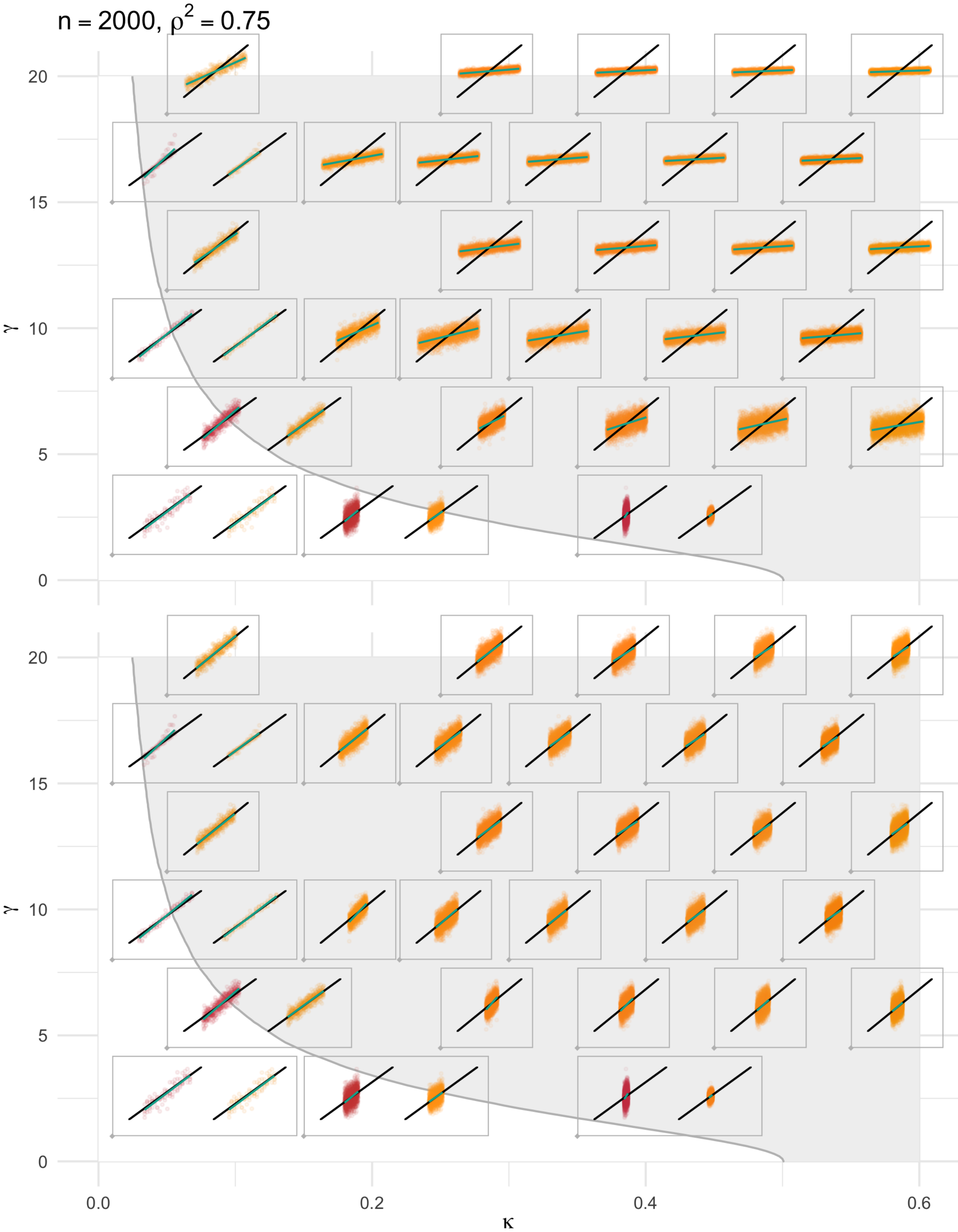


Figure S11: Setting a with $n = 2000$ and $\rho^2 = 0.8$.

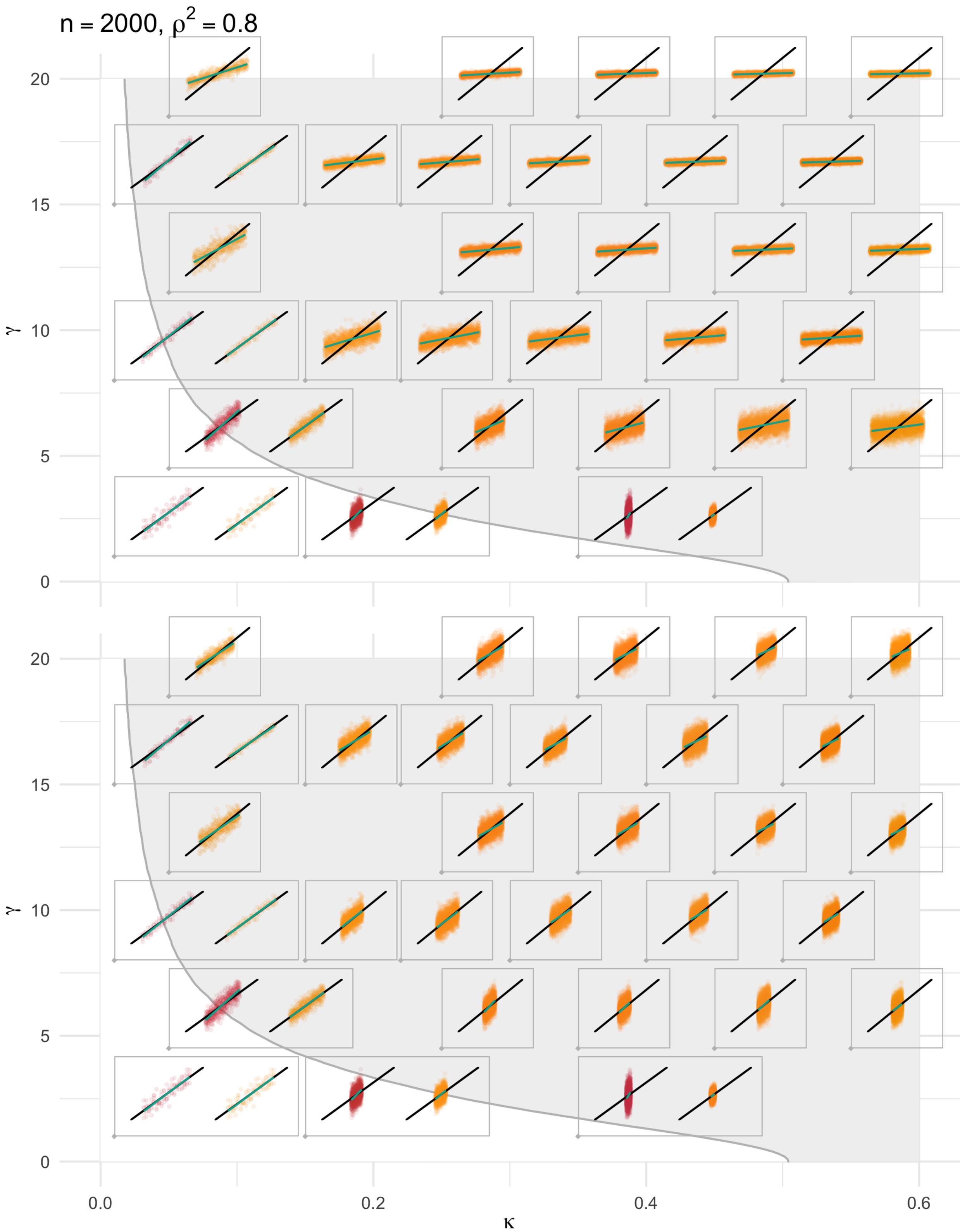


Figure S12: Setting a with $n = 2000$ and $\rho^2 = 0.9$.

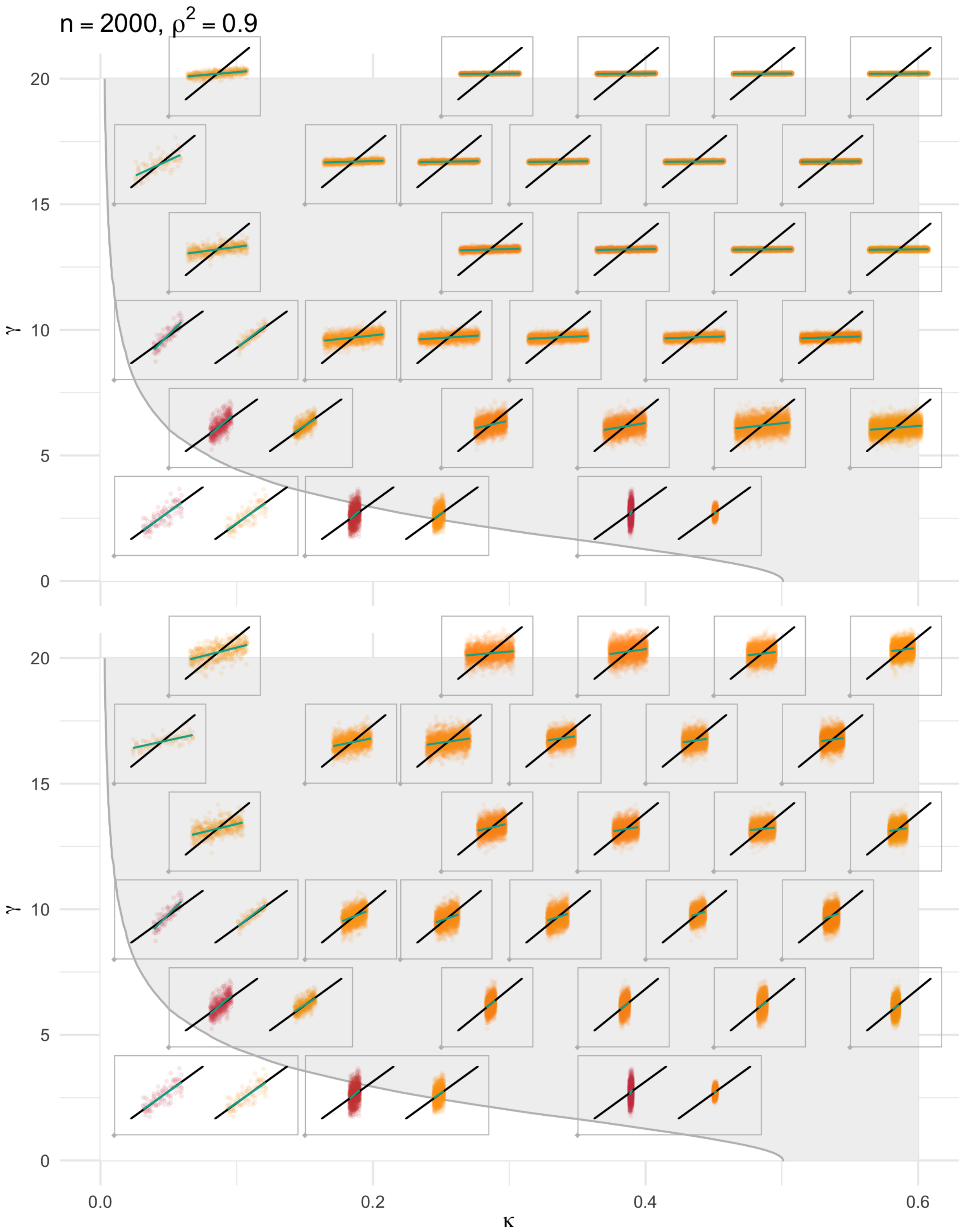


Figure S13: Setting a with $n = 3000$ and $\rho^2 = 0$.

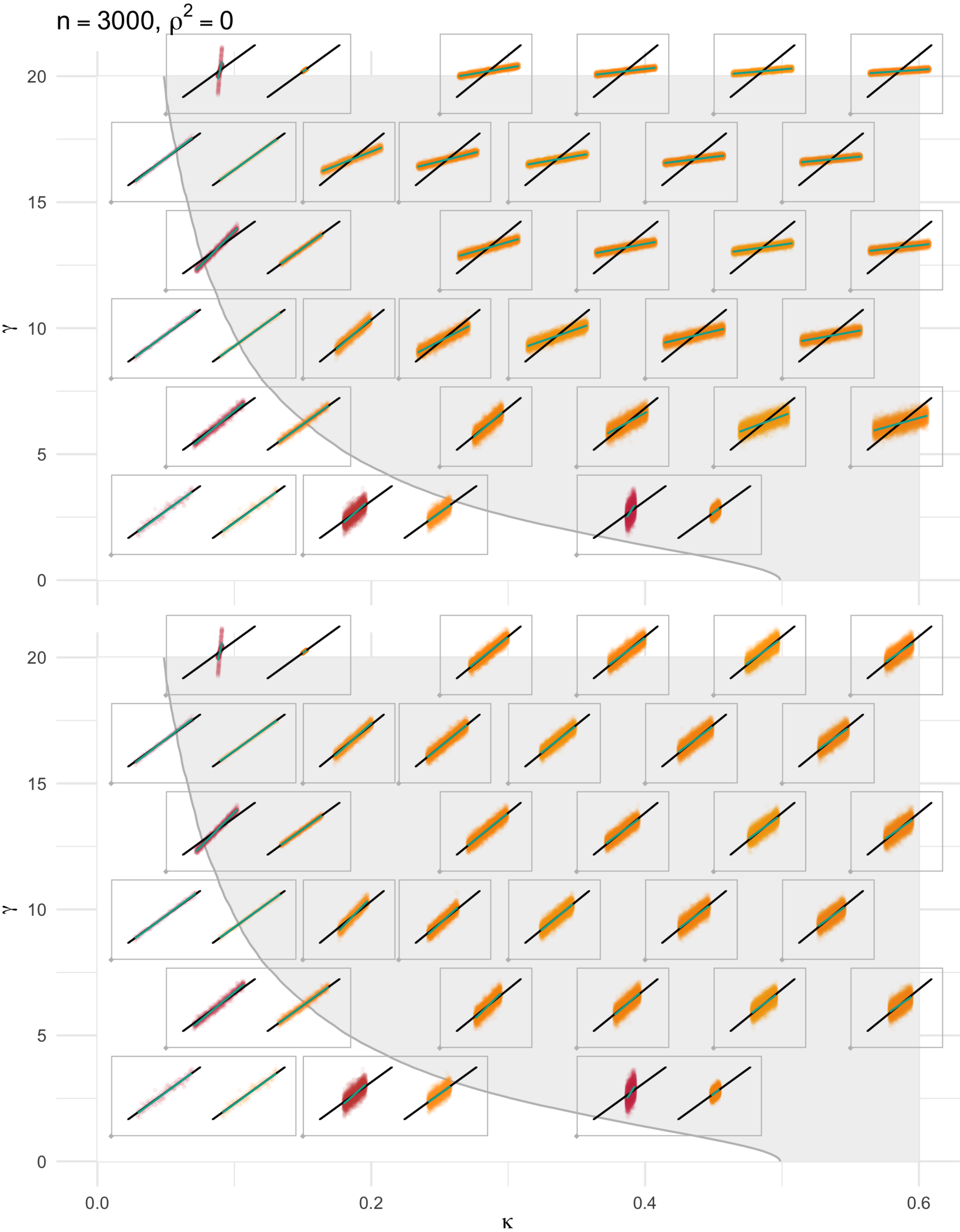


Figure S14: Setting a with $n = 3000$ and $\rho^2 = 0.25$.

