Statistica Sinica: Supplement

MULTISCALE BERNSTEIN POLYNOMIALS FOR DENSITIES

Antonio Canale[1,2] and David B. Dunson[3]

[1] *University of Turin,* [2] *Collegio Carlo Alberto, and* [3] *Duke University*

Supplementary Material

S1 Additional plots for Section 2

Figure 1 shows realizations from the prior for different a values. To better isolate the contribution of the a hyperparameter, we fixed the realizations of $R_{s,h} \sim \text{Be}(1,1)$ for all subplots. Figure 2 shows the prior expectation and 95% prior credible intervals of the total weight assigned to each scale for some alternative hyperparameter values.

S2 Additional plots for the Simulations of Section 4

Tables 1–2 report the mean Kolmogorov-Smirnoff, L_1, and L_2 distances between the true densities and all the models compared in Section 4. Figures 3–5 show graphically the same information. Table 3 reports the mean execution times over the simulations.

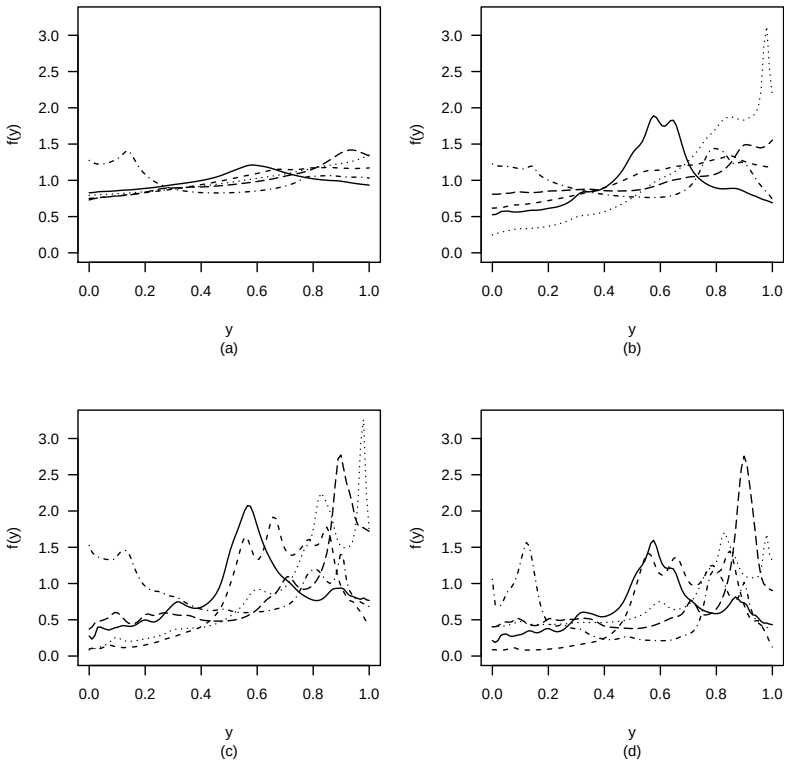

Figure 1: Five realizations from a multiscale Bernstein polynomial prior with $b = 1$ and (a) $a = 1$, (b) $a = 2$, (c) $a = 5$, and (d) $a = 10$.

