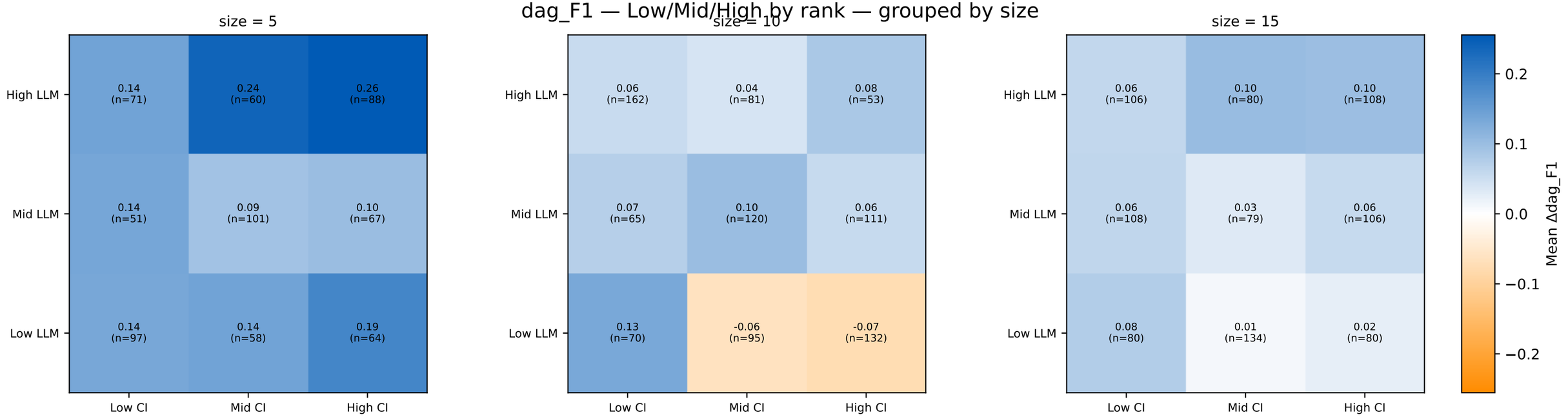
Factor Panels — Rank-based Low/Mid/High

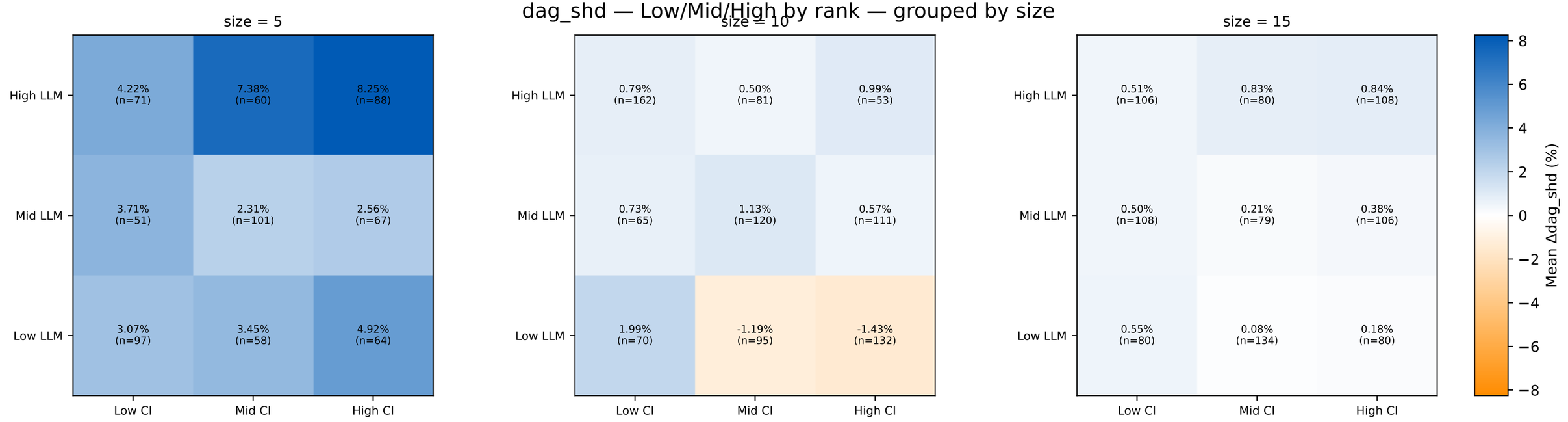
Panels by size/type/density; shared color scales per page.