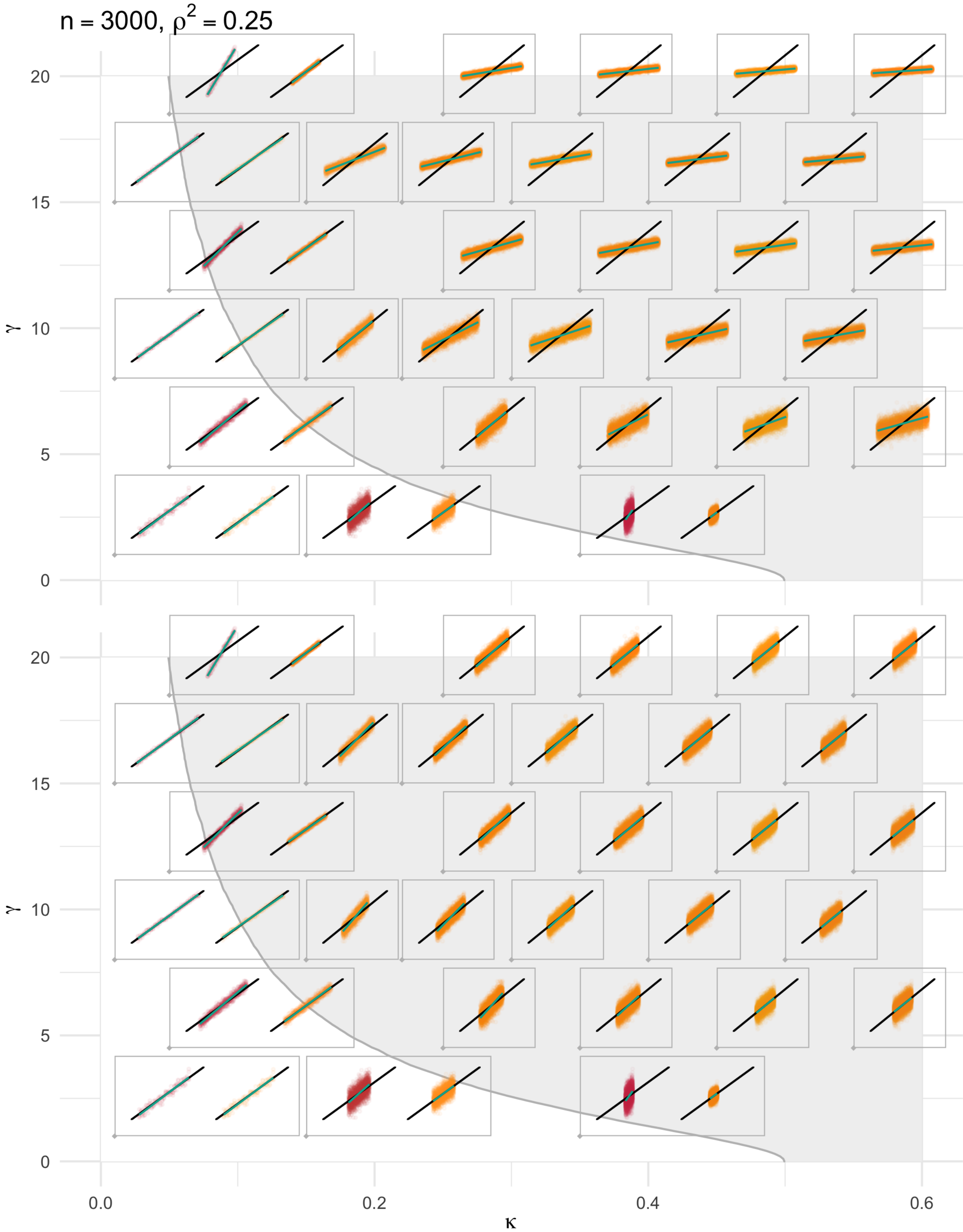


Figure S15: Setting a with $n = 3000$ and $\rho^2 = 0.5$.

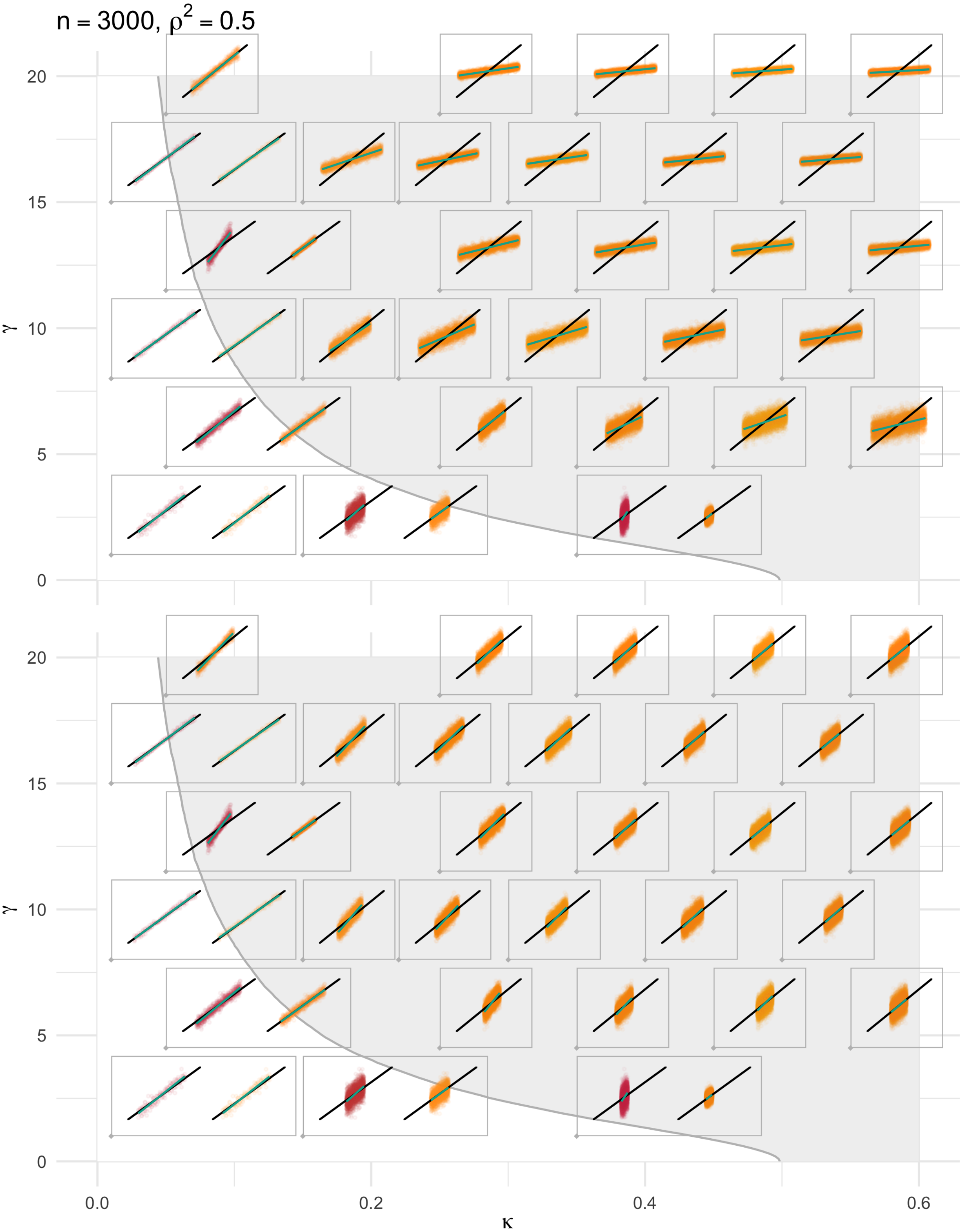


Figure S16: Setting a with $n = 3000$ and $\rho^2 = 0.75$.

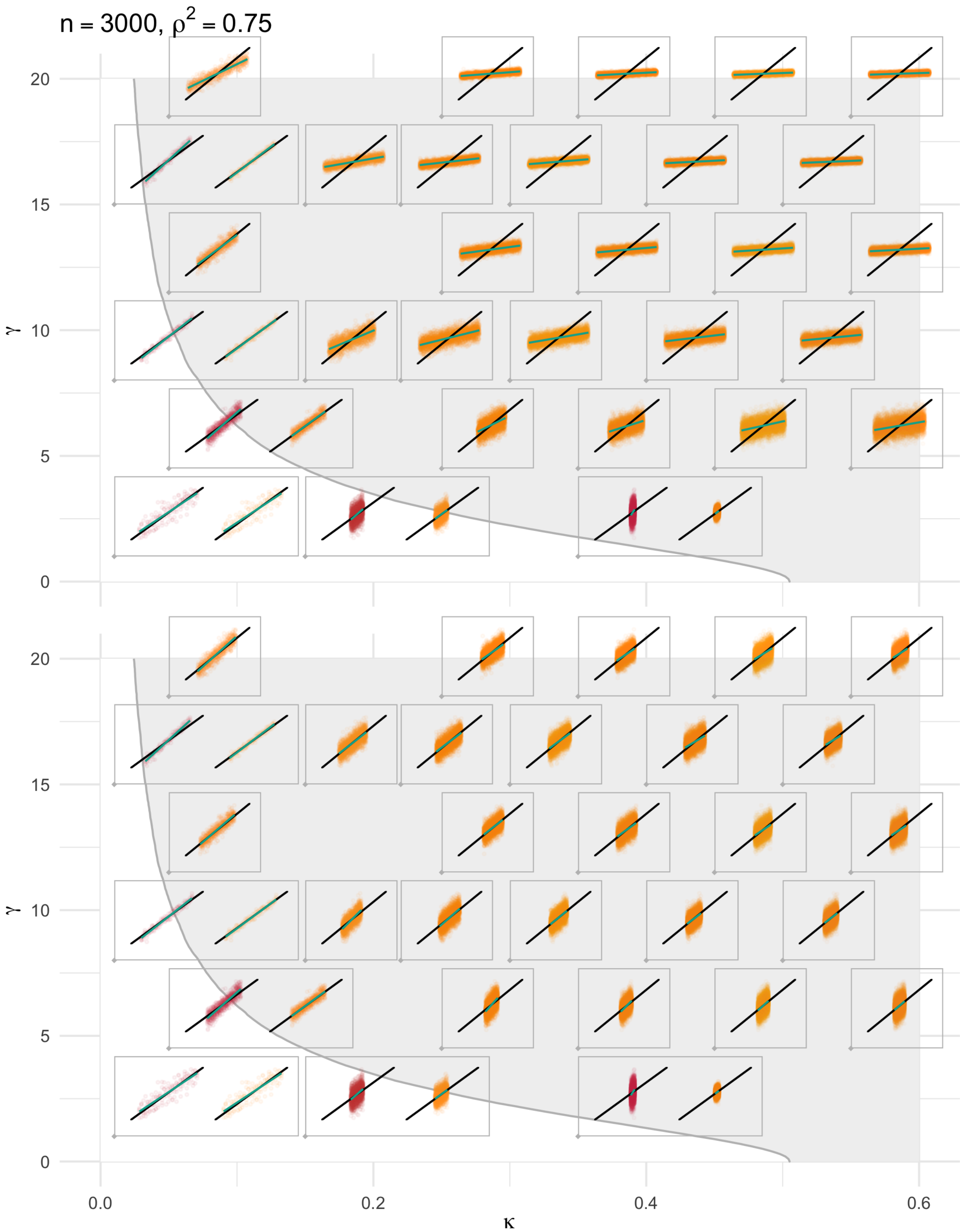


Figure S17: Setting a with $n = 3000$ and $\rho^2 = 0.8$.