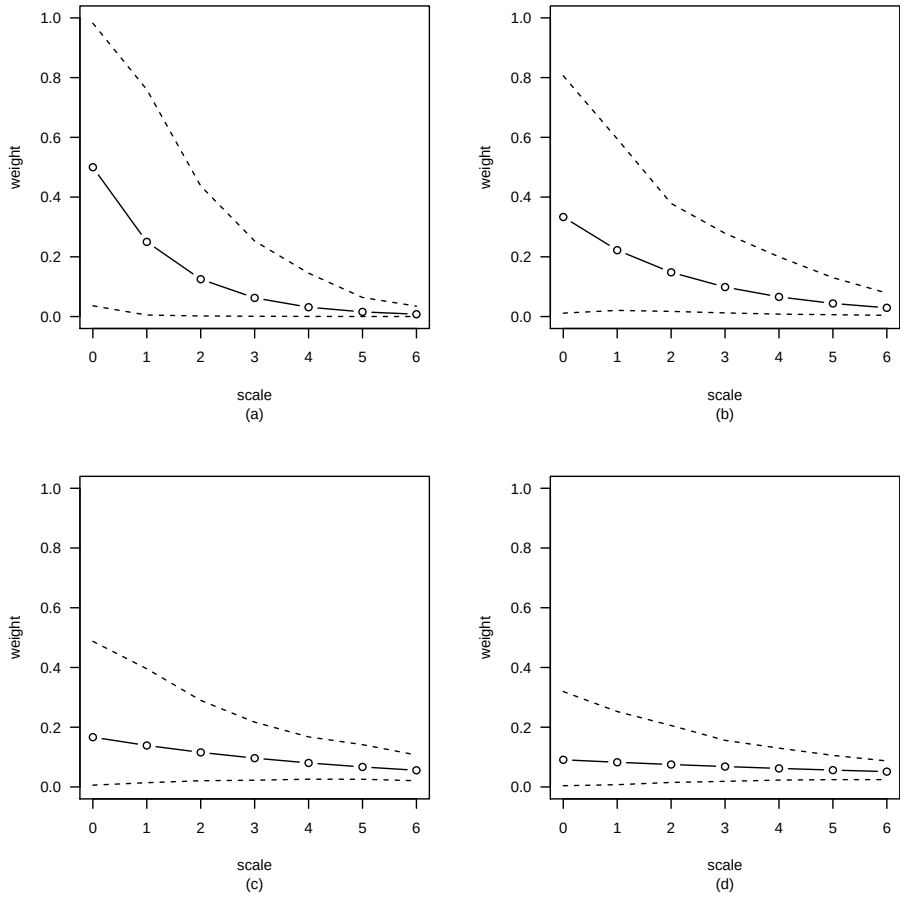

Figure 2: A priori expected weight for each scale and 95% credible intervals (dashed line) for: (a) $a = 1$ and $b = 1$, (b) $a = 2$ and $b = 1$, (c) $a = 5$ and $b = 1$, and (d) $a = 10$ and $b = 1$.

Table 1: Mean Kolmogorov Smirnoff (KS) distance, mean L_1 distance (L_1), and mean L_2 distance (L_2) between the true densities and the posterior msBP estimate (msBP), posterior DP mixture of Gaussians estimate (DPM), posterior DP mixture of Bernstein Polynomials estimate (DPBP), posterior mixture of Polya's tree estimate (PT), frequentist wavelet estimate (W), and frequentist kernel smoothing estimate (K) for Scenario 1 (S1), Scenario 2 (S2), Scenario 3 (S3), and Scenario 4 (S4). Mean distances computed over 200 samples, with Monte Carlo error in parenthesis.