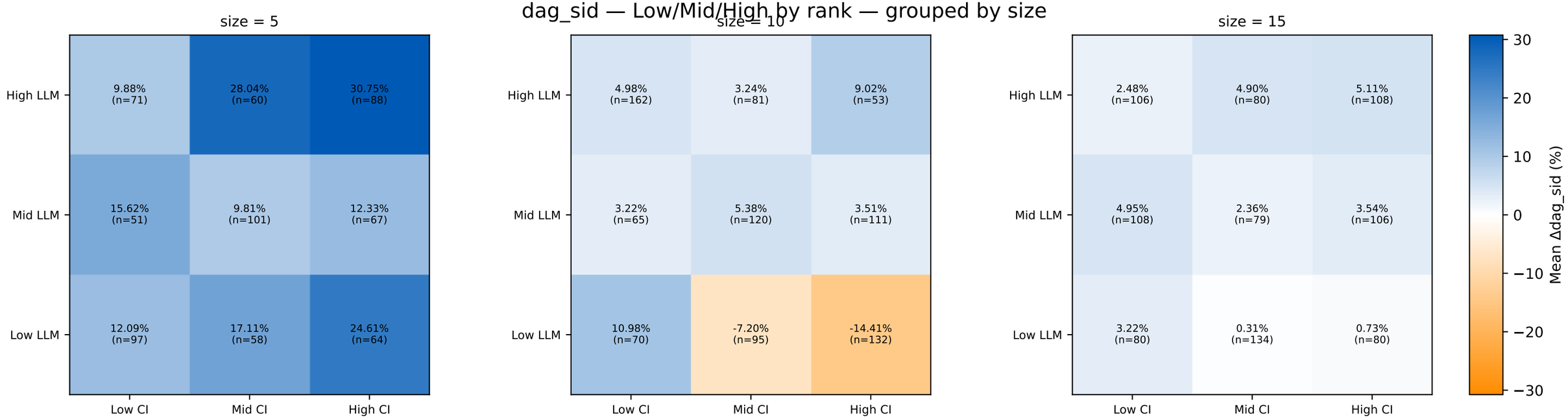
2025-10-02 12:50

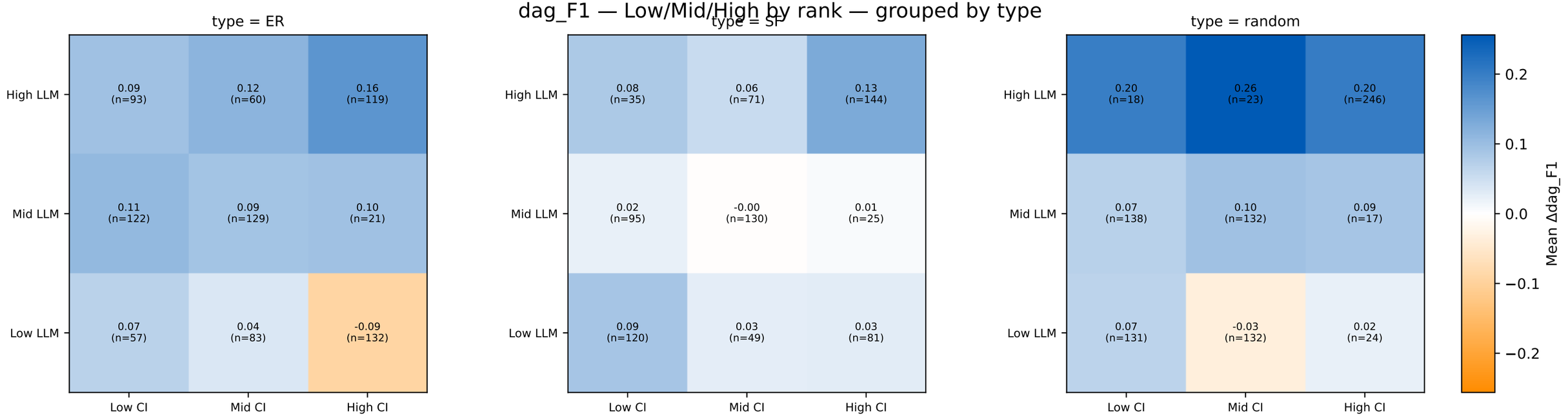
DAG Metrics

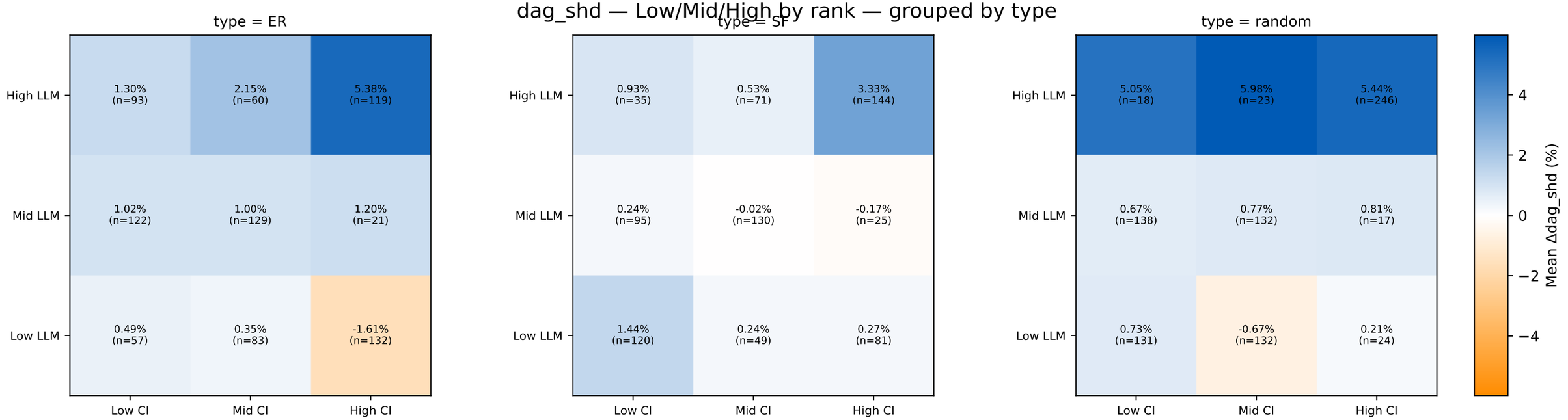
2025-10-02 12:50

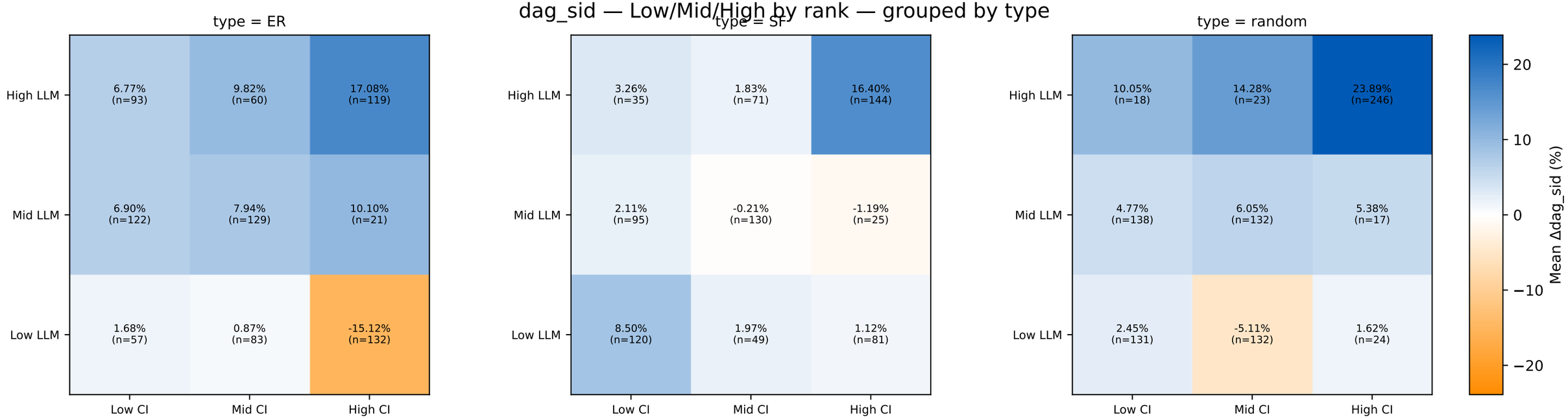

dag_F1 — Low/Mid/High by rank — grouped by size
size = 5
size = 10
size = 15
High LLM
Mid LLM
Low LLM
Low CI
Mid CI
High CI
0.14 (n=71)
0.24 (n=60)
0.26 (n=88)
0.14 (n=51)
0.09 (n=101)
0.10 (n=67)
0.14 (n=97)
0.14 (n=58)
0.19 (n=64)
0.06 (n=162)
0.04 (n=81)
0.08 (n=53)
0.07 (n=65)
0.10 (n=120)
0.06 (n=111)
0.13 (n=70)
-0.06 (n=95)
-0.07 (n=132)
0.06 (n=106)
0.10 (n=80)
0.10 (n=108)
0.06 (n=108)
0.03 (n=79)
0.06 (n=106)