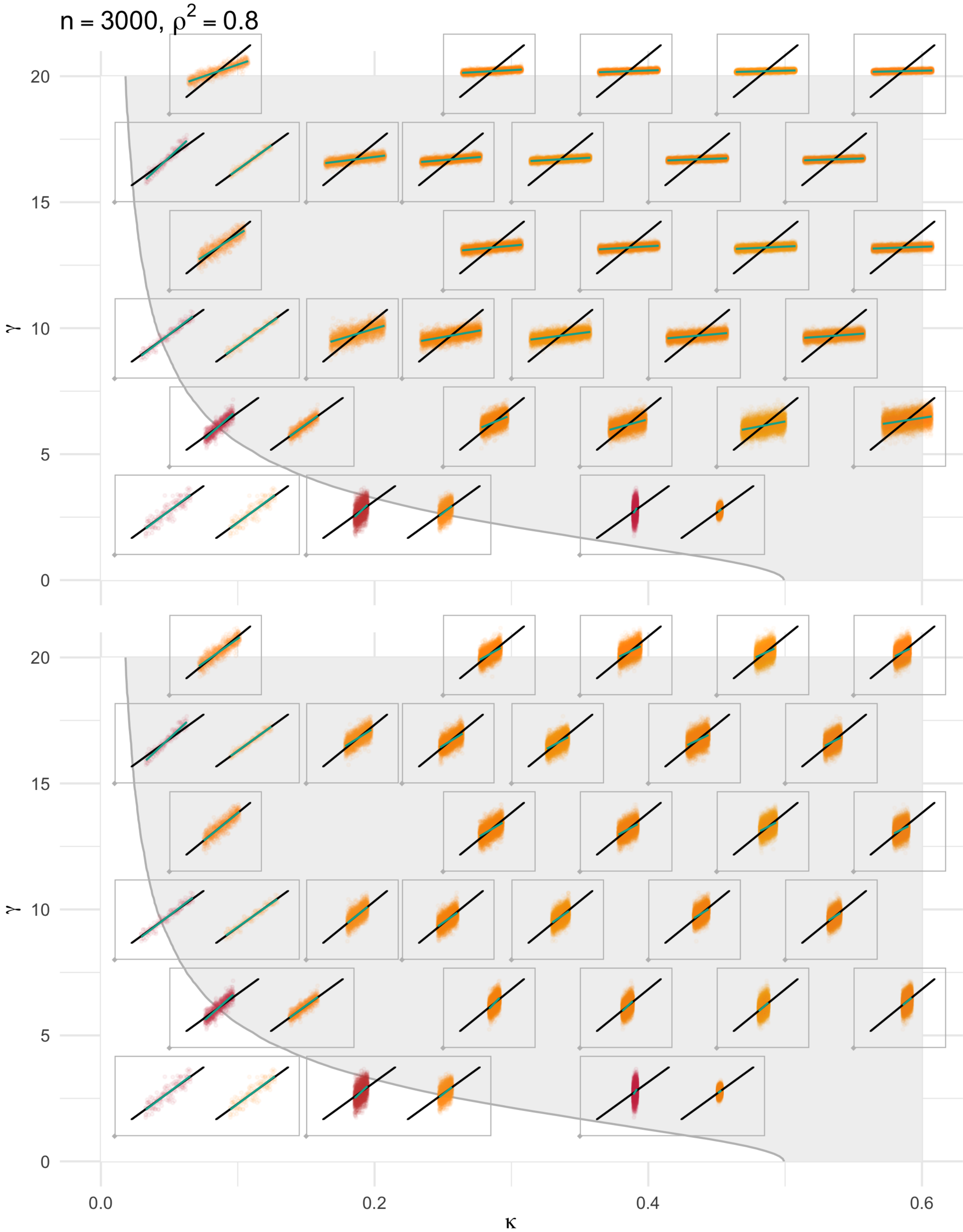
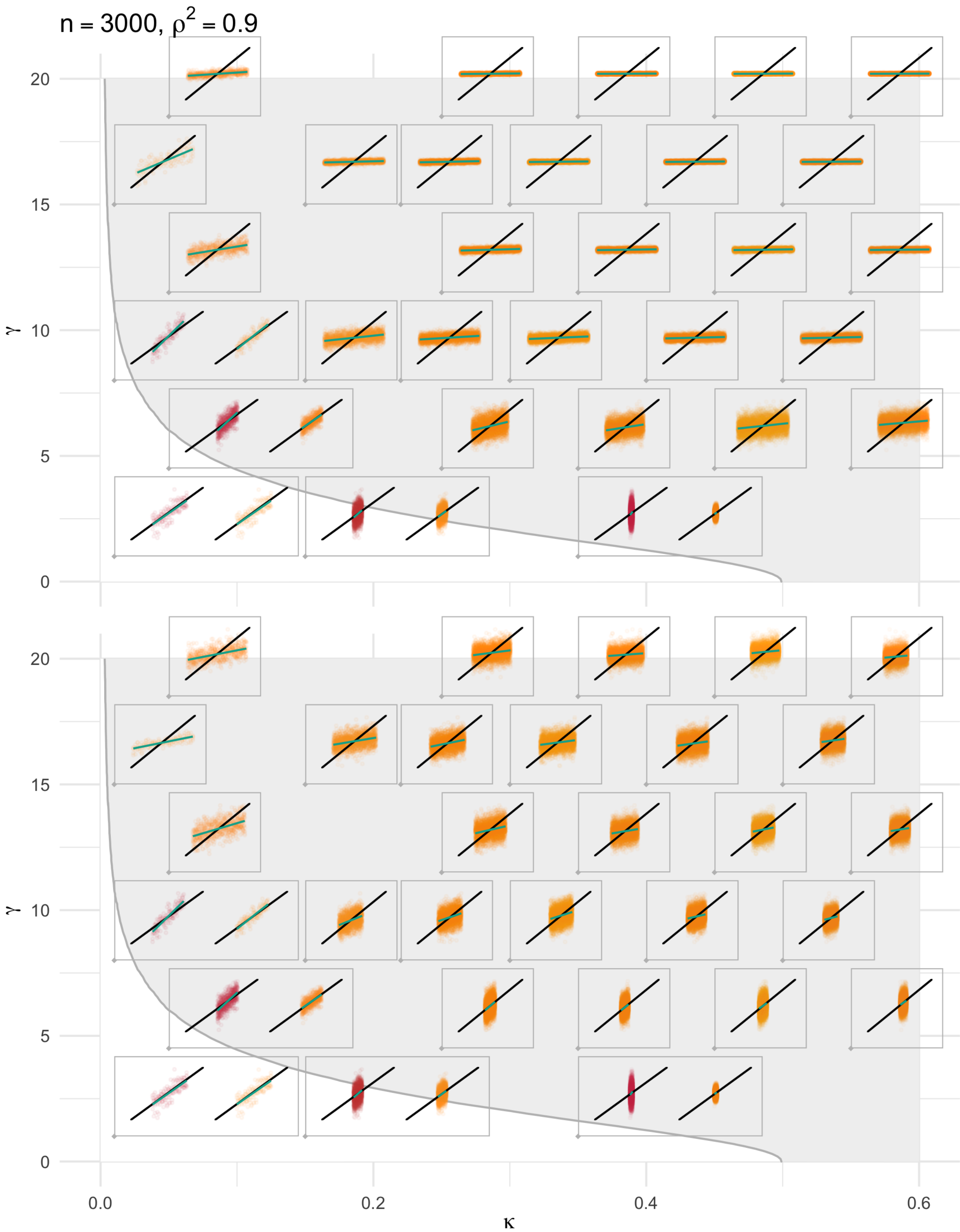


Figure S18: Setting a with $n = 3000$ and $\rho^2 = 0.9$.



S3 Simulation experiment in Section 5: Setting b

Figures S19-S36 show the phase transition curves derived using Candès and Sur (2020, Theorem 2.1) for $\rho^2 \in \{0, 0.25, 0.5, 0.75, 0.8, 0.9\}$ for the computer experiment in Section 5, overlaid with scatterplots of estimates of β from 5 independent realizations of y and X vs the true β, at 30 (κ, γ) points, for $n \in \{1000, 2000, 3000\}$, and with β^* set to have length $p = \lceil n\kappa \rceil$ with 20% of its values set to -10, 20% of its values set to 10, and the remaining set to 0. The top figure shows the ML estimator $\hat{\beta}$ (red) and mJPL estimator $\tilde{\beta}$ (orange), and the bottom figure shows $\hat{\beta}$ (red) and the estimator $\beta^{\dagger} = q(\kappa, \gamma, \rho)\tilde{\beta}$ (orange) with $q(\kappa, \gamma, \rho)$ as in (13). The black lines are intercept zero and slope one reference lines, and the green lines are estimated from a simple linear regression of the estimates on the truth.

Figure S19: Setting b with $n = 1000$ and $\rho^2 = 0$.

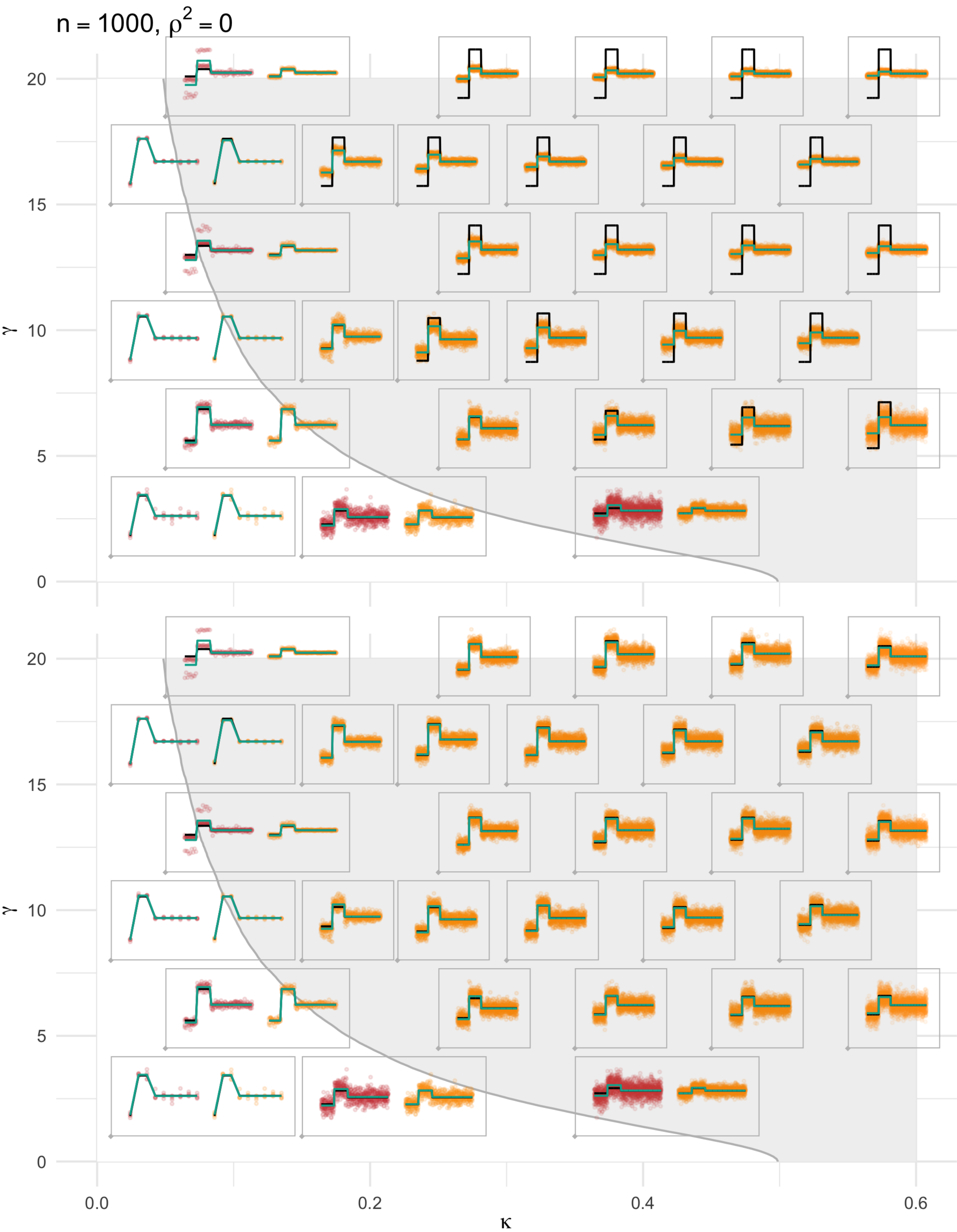


Figure S20: Setting b with $n = 1000$ and $\rho^2 = 0.25$.

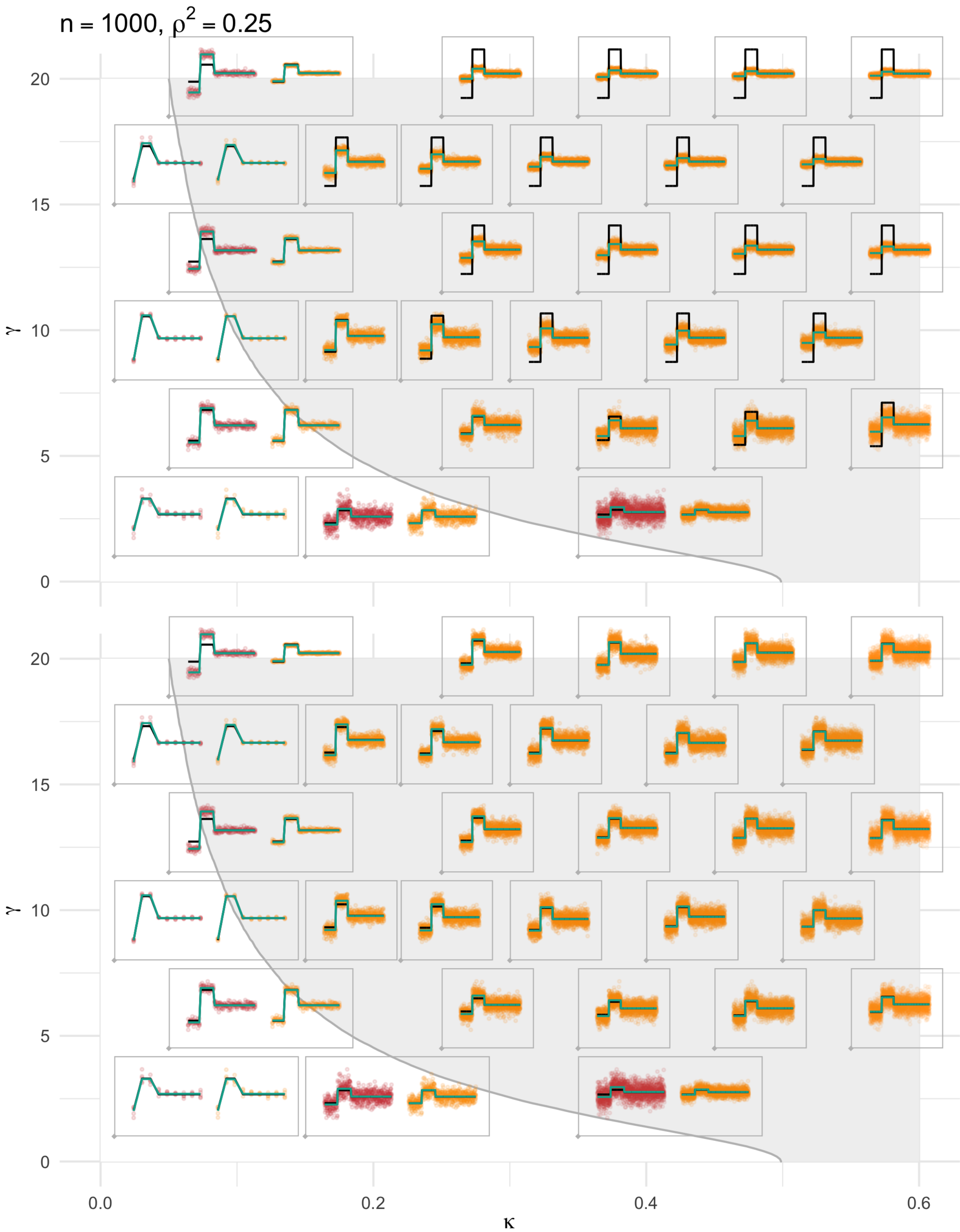


Figure S21: Setting b with $n = 1000$ and $\rho^2 = 0.5$.