		$n = 25$			$n = 50$		
		KS	L_1	L_2	KS	L_1	L_2
S1	msBP	1.1991 (0.15)	18.6393 (1.77)	12.0877 (2.46)	1.1207 (0.15)	16.7618 (1.98)	10.5297 (2.65)
	msBP-EB	0.9616 (0.28)	15.3337 (3.79)	9.0286 (4.06)	0.8529 (0.20)	12.4909 (2.78)	5.8835 (2.59)
	DPM	1.5785 (0.16)	18.1684 (1.78)	15.659 (2.76)	1.4137 (0.15)	18.1139 (1.50)	13.4228 (2.39)
	DPBP	1.2443 (0.19)	22.6341 (2.42)	15.9829 (3.53)	0.9245 (0.27)	15.3053 (3.83)	8.2186 (4.03)
	PT	2.4917 (0.01)	952.1645 (2.14)	1391.3997 (2.74)	2.4917 (0.01)	951.1084 (1.26)	1389.2295 (1.43)
	W	1.6867 (0.05)	26.5373 (0.81)	23.2277 (1.27)	1.6481 (0.04)	25.864 (0.73)	22.1622 (1.15)
	K	1.0629 (0.24)	15.8933 (3.44)	9.4448 (3.47)	0.8812 (0.21)	12.6056 (2.78)	6.0769 (2.70)
S2	msBP	0.1375 (0.04)	2.3126 (0.48)	0.1852 (0.08)	0.0961 (0.03)	1.6308 (0.35)	0.0888 (0.04)
	msBP-EB	0.0947 (0.03)	1.5028 (0.32)	0.0812 (0.03)	0.0742 (0.02)	1.106 (0.26)	0.0441 (0.02)
	DPM	0.1385 (0.06)	1.7884 (0.53)	0.1389 (0.08)	0.1012 (0.04)	1.3192 (0.4)	0.0728 (0.04)
	DPBP	0.2339 (0.01)	4.3513 (0.05)	0.6461 (0.01)	0.2339 (0.01)	4.488 (0.07)	0.6648 (0.01)
	PT	0.2347 (0.01)	94.2408 (0.44)	13.9915 (0.06)	0.2339 (0.01)	93.9891 (0.33)	13.9568 (0.03)
	W	0.1424 (0.05)	2.1501 (0.66)	0.1756 (0.1)	0.1027 (0.03)	1.562 (0.44)	0.0917 (0.05)
	K	0.0931 (0.02)	1.4714 (0.31)	0.0767 (0.03)	0.0778 (0.02)	1.173 (0.26)	0.0485 (0.02)
S3	msBP	0.2806 (0.05)	2.7758 (0.77)	0.3854 (0.18)	0.2571 (0.04)	2.2984 (0.64)	0.277 (0.13)
	DPM	0.2494 (0.07)	2.8651 (0.70)	0.3922 (0.20)	0.2276 (0.06)	2.3452 (0.58)	0.276 (0.14)
	DPBP	0.5137 (0.04)	6.8264 (0.21)	2.1555 (0.16)	0.5735 (0.03)	7.0762 (0.21)	2.4045 (0.15)
	PT	0.6621 (0.01)	157.7443 (0.85)	65.6996 (0.31)	0.6621 (0.01)	157.2414 (0.5)	65.5554 (0.18)
	W	0.2982 (0.05)	3.4876 (0.71)	0.4979 (0.17)	0.2759 (0.04)	3.149 (0.41)	0.4145 (0.10)
	K	0.2802 (0.05)	2.963 (0.92)	0.4318 (0.23)	0.2521 (0.04)	2.42 (0.69)	0.3006 (0.14)
S4	msBP	0.2942 (0.04)	4.3193 (0.82)	0.5608 (0.12)	0.2943 (0.03)	3.6779 (0.57)	0.5092 (0.05)
	DPM	0.3203 (0.06)	5.0048 (0.8)	0.7094 (0.25)	0.3037 (0.05)	4.4272 (0.63)	0.5958 (0.17)
	DPBP	0.4995 (0.01)	8.9148 (0.16)	1.8019 (0.05)	0.4995 (0.01)	9.0004 (0.13)	1.7803 (0.05)
	PT	0.4995 (0.01)	93.1303 (0.78)	20.5193 (0.17)	0.4995 (0.01)	92.7538 (0.65)	20.4479 (0.17)
	W	0.2990 (0.06)	5.4053 (0.76)	0.7075 (0.23)	0.2831 (0.04)	4.613 (0.53)	0.5752 (0.12)
	K	0.3000 (0.04)	4.322 (0.83)	0.5834 (0.14)	0.2924 (0.03)	3.799 (0.61)	0.5143 (0.08)

Note: 0.00 stands for "< 0.01"