0.08 (n=80)
0.01 (n=134)
0.02 (n=80)
Mean Δdag_F1
0.2
0.1
0.0
−0.1
−0.2

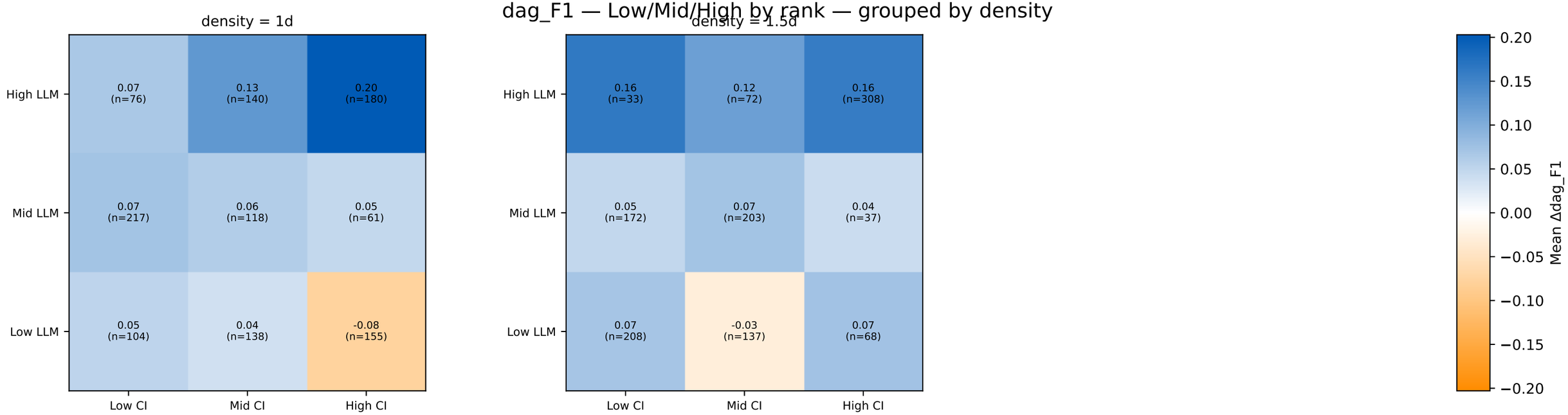

dag_shd — Low/Mid/High by rank — grouped by size
size = 5
size = 10
size = 15
High LLM
Mid LLM
Low LLM
Low CI
Mid CI
High CI
4.22% (n=71)
7.38% (n=60)
8.25% (n=88)
3.71% (n=51)
2.31% (n=101)
2.56% (n=67)
3.07% (n=97)
3.45% (n=58)
4.92% (n=64)
0.79% (n=162)
0.50% (n=81)
0.99% (n=53)
0.73% (n=65)
1.13% (n=120)
0.57% (n=111)
1.99% (n=70)
-1.19% (n=95)
-1.43% (n=132)
0.51% (n=106)
0.83% (n=80)
0.84% (n=108)
0.50% (n=108)
0.21% (n=79)
0.38% (n=106)
0.55% (n=80)
0.08% (n=134)
0.18% (n=80)
Mean Δdag_shd (%)
8
6
4
2
0
−2
−4
−6
−8