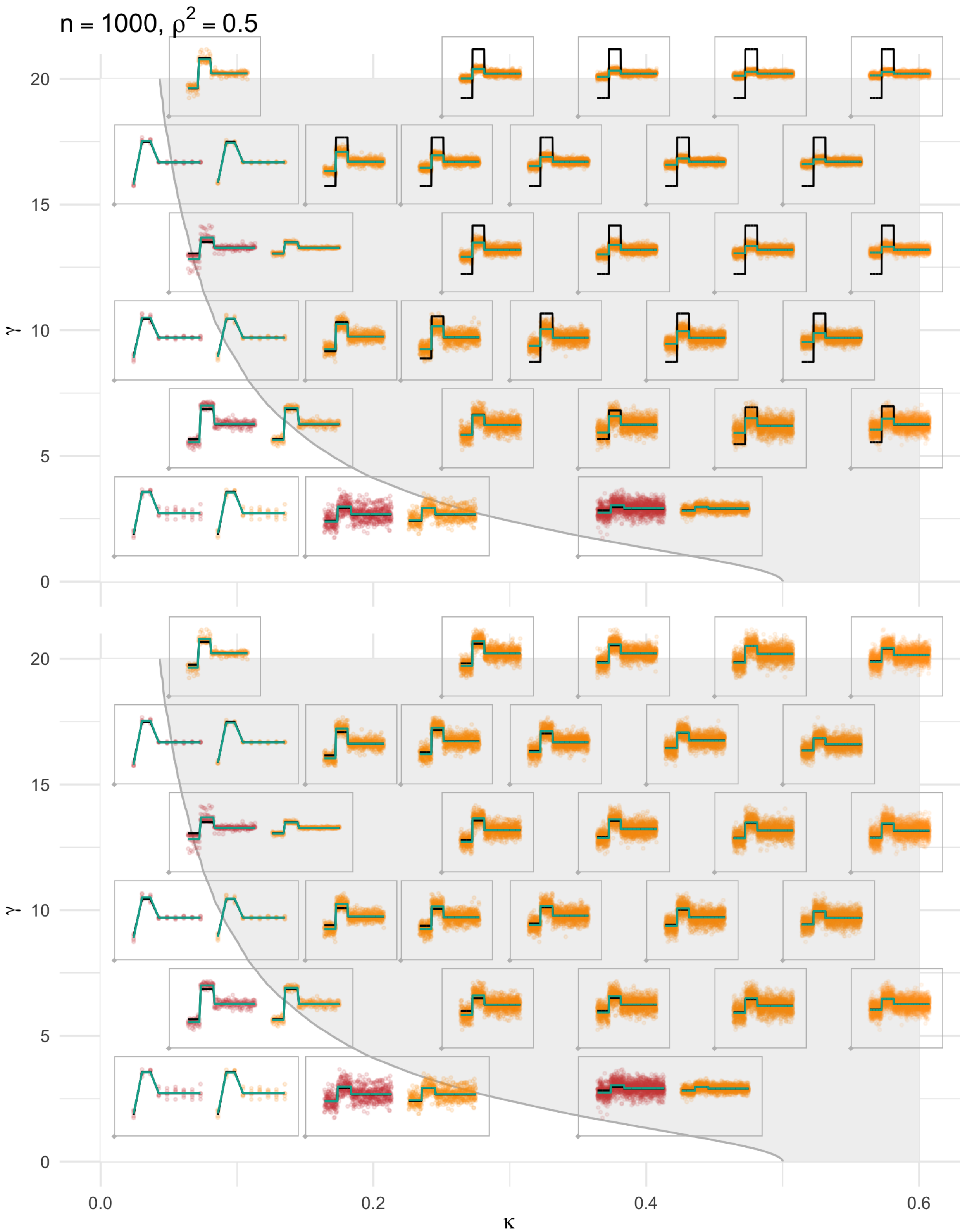


Figure S22: Setting b with $n = 1000$ and $\rho^2 = 0.75$.

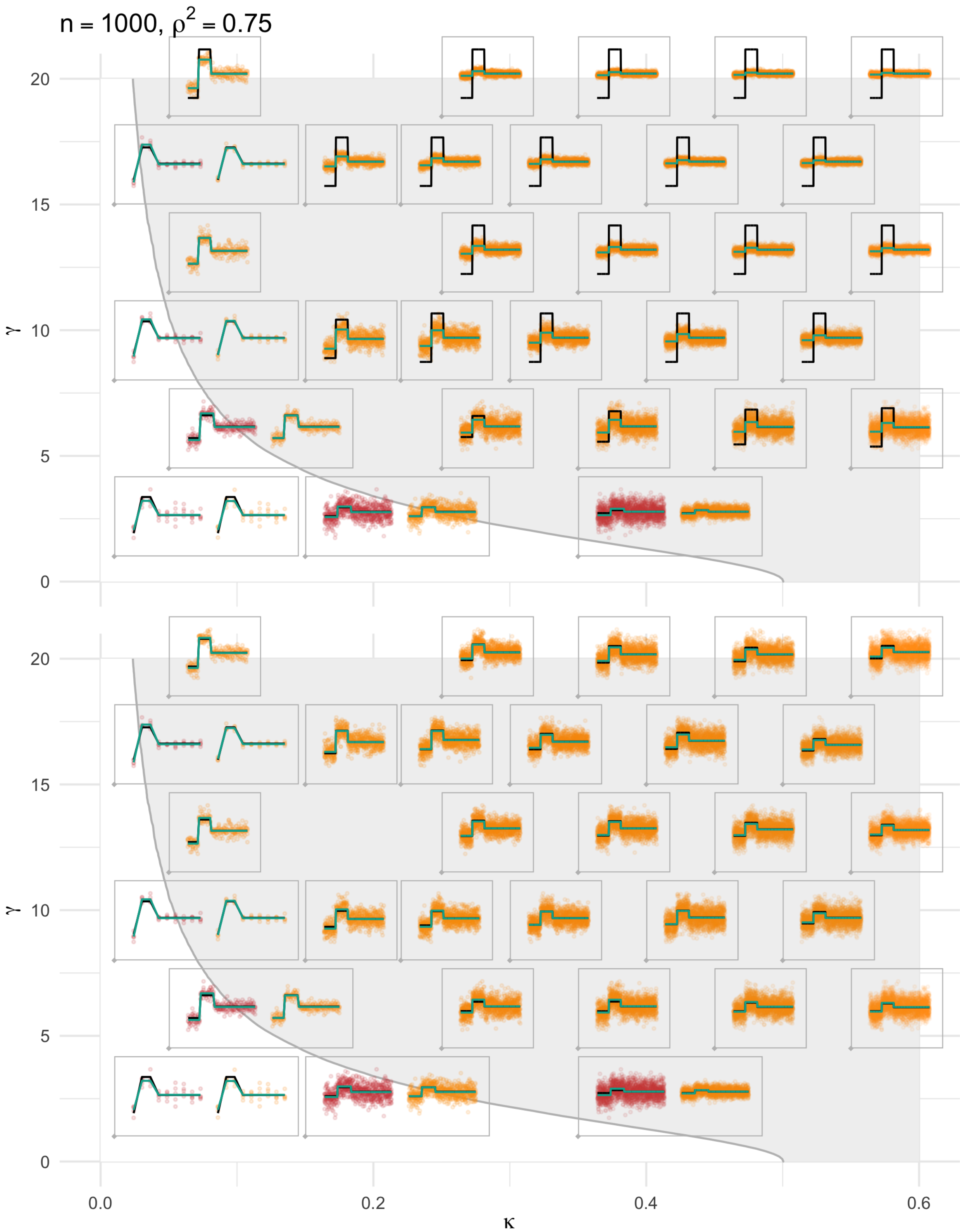


Figure S23: Setting b with $n = 1000$ and $\rho^2 = 0.8$.

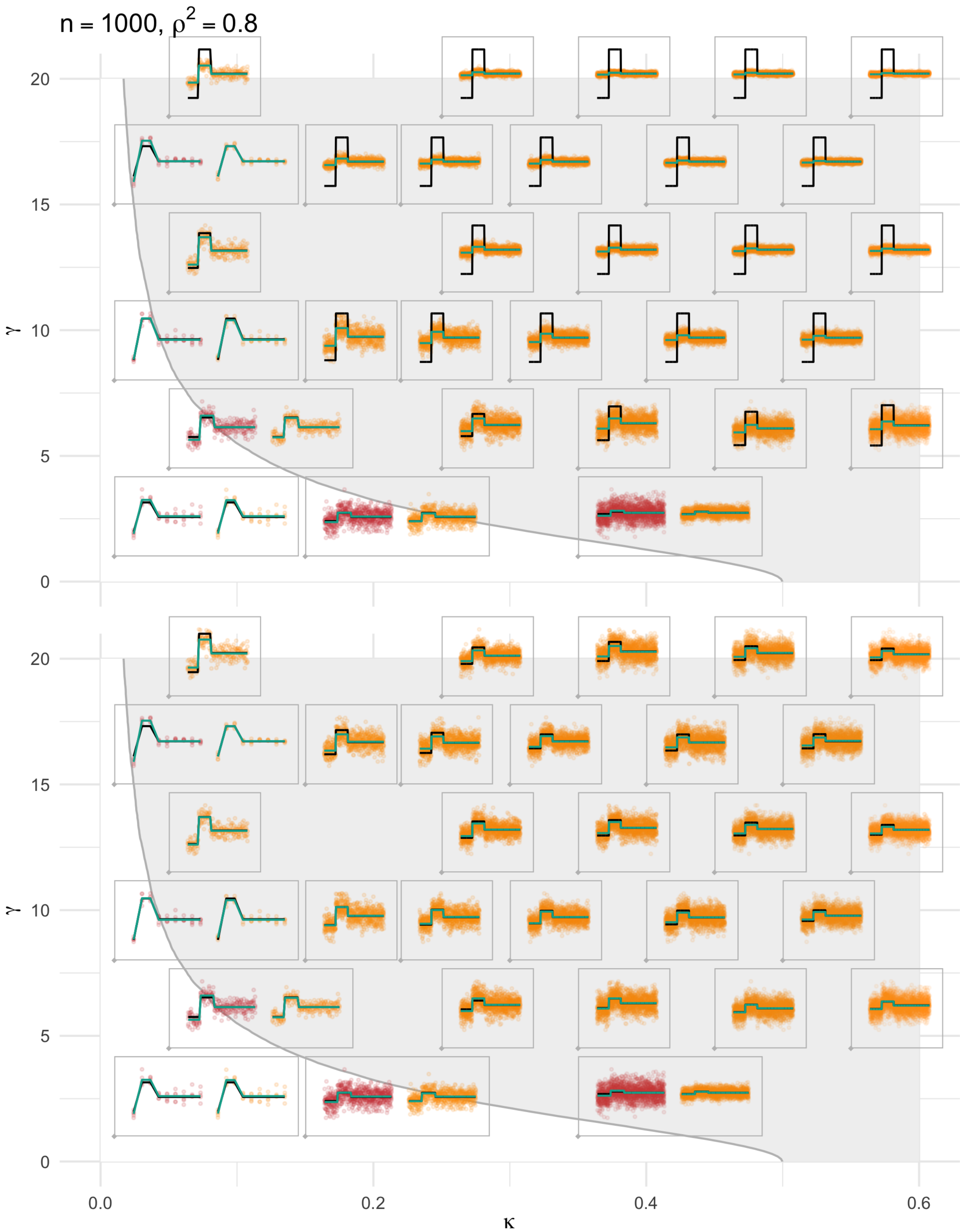


Figure S24: Setting b with $n = 1000$ and $\rho^2 = 0.9$.

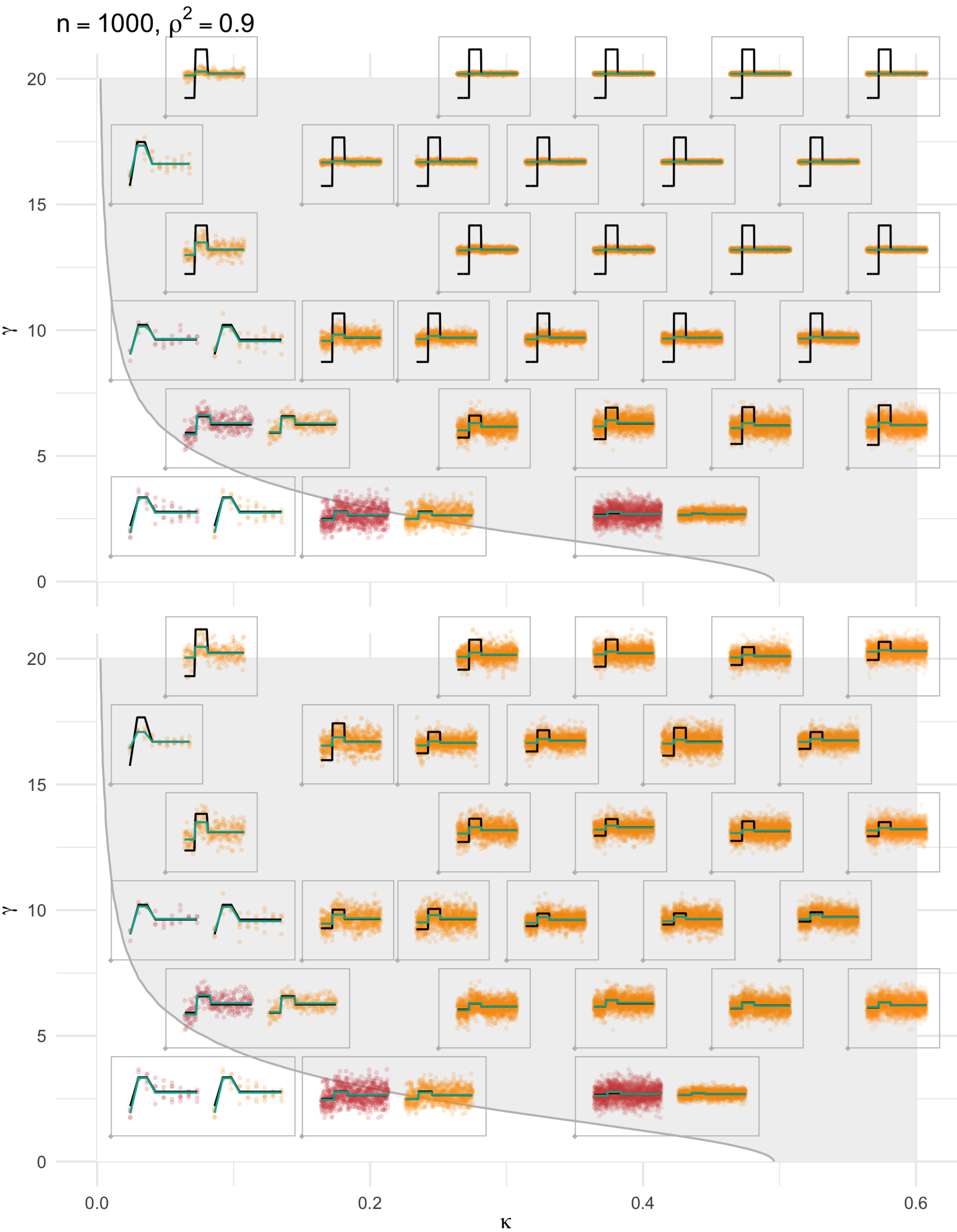


Figure S25: Setting b with $n = 2000$ and $\rho^2 = 0$.