Table 2: Mean Kolmogorov Smirnoff (KS) distance, mean L_1 distance (L_1), and mean L_2 distance (L_2) between the true densities and the posterior msBP estimate (msBP), posterior DP mixture of Gaussians estimate (DPM), posterior DP mixture of Bernstein Polynomials estimate (DPBP), posterior mixture of Polya's tree estimate (PT), frequentist wavelet estimate (W), and frequentist kernel smoothing estimate (K) for Scenario 1 (S1), Scenario 2 (S2), Scenario 3 (S3), and Scenario 4 (S4). Mean distances computed over 200 samples, with Monte Carlo error in parenthesis.

		$n = 100$			$n = 250$		
		KS	L_1	L_2	KS	L_1	L_2
S1	msBP	0.7318 (0.20)	10.2247 (2.44)	4.1602 (1.96)	0.4950 (0.11)	6.9636 (1.38)	1.8262 (0.74)
	DPM	1.3558 (0.17)	18.2278 (1.47)	13.0673 (2.46)	0.9307 (0.30)	12.1219 (4.02)	6.5113 (4.35)
	DPBP	0.6147 (0.24)	9.7378 (3.12)	3.3916 (2.16)	0.4092 (0.16)	6.4774 (1.82)	1.5369 (0.95)
	PT	2.4917 (0.01)	951.527 (0.94)	1389.741 (1.11)	2.4917 (0.01)	951.0927 (0.79)	1,389.0039 (1.39)
	W	1.6425 (0.03)	25.7625 (0.54)	21.9891 (0.84)	1.6365 (0.02)	25.6550 (0.33)	21.8122 (0.51)
	K	0.7623 (0.19)	10.396 (2.53)	4.3419 (1.99)	0.5702 (0.12)	7.6127 (1.55)	2.3262 (0.96)
S2	msBP	0.0642 (0.01)	0.9616 (0.18)	0.0327 (0.01)	0.0569 (0.01)	0.7717 (0.10)	0.0230 (0.01)
	DPM	0.0700 (0.03)	0.9485 (0.30)	0.0372 (0.02)	0.0514 (0.02)	0.6809 (0.19)	0.0194 (0.01)
	DPBP	0.2339 (0.01)	4.5672 (0.07)	0.6783 (0.01)	0.2339 (0.01)	4.6075 (0.06)	0.6862 (0.01)
	PT	0.2339 (0.01)	93.8067 (0.28)	13.9393 (0.02)	0.2339 (0.01)	93.4766 (0.15)	13.9142 (0.01)
	W	0.0717 (0.02)	1.141 (0.31)	0.0468 (0.02)	0.0574 (0.01)	0.8228 (0.15)	0.0251 (0.01)
	K	0.0665 (0.02)	0.9893 (0.18)	0.0344 (0.01)	0.0573 (0.02)	0.7455 (0.16)	0.0221 (0.01)
S3	msBP	0.2252 (0.03)	1.8722 (0.43)	0.1907 (0.07)	0.1753 (0.03)	1.5645 (0.24)	0.1347 (0.04)
	DPM	0.1938 (0.05)	1.8194 (0.31)	0.1735 (0.07)	0.1545 (0.03)	1.4024 (0.22)	0.1083 (0.03)
	DPBP	0.6019 (0.01)	7.1933 (0.20)	2.5392 (0.13)	0.6183 (0.01)	7.3056 (0.16)	2.6557 (0.11)
	PT	0.6621 (0.01)	156.9909 (0.26)	65.4821 (0.10)	0.6621 (0.01)	156.8642 (0.24)	65.4271 (0.07)
	W	0.2599 (0.02)	2.9623 (0.25)	0.3631 (0.05)	0.2432 (0.01)	2.9181 (0.17)	0.3451 (0.03)
	K	0.2231 (0.03)	1.9428 (0.45)	0.2042 (0.08)	0.1902 (0.02)	1.5340 (0.27)	0.1321 (0.04)
S4	msBP	0.2856 (0.02)	3.4838 (0.35)	0.4759 (0.04)	0.2658 (0.02)	3.4730 (0.21)	0.4371 (0.02)
	DPM	0.2966 (0.04)	3.9836 (0.59)	0.5428 (0.15)	0.2891 (0.03)	3.6104 (0.31)	0.4740 (0.08)
	DPBP	0.4995 (0.01)	9.0851 (0.08)	1.7881 (0.03)	0.4995 (0.01)	9.1581 (0.07)	1.8103 (0.04)
	PT	0.4995 (0.01)	92.4458 (0.52)	20.4152 (0.14)	0.4995 (0.01)	92.0746 (0.31)	20.3587 (0.10)
	W	0.2734 (0.03)	4.0647 (0.42)	0.5036 (0.08)	0.2671 (0.02)	3.7149 (0.22)	0.4663 (0.03)
	K	0.2834 (0.02)	3.5222 (0.44)	0.4744 (0.05)	0.2761 (0.02)	3.3871 (0.23)	0.4501 (0.02)