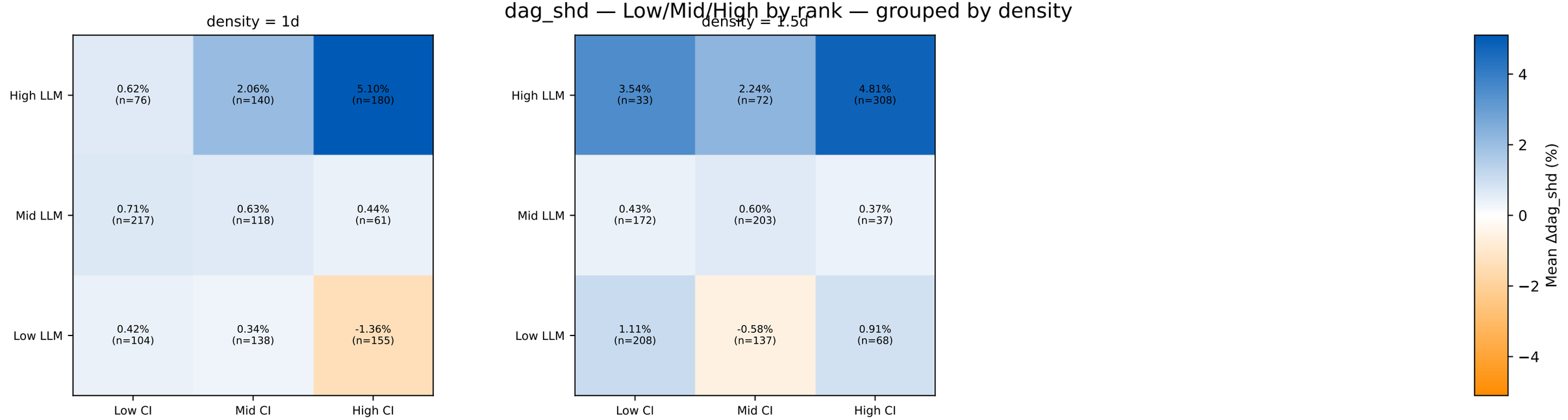

dag_sid — Low/Mid/High by rank — grouped by size
size = 5
size = 10
size = 15
High LLM
Mid LLM
Low LLM
Low CI
Mid CI
High CI
9.88% (n=71)
28.04% (n=60)
30.75% (n=88)
15.62% (n=51)
9.81% (n=101)
12.33% (n=67)
12.09% (n=97)
17.11% (n=58)
24.61% (n=64)
4.98% (n=162)
3.24% (n=81)
9.02% (n=53)
3.22% (n=65)
5.38% (n=120)
3.51% (n=111)
10.98% (n=70)
-7.20% (n=95)
-14.41% (n=132)
2.48% (n=106)
4.90% (n=80)
5.11% (n=108)
4.95% (n=108)
2.36% (n=79)
3.54% (n=106)
3.22% (n=80)
0.31% (n=134)
0.73% (n=80)
Mean Δdag_sid (%)
30
20
10
0
−10
−20
−30

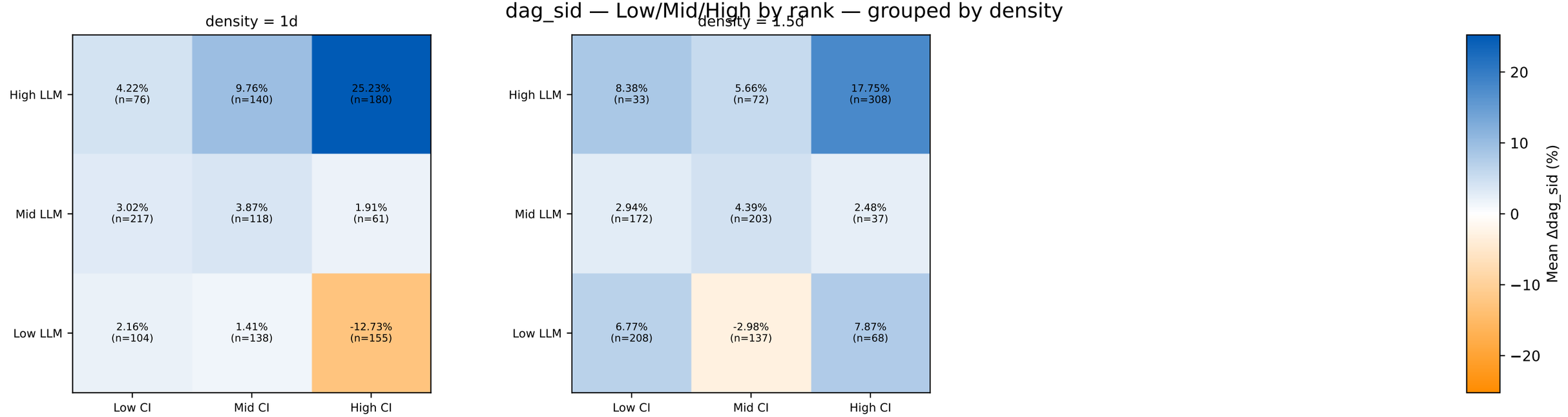

dag_F1 — Low/Mid/High by rank — grouped by type
type = ER
type = SF
type = random
High LLM
Mid LLM
Low LLM
Low CI
Mid CI
High CI
0.09 (n=93)
0.12 (n=60)
0.16 (n=119)
0.11 (n=122)
0.09 (n=129)
0.10 (n=21)
0.07 (n=57)
0.04 (n=83)
-0.09 (n=132)
0.08 (n=35)
0.06 (n=71)
0.13 (n=144)
0.02 (n=95)
-0.00 (n=130)
0.01 (n=25)
0.09 (n=120)
0.03 (n=49)
0.03 (n=81)
0.20 (n=18)
0.26 (n=23)
0.20 (n=246)
0.07 (n=138)
0.10 (n=132)
0.09 (n=17)
0.07 (n=131)
-0.03 (n=132)
0.02 (n=24)
Mean Δdag_F1
0.2
0.1
0.0
−0.1
−0.2