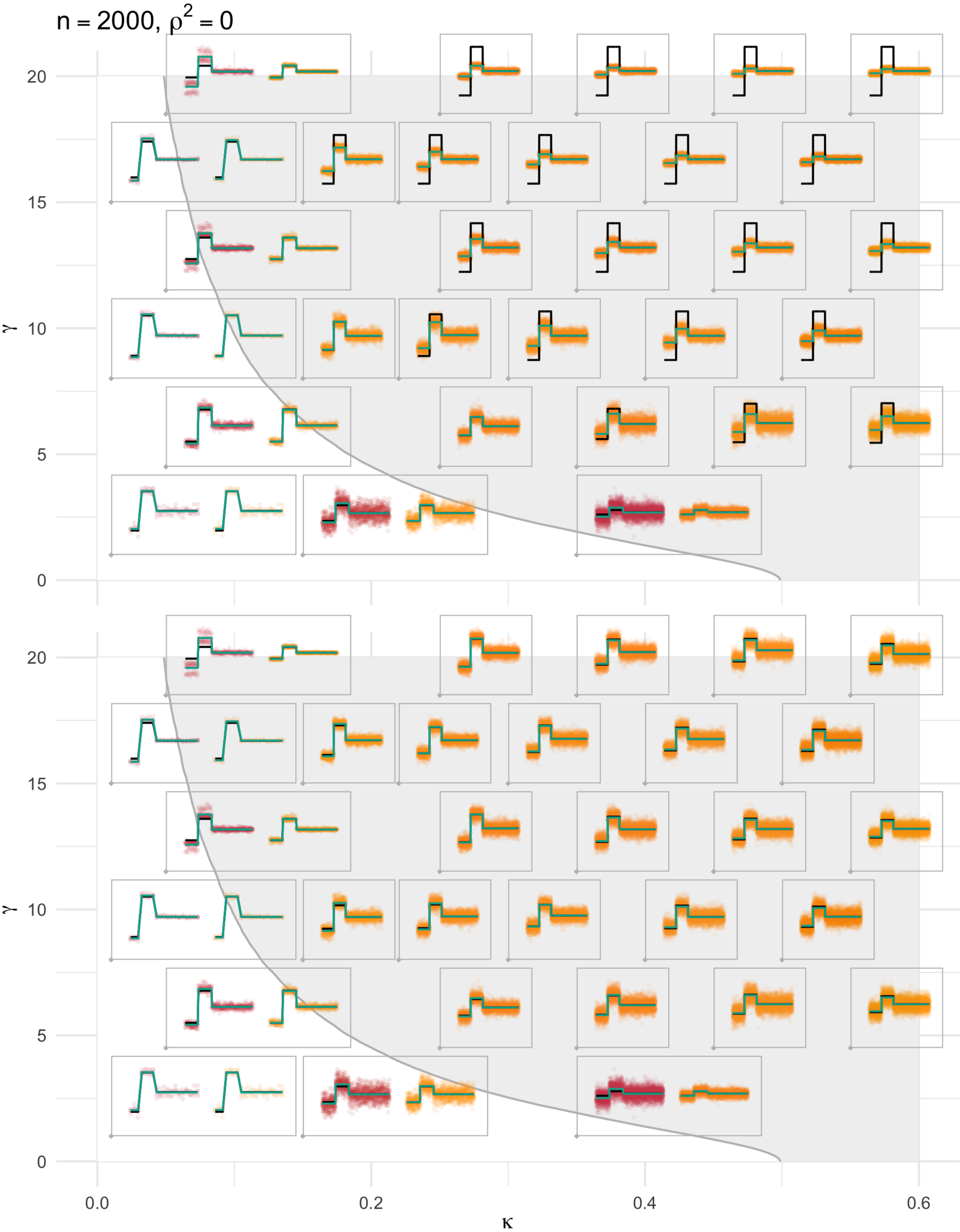


Figure S26: Setting b with $n = 2000$ and $\rho^2 = 0.25$.

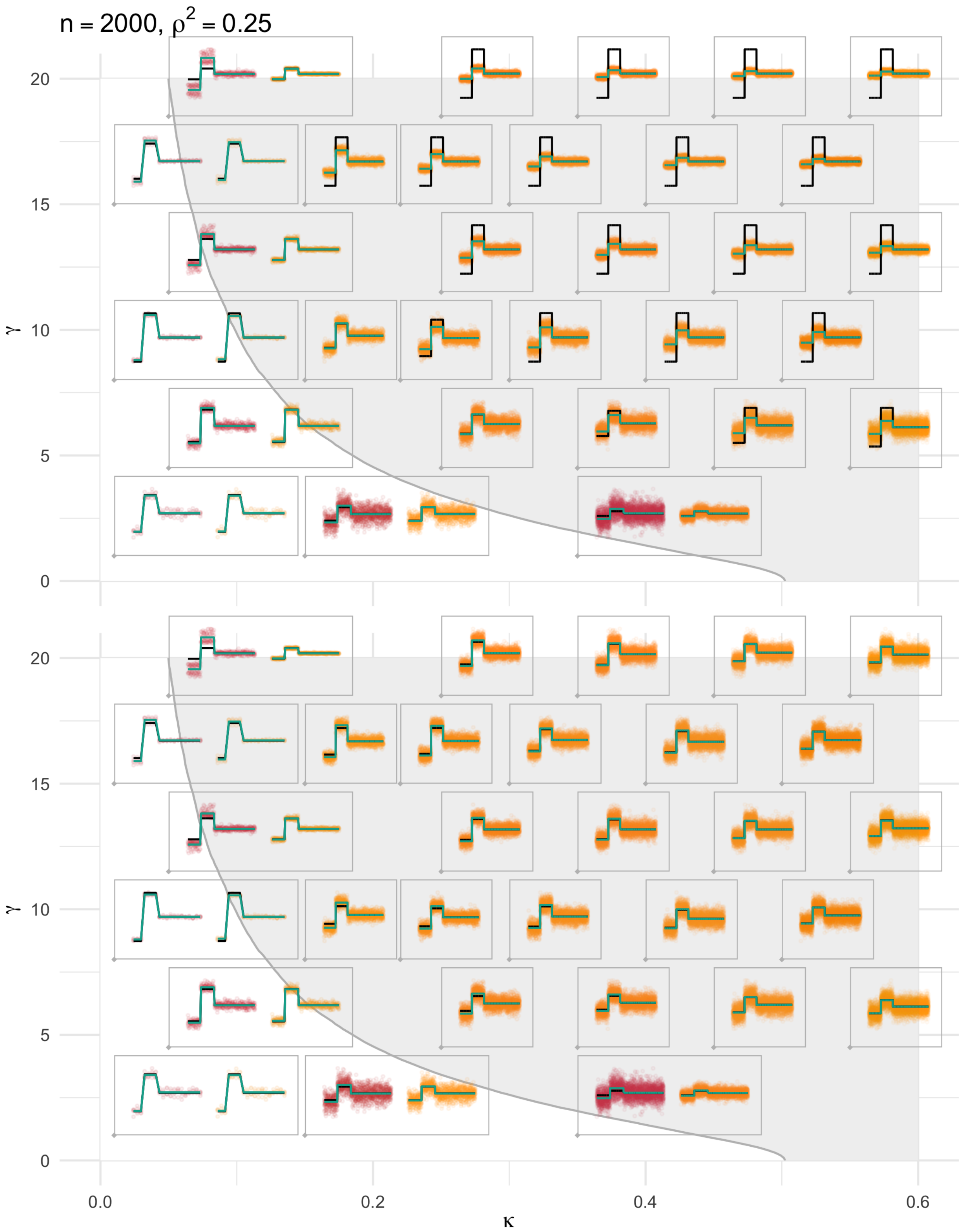


Figure S27: Setting b with $n = 2000$ and $\rho^2 = 0.5$.

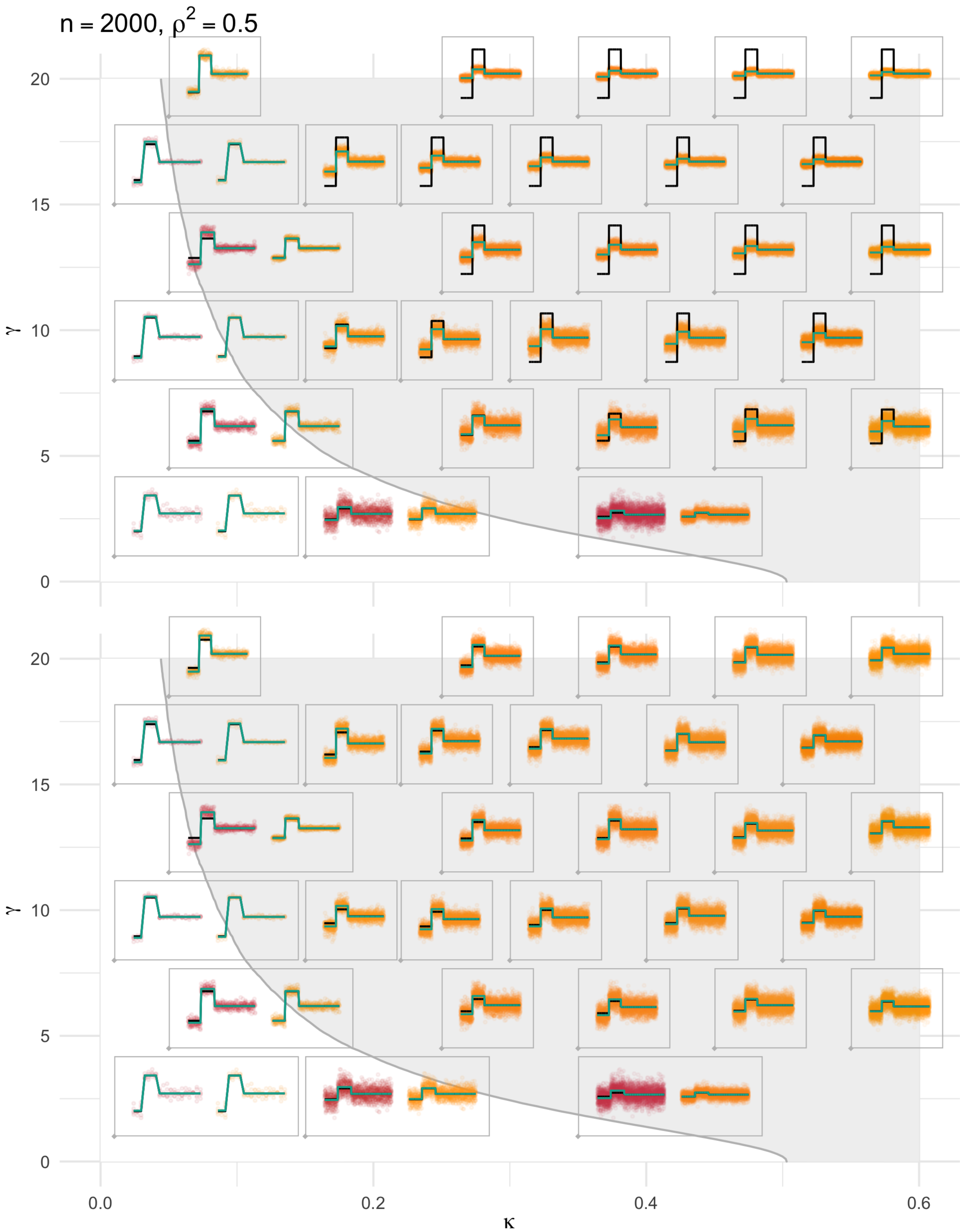


Figure S28: Setting b with $n = 2000$ and $\rho^2 = 0.75$.

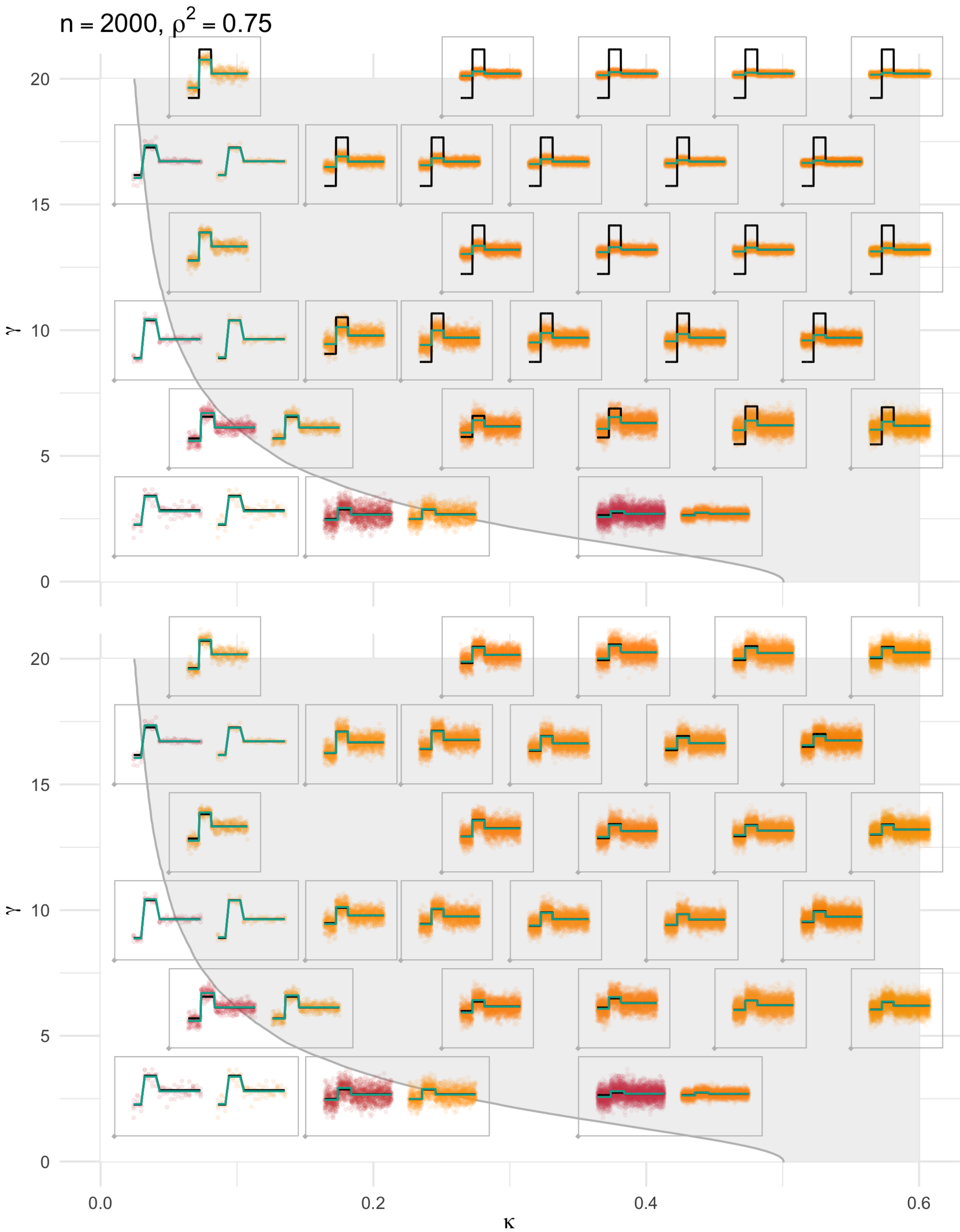


Figure S29: Setting b with $n = 2000$ and $\rho^2 = 0.8$.