Note: 0.00 stands for "< 0.01"

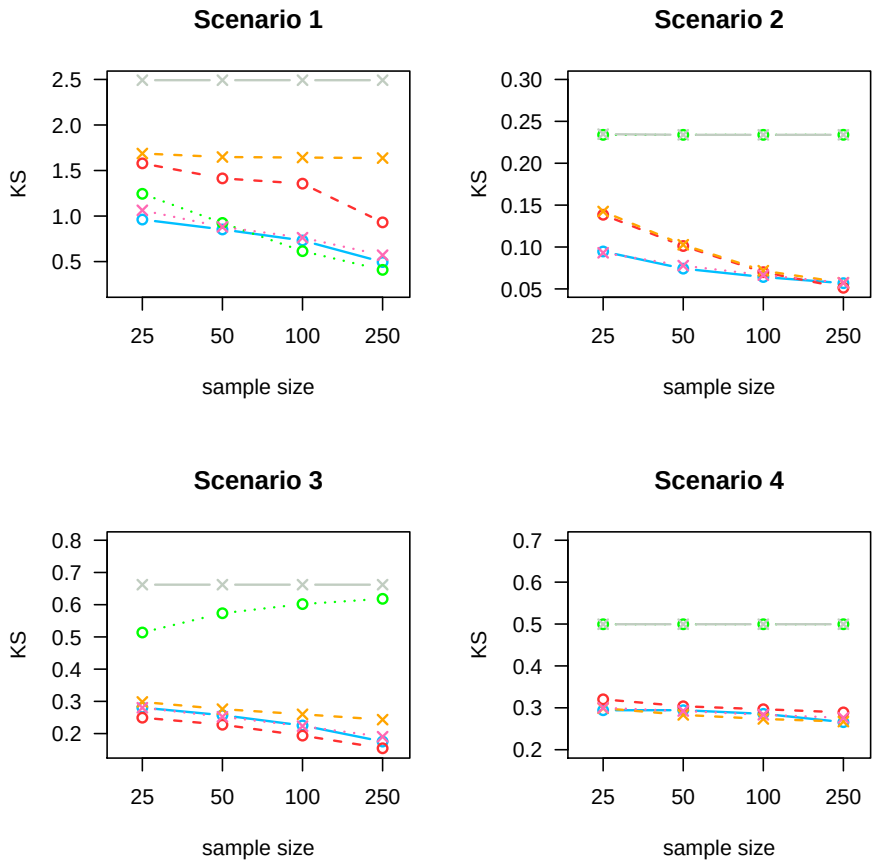

Figure 3: Mean Kolmogorov-Smirnoff distance between the true densities and the posterior msBP estimate (blue line), posterior DP mixture of Gaussians estimate (red line), posterior DP mixture of Bernstein Polynomials estimate (green line), posterior mixture of Polya trees estimate (gray), frequentist wavelet estimate (orange), and frequentist kernel smoothing estimate (pink line) under the four scenarios.