dag_shd — Low/Mid/High by rank — grouped by type
type = ER
type = SF
type = random
High LLM
Mid LLM
Low LLM
1.30%
(n=93)
2.15%
(n=60)
5.38%
(n=119)
1.02%
(n=122)
1.00%
(n=129)
1.20%
(n=21)
0.49%
(n=57)
0.35%
(n=83)
-1.61%
(n=132)
0.93%
(n=35)
0.53%
(n=71)
3.33%
(n=144)
0.24%
(n=95)
-0.02%
(n=130)
-0.17%
(n=25)
1.44%
(n=120)
0.24%
(n=49)
0.27%
(n=81)
5.05%
(n=18)
5.98%
(n=23)
5.44%
(n=246)
0.67%
(n=138)
0.77%
(n=132)
0.81%
(n=17)
0.73%
(n=131)
-0.67%
(n=132)
0.21%
(n=24)
Low CI
Mid CI
High CI
Mean Δdag_shd (%)
4
2
0
−2
−4


dag_sid — Low/Mid/High by rank — grouped by type
type = ER
type = SF
type = random
High LLM
Mid LLM
Low LLM
Low CI
Mid CI
High CI
6.77% (n=93)
9.82% (n=60)
17.08% (n=119)
6.90% (n=122)
7.94% (n=129)
10.10% (n=21)
1.68% (n=57)
0.87% (n=83)
-15.12% (n=132)
3.26% (n=35)
1.83% (n=71)
16.40% (n=144)
2.11% (n=95)
-0.21% (n=130)
-1.19% (n=25)
8.50% (n=120)
1.97% (n=49)
1.12% (n=81)
10.05% (n=18)
14.28% (n=23)
23.89% (n=246)
4.77% (n=138)
6.05% (n=132)
5.38% (n=17)
2.45% (n=131)
-5.11% (n=132)
1.62% (n=24)
Mean Δdag_sid (%)
20
10
0
−10
−20

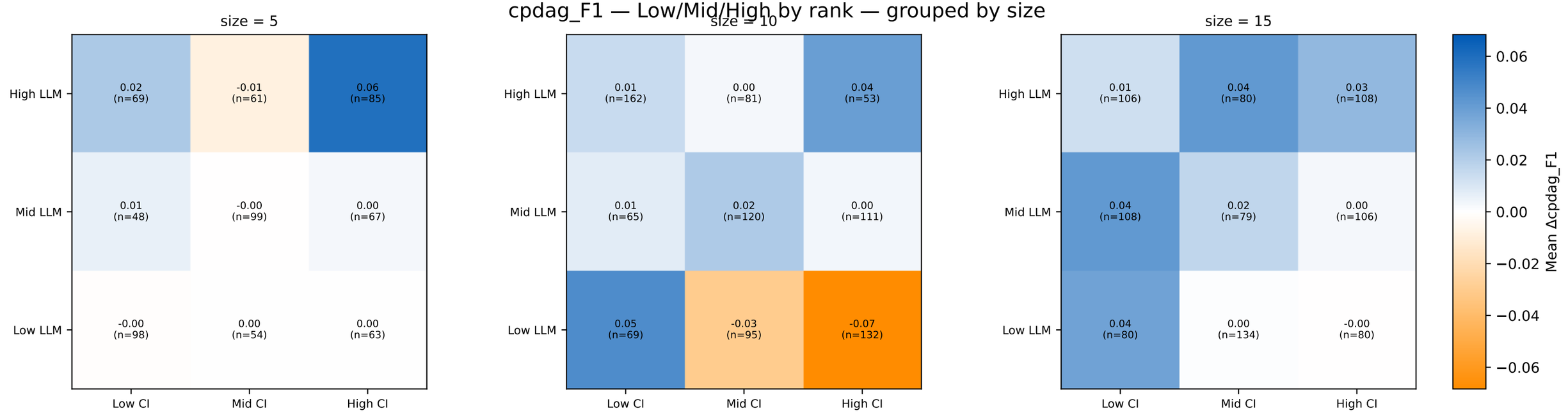

dag_F1 — Low/Mid/High by rank — grouped by density
density = 1d
density = 1.5d
High LLM
Mid LLM
Low LLM
Low CI
Mid CI
High CI
0.07 (n=76)
0.13 (n=140)
0.20 (n=180)
0.07 (n=217)
0.06 (n=118)
0.05 (n=61)
0.05 (n=104)
0.04 (n=138)
-0.08 (n=155)
0.16 (n=33)
0.12 (n=72)
0.16 (n=308)
0.05 (n=172)
0.07 (n=203)
0.04 (n=37)
0.07 (n=208)
-0.03 (n=137)
0.07 (n=68)
Mean Δdag_F1
0.20
0.15
0.10
0.05
0.00
−0.05
−0.10
−0.15
−0.20