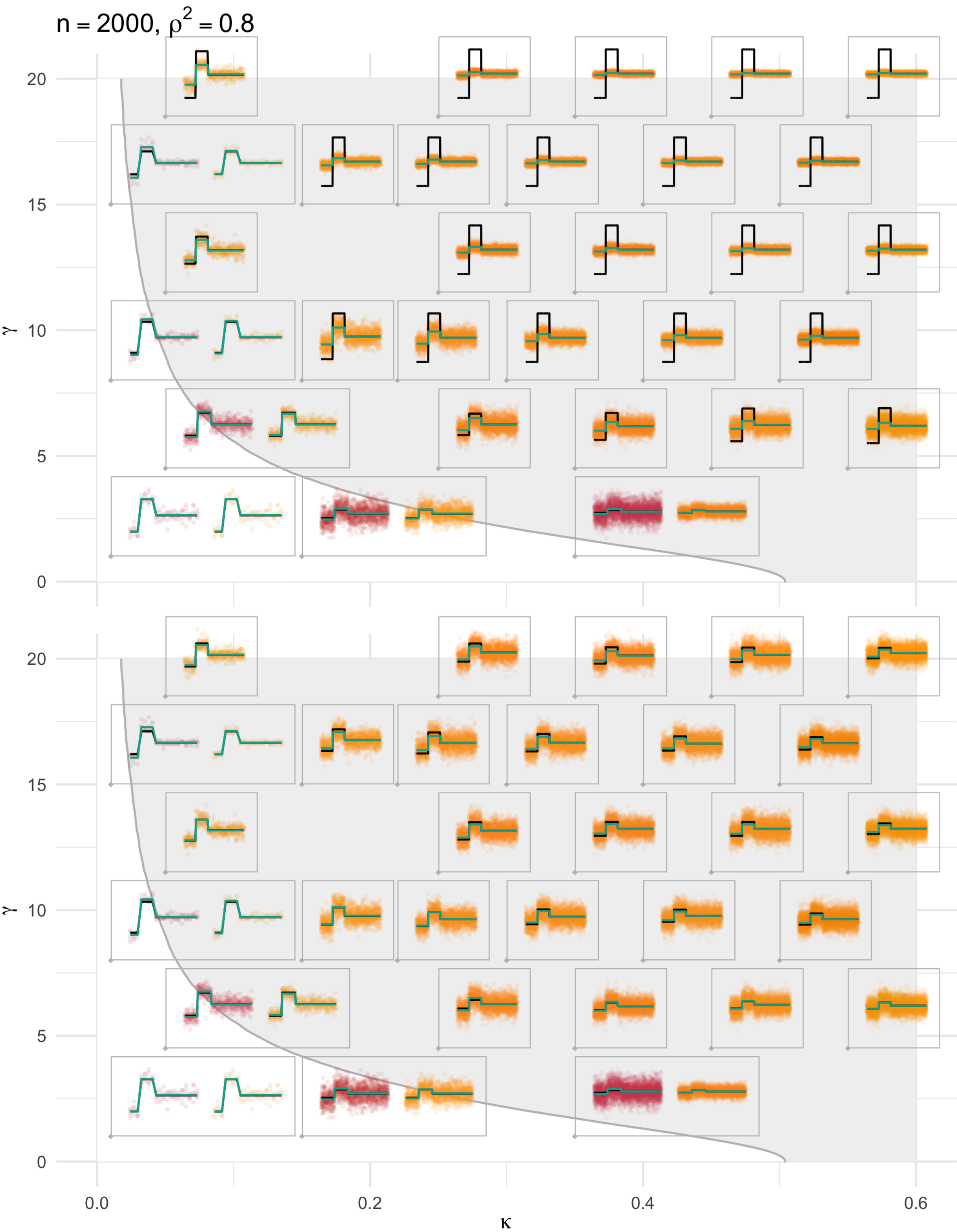


Figure S30: Setting b with $n = 2000$ and $\rho^2 = 0.9$.

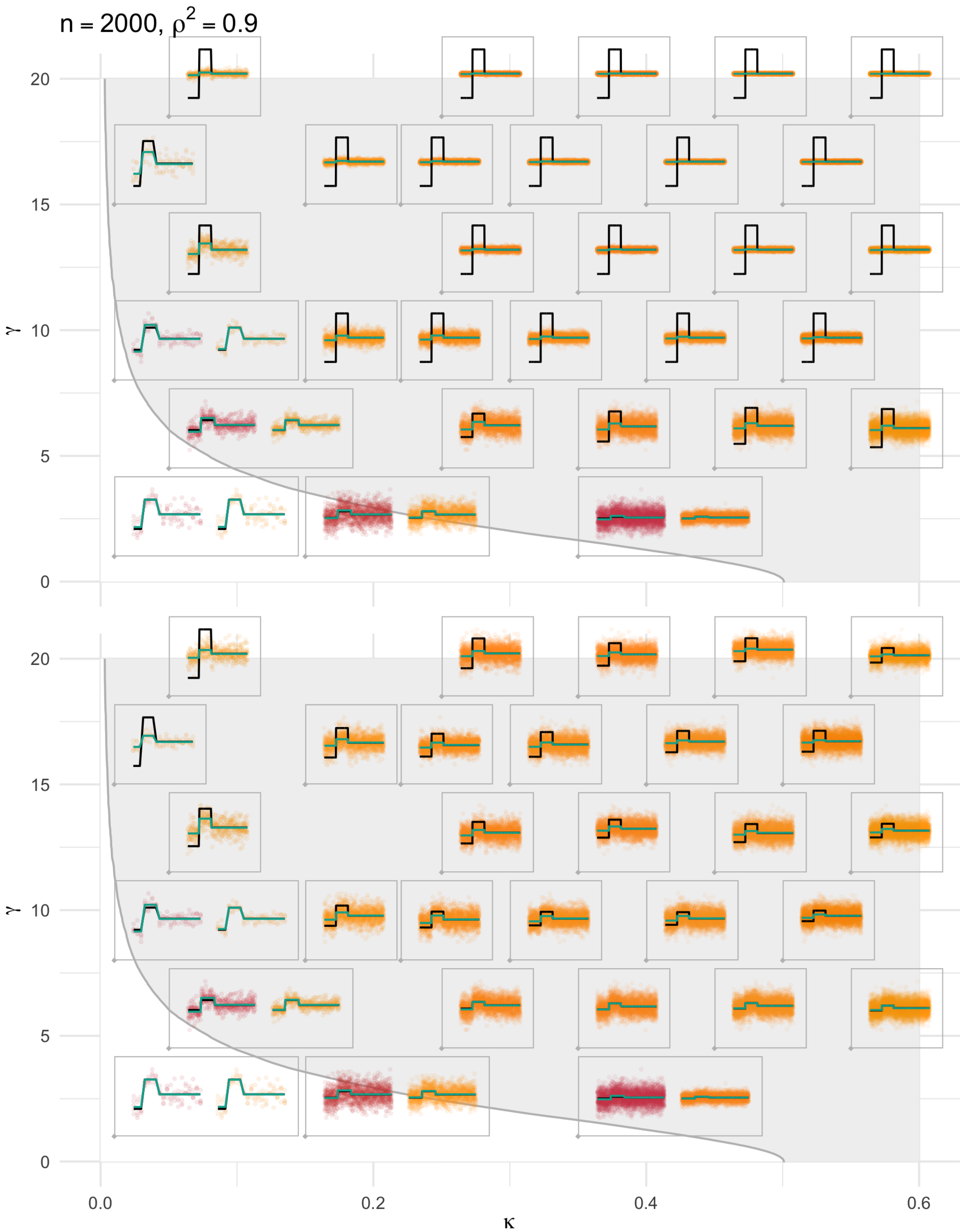


Figure S31: Setting b with $n = 3000$ and $\rho^2 = 0$.

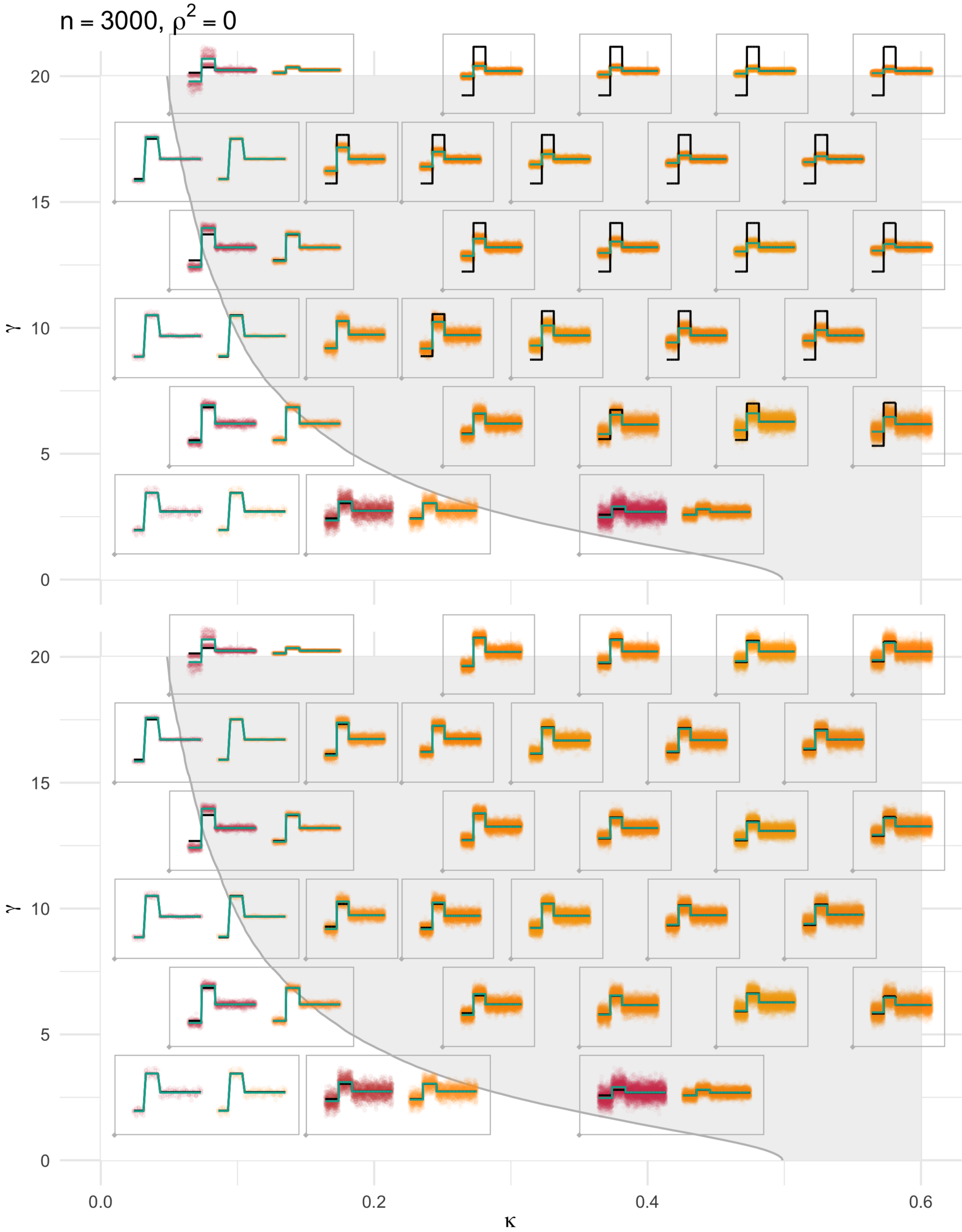


Figure S32: Setting b with $n = 3000$ and $\rho^2 = 0.25$.

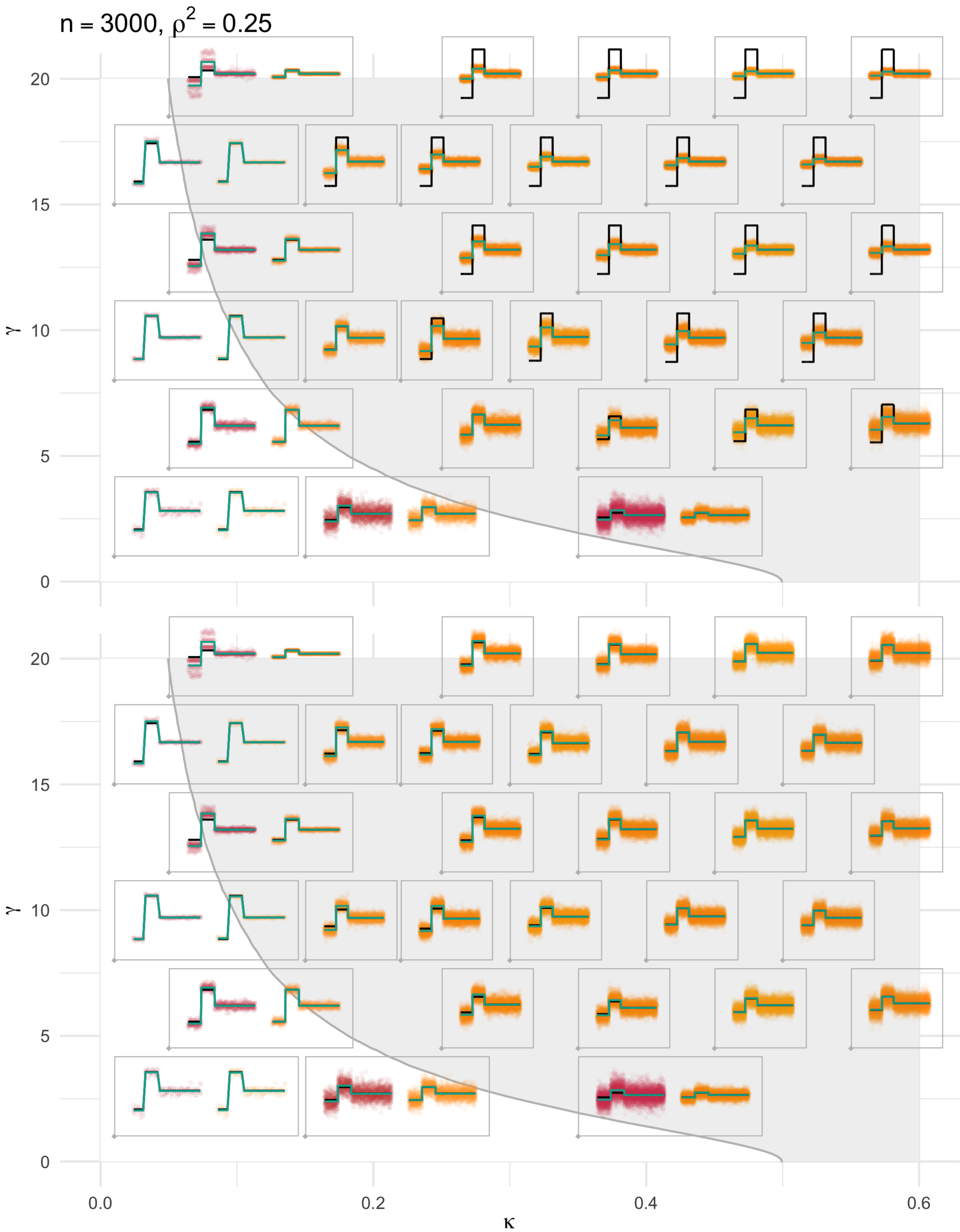


Figure S33: Setting b with $n = 3000$ and $\rho^2 = 0.5$.