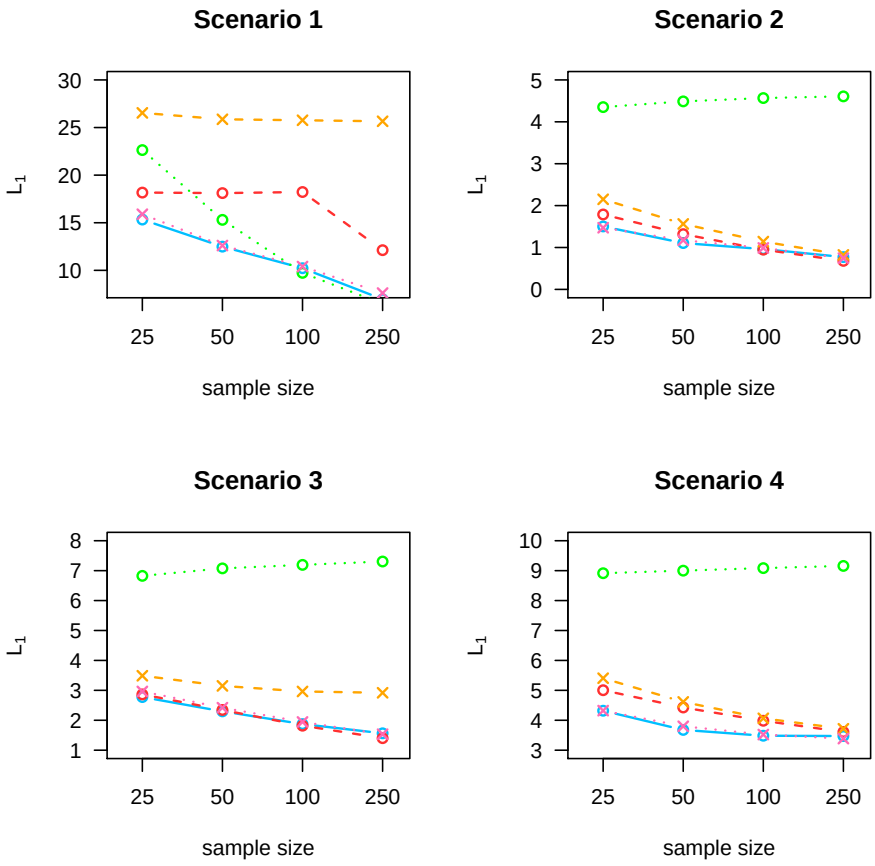

Figure 4: Mean L_1 distance between the true densities and the posterior multiscale Bernstein polynomial estimate (continuous line, circle dots), posterior DP mixture of Gaussians estimate (dashed line, circle dots), posterior DP mixture of Bernstein polynomials estimate (dotted line, circle dots), frequentist wavelet estimate (dashed line, "×" dots), and frequentist kernel smoothing estimate (dotted line, "×" dots) under the four scenarios. The posterior mixture of Polya trees estimate is far away and it is not reported for graphical reasons.