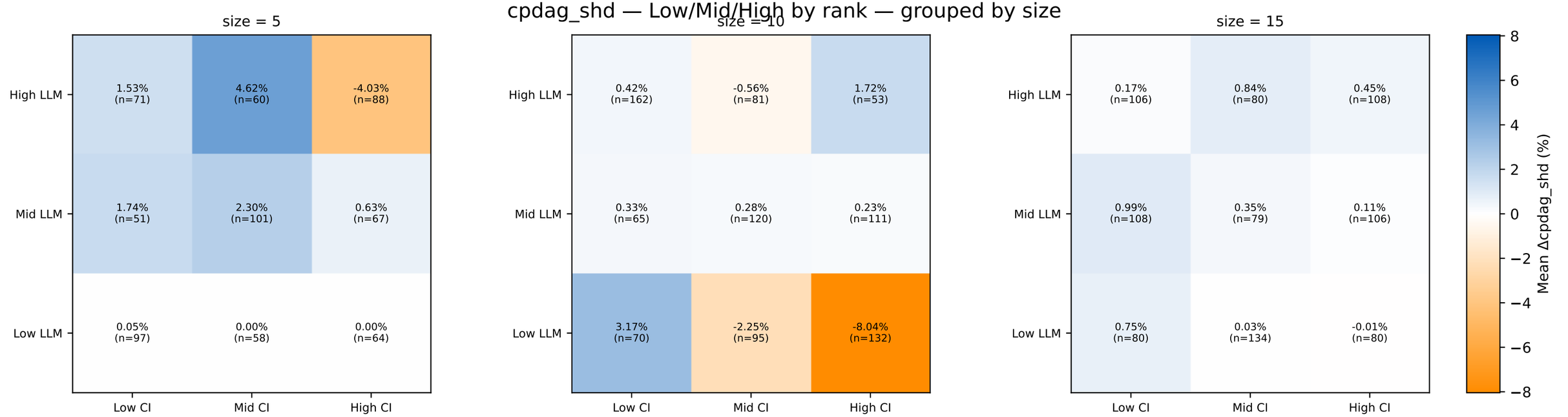

dag_shd — Low/Mid/High by rank — grouped by density
density = 1d
density = 1.5d
High LLM
Mid LLM
Low LLM
Low CI
Mid CI
High CI
0.62%
(n=76)
2.06%
(n=140)
5.10%
(n=180)
0.71%
(n=217)
0.63%
(n=118)
0.44%
(n=61)
0.42%
(n=104)
0.34%
(n=138)
-1.36%
(n=155)
3.54%
(n=33)
2.24%
(n=72)
4.81%
(n=308)
0.43%
(n=172)
0.60%
(n=203)
0.37%
(n=37)
1.11%
(n=208)
-0.58%
(n=137)
0.91%
(n=68)
Mean Δdag_shd (%)
4
2
0
−2
−4

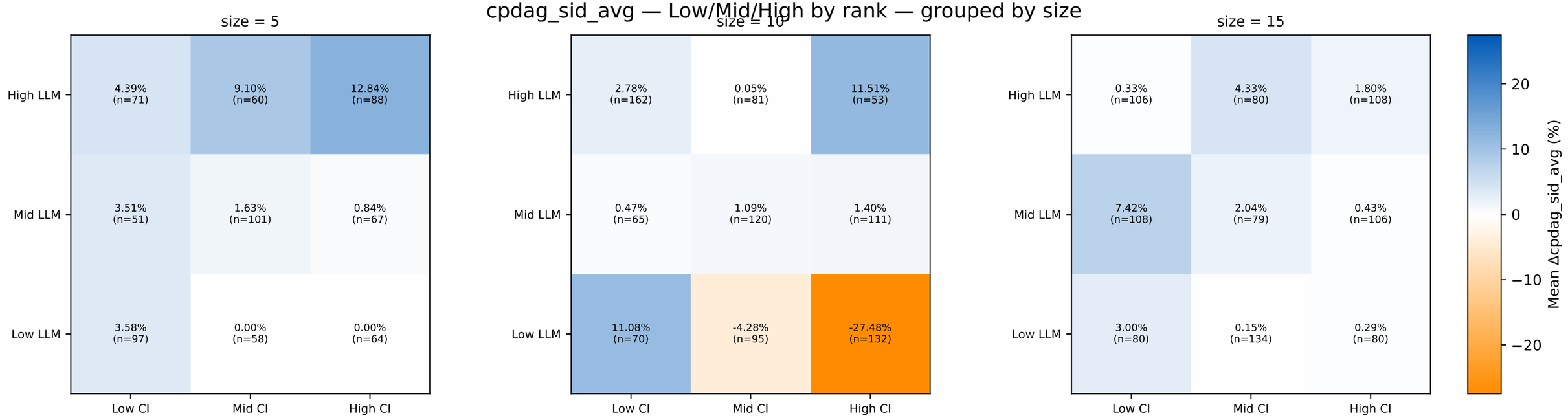

dag_sid — Low/Mid/High by rank — grouped by density
density = 1d
density = 1.5d
High LLM
Mid LLM
Low LLM
Low CI
Mid CI
High CI
4.22%
(n=76)
9.76%
(n=140)
25.23%
(n=180)
3.02%
(n=217)
3.87%
(n=118)
1.91%
(n=61)
2.16%
(n=104)
1.41%
(n=138)
-12.73%
(n=155)
8.38%
(n=33)
5.66%
(n=72)
17.75%
(n=308)
2.94%
(n=172)
4.39%
(n=203)
2.48%
(n=37)
6.77%
(n=208)
-2.98%
(n=137)
7.87%
(n=68)
Mean Δdag_sid (%)
20
10
0
−10
−20