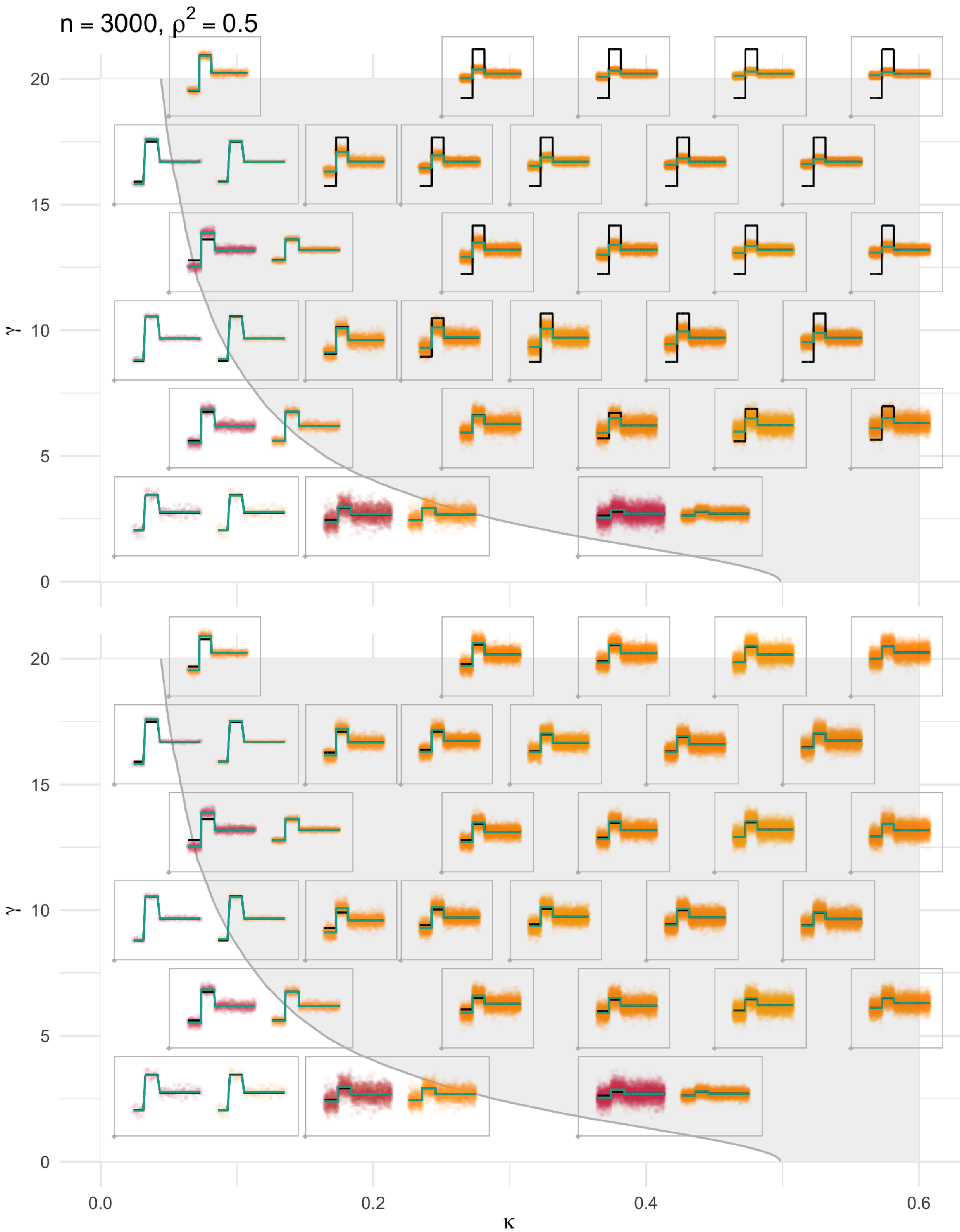


Figure S34: Setting b with $n = 3000$ and $\rho^2 = 0.75$.

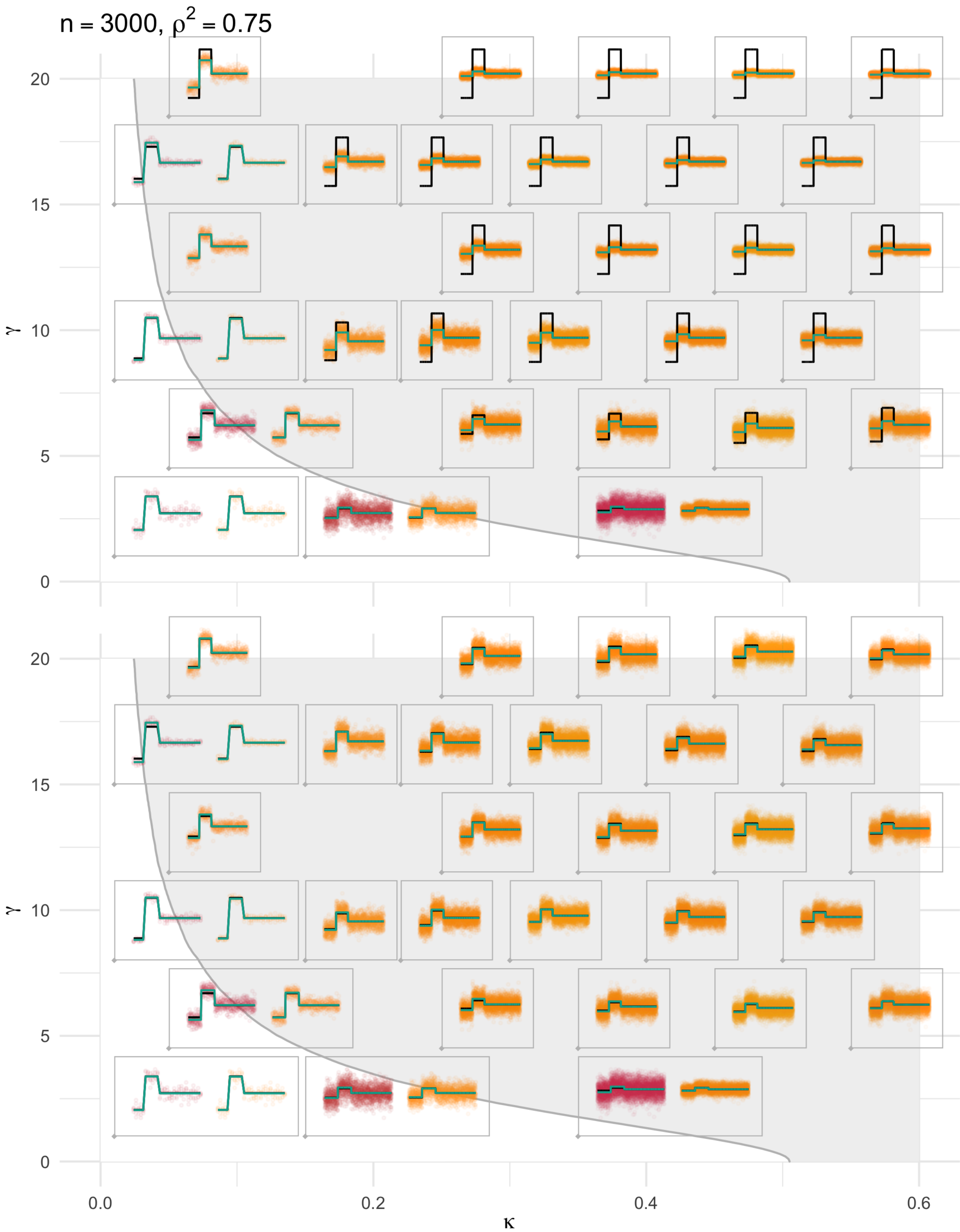


Figure S35: Setting b with $n = 3000$ and $\rho^2 = 0.8$.

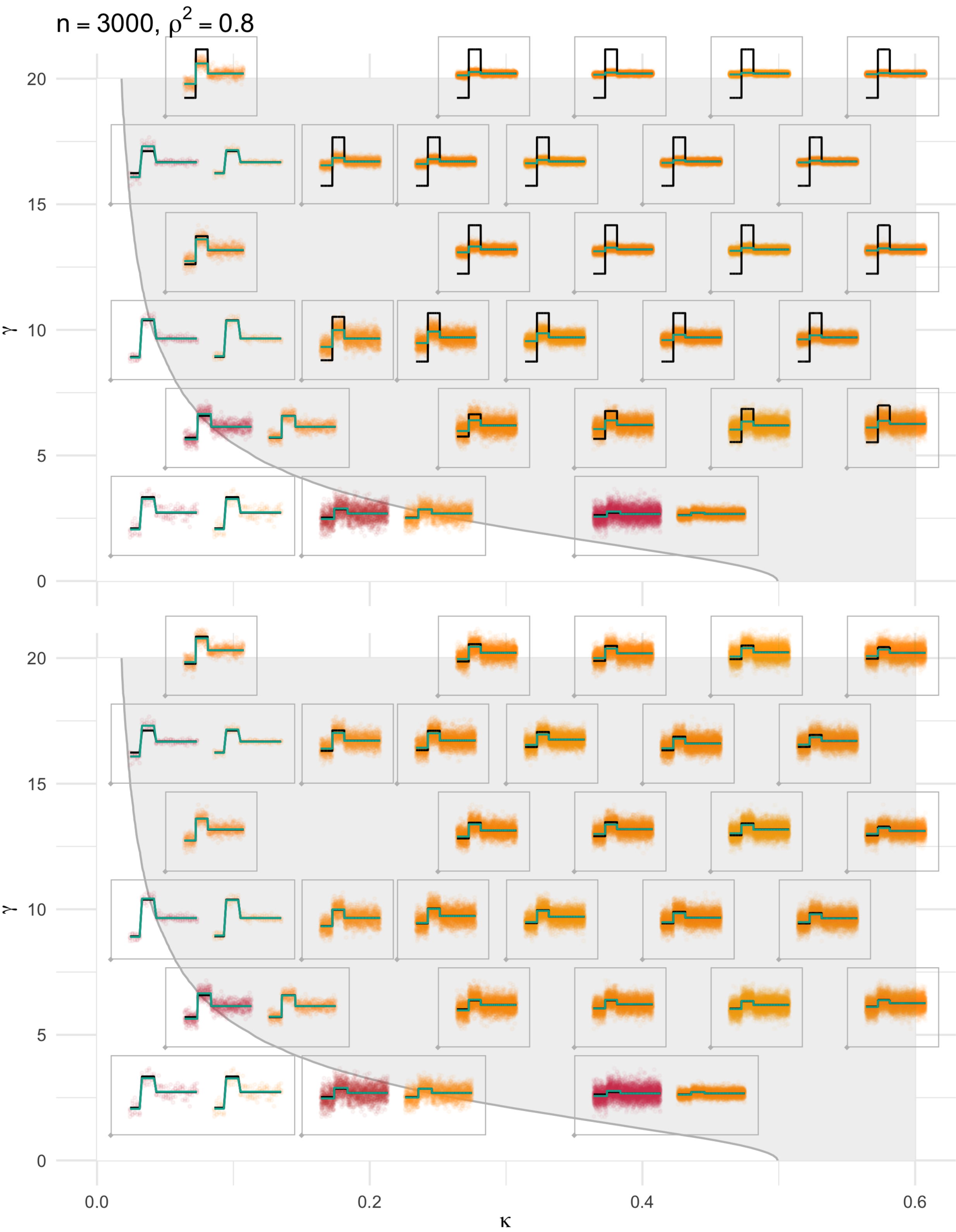
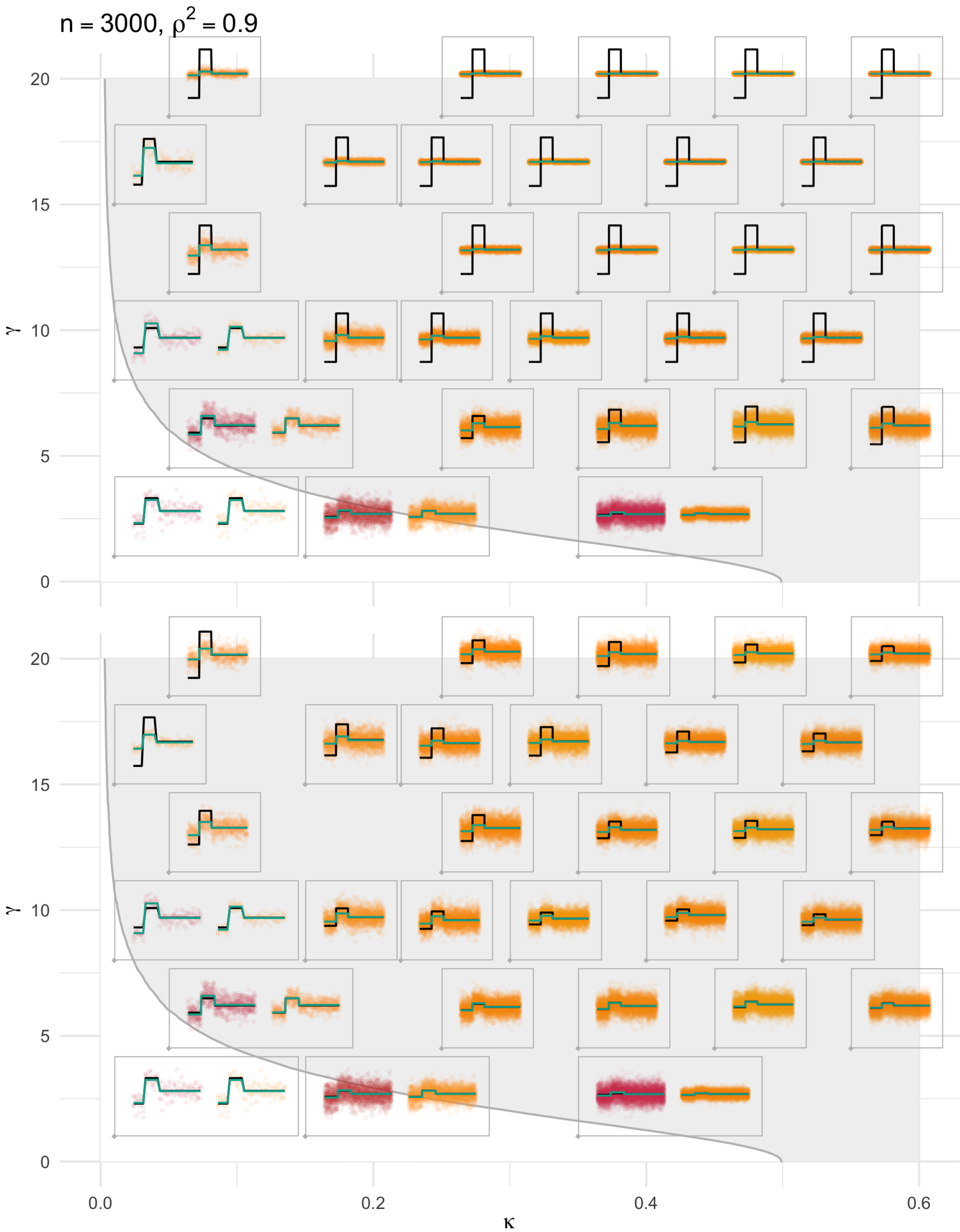


Figure S36: Setting b with $n = 3000$ and $\rho^2 = 0.9$.