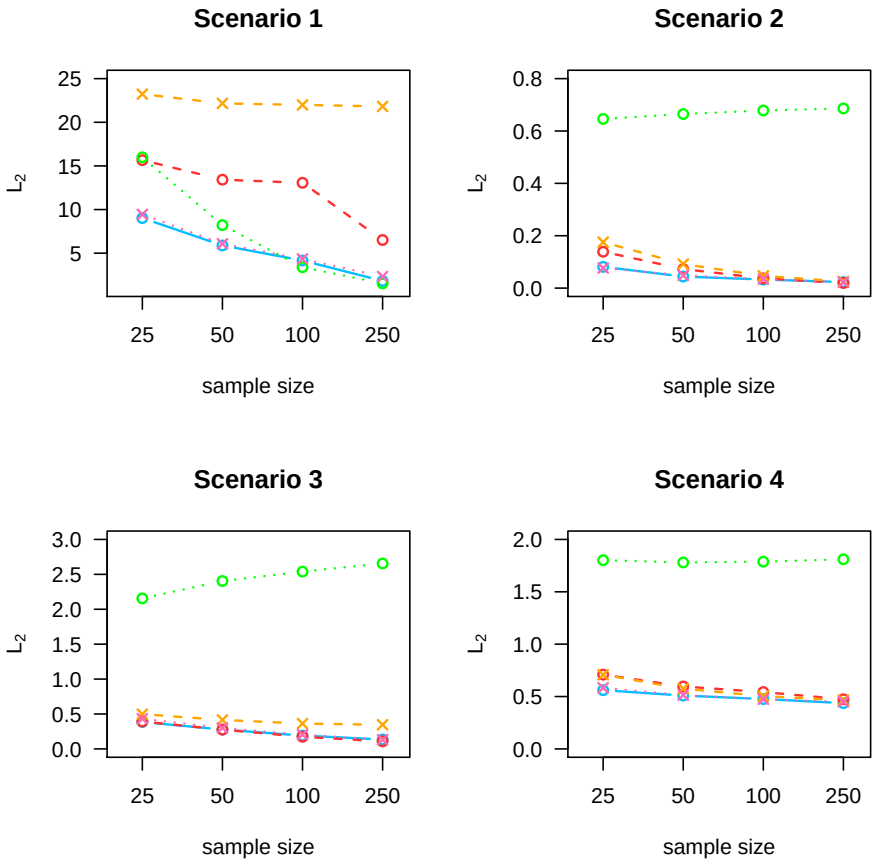

Figure 5: Mean L_2 distance between the true densities and the posterior msBP estimate (blue line), posterior DP mixture of Gaussians estimate (red line), posterior DP mixture of Bernstein Polynomials estimate (green line), frequentist wavelet estimate (orange), and frequentist kernel smoothing estimate (pink line) under the four scenarios. The posterior mixture of Polya trees estimate is far away and it is not reported for graphical reasons.

Table 3: Mean execution times (in seconds) for the msBP model (msBP), DP mixture of Gaussians (DPM), DP mixture of Bernstein Polynomials (DPBP), mixture of Polya's tree (PT), wavelet (W), and frequentist kernel smoothing (K) for Scenario 1 (S1), Scenario 2 (S2), Scenario 3 (S3), and Scenario 4 (S4).

	S1				S2			
	$n = 25$	$n = 50$	$n = 100$	$n = 250$	$n = 25$	$n = 50$	$n = 100$	$n = 250$
msBP	1.302	2.652	3.242	4.639	1.128	1.211	0.797	0.231
DPM	0.461	0.953	1.793	5.098	0.536	1.004	1.985	4.924
DPBP	5.507	10.419	19.122	40.746	6.662	10.431	17.177	39.629
PT	0.162	0.193	0.308	0.436	0.186	0.216	0.275	0.467
W	0.039	0.031	0.035	0.031	0.027	0.037	0.029	0.030
K	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001
	S3				S4			
	$n = 25$	$n = 50$	$n = 100$	$n = 250$	$n = 25$	$n = 50$	$n = 100$	$n = 250$
msBP	1.683	1.853	2.617	3.853	1.379	2.312	3.102	1.913
DPM	0.523	1.012	1.953	4.911	0.586	1.070	2.089	5.223
DPBP	4.882	7.995	18.080	46.881	5.989	9.970	17.770	45.569
PT	0.164	0.194	0.264	0.458	0.203	0.239	0.286	0.473
W	0.028	0.030	0.028	0.031	0.032	0.029	0.032	0.031
K	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001