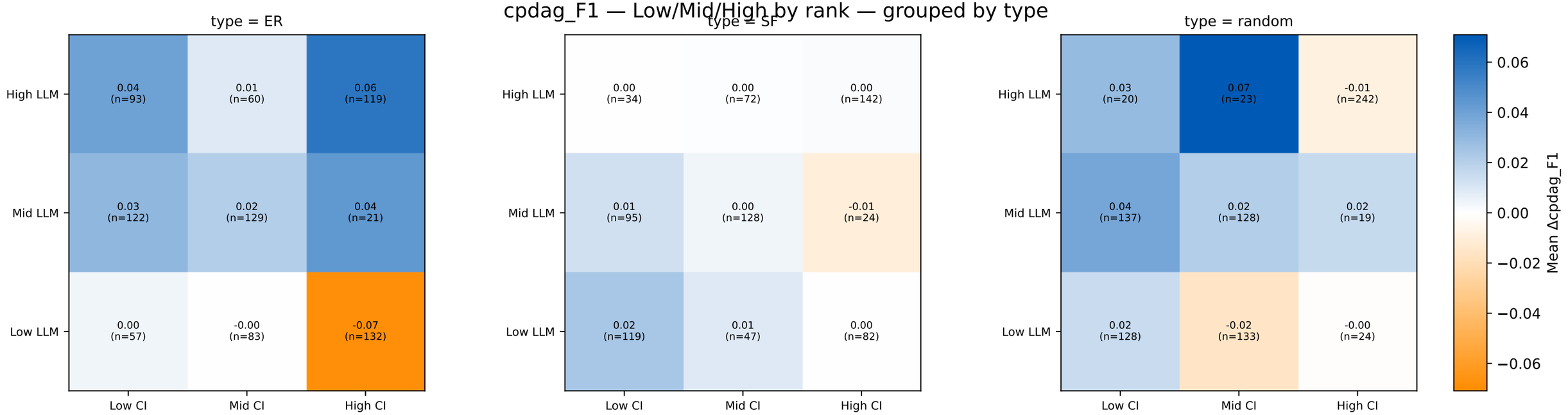
CPDAG Metrics

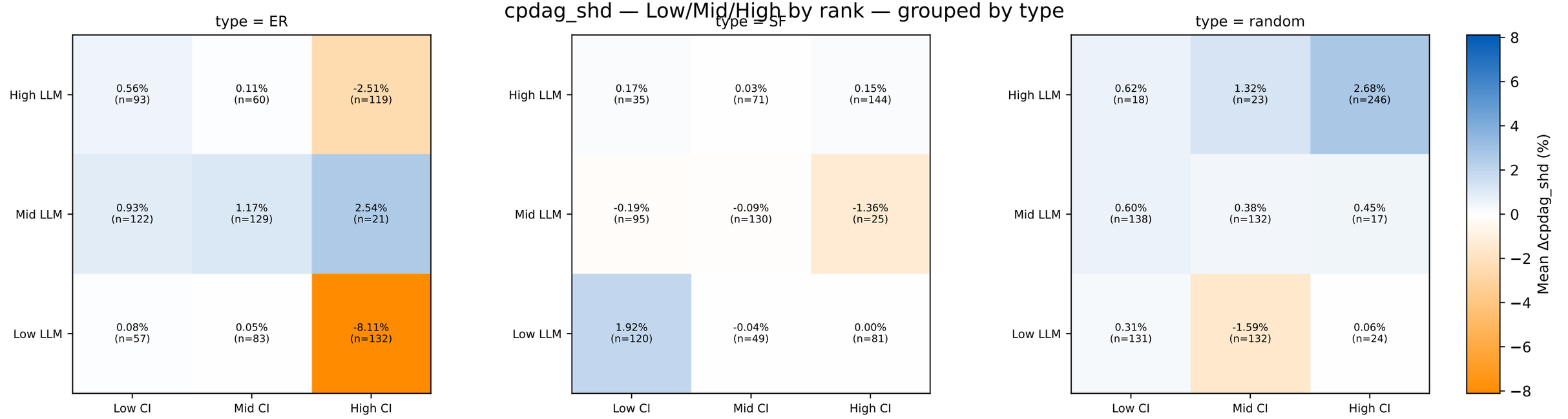
2025-10-02 12:50

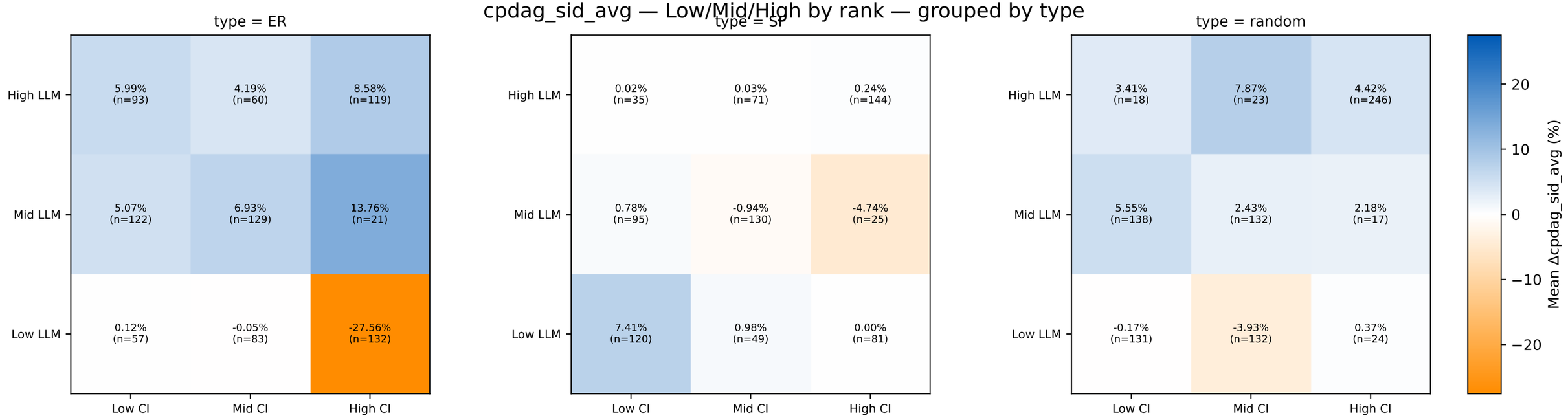

cpdag_F1 — Low/Mid/High by rank — grouped by size
size = 5
size = 10
size = 15
High LLM
Mid LLM
Low LLM
Low CI
Mid CI
High CI
0.02 (n=69)
-0.01 (n=61)
0.06 (n=85)
0.01 (n=48)
-0.00 (n=99)
0.00 (n=67)
-0.00 (n=98)
0.00 (n=54)
0.00 (n=63)
0.01 (n=162)
0.00 (n=81)
0.04 (n=53)
0.01 (n=65)
0.02 (n=120)
0.00 (n=111)
0.05 (n=69)
-0.03 (n=95)
-0.07 (n=132)
0.01 (n=106)
0.04 (n=80)
0.03 (n=108)
0.04 (n=108)
0.02 (n=79)
0.00 (n=106)
0.04 (n=80)
0.00 (n=134)
-0.00 (n=80)
Mean Δcpdag_F1
0.06
0.04
0.02
0.00
−0.02
−0.04
−0.06