S4 Computational performance for $\rho^2 = 0$: Setting a

Runtime summaries for the two-pass IWLS implementation of mJPL (average computational time in seconds, and minimum, average and maximum required number of iterations) across the 5 samples at each of 30 (κ, γ) settings, with β^* set to an equi-spaced grid of length $p = \lceil n\kappa \rceil$ between -10 and 10, for $n \in \{1000, 2000, 3000\}$, and $\rho^2 = 0$.

Table S2, Table S3, and Table S4 report the runtime summaries for the mJPL fits in Figure S1, Figure S7, and Figure S13, respectively.

Table S2: Setting a with $n = 1000$ and $\rho^2 = 0$.

κ	γ	p	average time (sec)	minimum iterations	average iterations	maximum iterations
0.01	1.0	10	0.01	4	4.0	4
0.01	8.0	10	0.02	9	9.0	9
0.01	15.0	10	0.03	11	11.4	12
0.05	4.5	50	0.12	7	7.8	8
0.05	11.5	50	0.21	12	13.8	15
0.05	18.5	50	0.49	19	28.0	48
0.15	1.0	150	0.36	4	4.0	4
0.15	8.0	150	4.42	38	46.2	53
0.15	15.0	150	3.39	25	35.8	49
0.22	8.0	220	7.48	32	39.0	56
0.22	15.0	220	4.59	23	24.2	29
0.25	4.5	250	6.47	18	27.6	32
0.25	11.5	250	5.27	19	22.2	26
0.25	18.5	250	4.77	18	20.4	23
0.30	8.0	300	6.85	16	20.8	25
0.30	15.0	300	5.73	15	17.2	19
0.35	1.0	350	3.03	6	7.0	9
0.35	4.5	350	12.04	21	28.0	48
0.35	11.5	350	7.93	14	18.4	26
0.35	18.5	350	7.10	14	16.0	18
0.40	8.0	400	11.29	13	15.2	19
0.40	15.0	400	9.65	14	15.8	19
0.45	4.5	450	14.94	16	19.6	25
0.45	11.5	450	12.12	12	14.4	18
0.45	18.5	450	10.38	11	14.4	17
0.50	8.0	500	11.39	13	13.8	16
0.50	15.0	500	10.90	12	13.2	14
0.55	4.5	550	16.01	11	15.8	21
0.55	11.5	550	13.13	11	12.0	13
0.55	18.5	550	13.98	10	12.0	15

Table S3: Setting a with $n = 2000$ and $\rho^2 = 0$.

κ	γ	p	average time (sec)	minimum iterations	average iterations	maximum iterations
0.01	1.0	20	0.04	4	4.0	4
0.01	8.0	20	0.10	8	8.6	9
0.01	15.0	20	0.14	11	11.2	12
0.05	4.5	100	0.87	7	7.2	8
0.05	11.5	100	1.29	12	13.0	14
0.05	18.5	100	2.58	17	26.2	36
0.15	1.0	300	3.17	4	4.0	4
0.15	8.0	300	27.49	27	38.4	55
0.15	15.0	300	24.22	27	31.8	40
0.22	8.0	440	45.82	24	29.6	42
0.22	15.0	440	32.11	18	21.4	25
0.25	4.5	500	74.22	23	40.0	91
0.25	11.5	500	41.39	18	21.2	29
0.25	18.5	500	37.15	17	19.0	22
0.30	8.0	600	59.52	17	21.4	25
0.30	15.0	600	51.22	16	18.0	21
0.35	1.0	700	21.58	6	6.6	7
0.35	4.5	700	101.74	22	29.0	40
0.35	11.5	700	63.17	16	17.0	18
0.35	18.5	700	56.39	14	15.4	17
0.40	8.0	800	73.69	14	15.6	20
0.40	15.0	800	65.75	13	13.8	15
0.45	4.5	900	102.83	17	18.0	19
0.45	11.5	900	80.31	13	13.0	13
0.45	18.5	900	69.78	10	11.8	13
0.50	8.0	1000	86.58	11	12.0	13
0.50	15.0	1000	83.12	10	11.8	15
0.55	4.5	1100	135.47	13	15.6	18
0.55	11.5	1100	96.42	10	10.8	11
0.55	18.5	1100	92.16	10	10.8	12

Table S4: Setting a with $n = 3000$ and $\rho^2 = 0$.

κ	γ	p	average time (sec)	minimum iterations	average iterations	maximum iterations
0.01	1.0	30	0.09	4	4.0	4
0.01	8.0	30	0.24	8	8.6	9
0.01	15.0	30	0.31	10	10.6	11
0.05	4.5	150	2.31	7	7.2	8
0.05	11.5	150	3.72	12	12.6	13
0.05	18.5	150	6.73	19	23.6	30
0.15	1.0	450	8.66	4	4.0	4
0.15	8.0	450	80.38	22	37.2	47
0.15	15.0	450	61.68	24	28.6	34
0.22	8.0	660	127.59	24	28.8	38
0.22	15.0	660	100.75	20	22.8	27
0.25	4.0	750	222.02	25	39.4	54
0.25	11.0	750	114.10	18	20.2	24
0.25	18.5	750	109.69	18	19.4	21
0.30	8.0	900	166.44	19	20.8	23
0.30	15.0	900	124.74	14	15.6	17
0.35	1.0	1050	66.40	6	6.2	7
0.35	4.5	1050	294.92	21	27.8	44
0.35	11.5	1050	158.85	13	14.6	16
0.35	18.5	1050	154.65	13	14.4	15
0.40	8.0	1200	211.97	15	15.4	16
0.40	15.0	1200	181.69	12	13.2	14
0.45	4.5	1350	293.02	16	17.0	20
0.45	11.5	1350	214.48	12	12.4	13
0.45	18.5	1350	199.81	10	11.6	13
0.50	8.0	1500	272.56	11	12.8	17
0.50	15.0	1500	231.55	10	10.8	11
0.55	4.5	1651	390.04	13	14.4	18
0.55	11.5	1651	288.30	10	10.4	12
0.55	18.5	1651	272.82	9	10.0	11

S5 Computational performance for $\rho^2 = 0$: Setting b

Runtime summaries for the two-pass IWLS implementation of mJPL (average computational time in seconds, and minimum, average and maximum required number of iterations) across the 5 samples at each of 30 (κ, γ) settings, with β^* set to have length $p = \lceil n\kappa \rceil$ with 20% of its values set to -10, 20% of its values set to 10, and the remaining set to 0, for $n \in \{1000, 2000, 3000\}$, and $\rho^2 = 0$.

Table S5, Table S6, and Table S7 report the runtime summaries for the mJPL fits in Figure S19, Figure S25, and Figure S31, respectively.

Table S5: Setting b with $n = 1000$ and $\rho^2 = 0$.

κ	γ	p	average time (sec)	minimum iterations	average iterations	maximum iterations
0.01	1.0	10	0.01	4	4.0	4
0.01	8.0	10	0.02	9	9.0	9
0.01	15.0	10	0.03	11	11.0	11
0.05	4.5	50	0.12	7	7.8	8
0.05	11.5	50	0.27	12	15.0	21
0.05	18.5	50	0.57	20	27.8	54
0.15	1.0	150	0.42	4	4.0	4
0.15	8.0	150	5.09	25	48.4	96
0.15	15.0	150	3.25	22	30.0	39
0.22	8.0	220	6.77	24	30.2	40
0.22	15.0	220	5.02	22	25.4	30
0.25	4.5	250	8.76	24	29.6	44
0.25	11.5	250	6.28	18	22.4	28
0.25	18.5	250	6.16	18	22.4	27
0.30	8.0	300	10.17	18	26.8	32
0.30	15.0	300	7.84	18	19.8	25
0.35	1.0	350	3.32	6	6.8	8
0.35	4.5	350	13.59	22	28.0	36
0.35	11.5	350	8.01	14	16.2	19
0.35	18.5	350	8.50	16	17.4	18
0.40	8.0	400	10.22	14	16.2	18
0.40	15.0	400	8.29	11	13.6	16
0.45	4.5	450	16.13	14	21.0	35
0.45	11.5	450	11.20	13	14.2	15
0.45	18.5	450	10.96	12	13.8	16
0.50	8.0	500	12.76	12	13.2	16
0.50	15.0	500	11.51	10	12.0	13
0.55	4.5	550	17.71	13	14.6	17
0.55	11.5	550	14.52	11	13.2	16
0.55	18.5	550	12.52	11	12.0	14

Table S6: Setting b with $n = 2000$ and $\rho^2 = 0$.

κ	γ	p	average time (sec)	minimum iterations	average iterations	maximum iterations
0.01	1.0	20	0.05	4	4.0	4
0.01	8.0	20	0.11	9	9.0	9
0.01	15.0	20	0.14	11	11.2	12
0.05	4.5	100	0.86	7	7.2	8
0.05	11.5	100	1.54	12	12.8	15
0.05	18.5	100	2.68	17	24.8	35
0.15	1.0	300	2.76	4	4.0	4
0.15	8.0	300	25.57	19	34.8	55
0.15	15.0	300	22.87	28	29.8	33
0.22	8.0	440	47.74	24	30.8	39
0.22	15.0	440	37.27	21	25.4	30
0.25	4.5	500	52.95	17	27.8	38
0.25	11.5	500	37.53	21	21.8	23
0.25	18.5	500	33.15	18	19.2	20
0.30	8.0	600	59.14	19	22.4	24
0.30	15.0	600	43.54	15	16.2	17
0.35	1.0	700	22.41	6	6.2	7
0.35	4.5	700	109.54	22	30.4	45
0.35	11.5	700	61.47	14	16.6	20
0.35	18.5	700	50.73	14	15.4	17
0.40	8.0	800	71.82	13	15.0	16
0.40	15.0	800	67.66	12	14.4	16
0.45	4.5	900	122.83	18	21.2	28
0.45	11.5	900	76.48	10	12.8	16
0.45	18.5	900	76.62	11	12.0	14
0.50	8.0	1000	92.12	12	12.6	14
0.50	15.0	1000	94.56	12	13.0	15
0.55	4.5	1100	127.65	13	14.8	16
0.55	11.5	1100	90.75	10	10.8	11
0.55	18.5	1100	94.43	9	10.6	12

Table S7: Setting b with $n = 3000$ and $\rho^2 = 0$.

κ	γ	p	average time (sec)	minimum iterations	average iterations	maximum iterations
0.01	1.0	30	0.10	4	4.0	4
0.01	8.0	30	0.21	8	8.2	9
0.01	15.0	30	0.34	10	10.8	11
0.05	4.5	150	2.34	7	7.2	8
0.05	11.5	150	3.90	12	13.2	14
0.05	18.5	150	6.19	17	21.4	26
0.15	1.0	450	8.66	4	4.0	4
0.15	8.0	450	74.05	24	34.2	46
0.15	15.0	450	59.41	23	27.6	36
0.22	8.0	660	122.98	21	27.8	33
0.22	15.0	660	93.83	18	21.2	24
0.25	4.5	750	181.15	26	32.0	35
0.25	11.5	750	115.06	20	20.4	22
0.25	18.5	750	110.68	18	19.6	21
0.30	8.0	900	155.57	14	19.4	21
0.30	15.0	900	138.84	16	17.4	18
0.35	1.0	1050	64.25	6	6.0	6
0.35	4.5	1050	257.86	22	24.2	27
0.35	11.5	1050	168.51	15	15.8	17
0.35	18.5	1050	158.72	13	14.8	17
0.40	8.0	1200	203.61	13	14.8	18
0.40	15.0	1200	187.69	13	13.6	14
0.45	4.5	1350	318.07	15	18.4	22
0.45	11.5	1350	206.77	11	12.0	13
0.45	18.5	1350	207.94	11	12.0	13
0.50	8.0	1500	247.68	11	11.8	12
0.50	15.0	1500	259.82	10	11.6	13
0.55	4.5	1651	375.09	13	13.4	14
0.55	11.5	1651	272.48	9	10.0	11
0.55	18.5	1651	282.12	10	10.4	11

References

Candès, E. J. and P. Sur (2020). The phase transition for the existence of the maximum likelihood estimate in high-dimensional logistic regression. *Annals of Statistics 48*(1), 27–42.

Firth, D. (1993). Bias reduction of maximum likelihood estimates. *Biometrika 80*(1), 27–38.

Kosmidis, I. and D. Firth (2021). Jeffreys-prior penalty, finiteness and shrinkage in binomial-response generalized linear models. *Biometrika 108*, 71–82.

Lumley, T. (2020). *biglm: Bounded Memory Linear and Generalized Linear Models*. R package version 0.9-2.1.