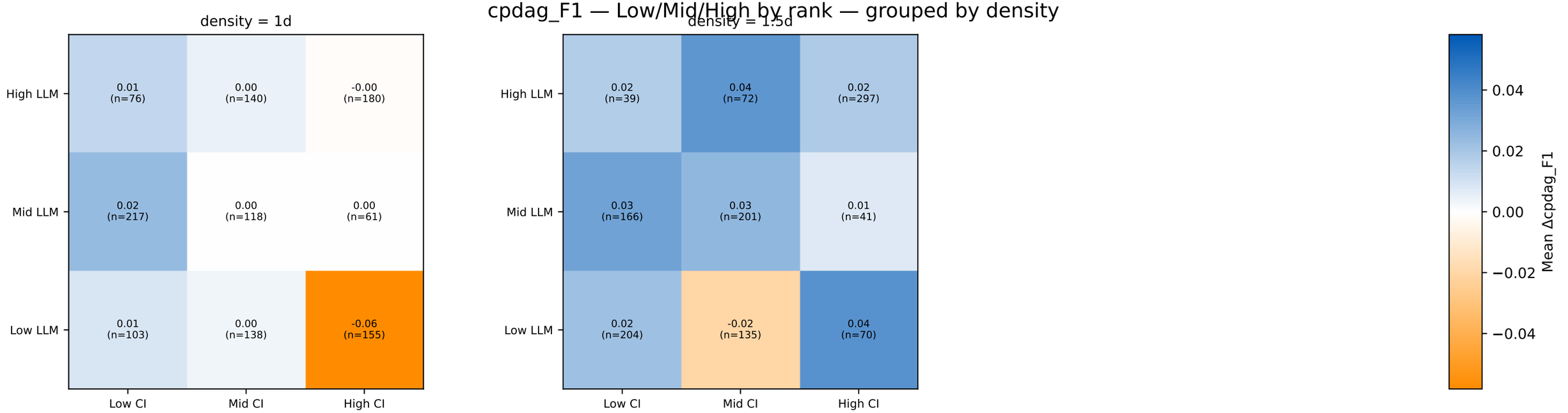

cpdag_shd — Low/Mid/High by rank — grouped by size
size = 5
size = 10
size = 15
High LLM
Mid LLM
Low LLM
Low CI
Mid CI
High CI
1.53% (n=71)
4.62% (n=60)
-4.03% (n=88)
1.74% (n=51)
2.30% (n=101)
0.63% (n=67)
0.05% (n=97)
0.00% (n=58)
0.00% (n=64)
0.42% (n=162)
-0.56% (n=81)
1.72% (n=53)
0.33% (n=65)
0.28% (n=120)
0.23% (n=111)
3.17% (n=70)
-2.25% (n=95)
-8.04% (n=132)
0.17% (n=106)
0.84% (n=80)
0.45% (n=108)
0.99% (n=108)
0.35% (n=79)
0.11% (n=106)
0.75% (n=80)
0.03% (n=134)
-0.01% (n=80)
Mean Δcpdag_shd (%)
8
6
4
2
0
−2
−4
−6
−8

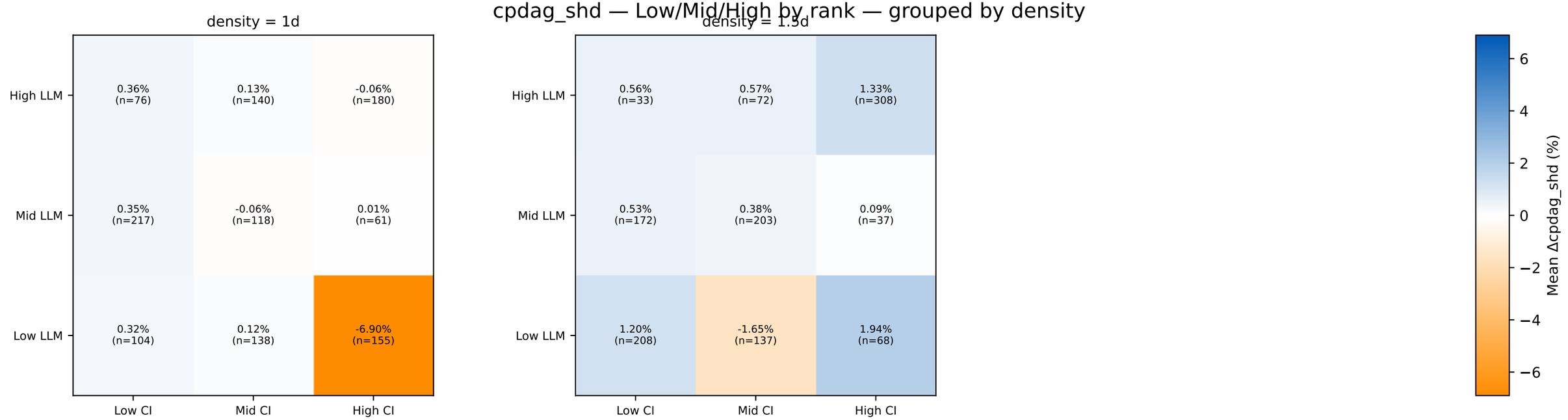

cpdag_sid_avg — Low/Mid/High by rank — grouped by size
size = 5
size = 10
size = 15
High LLM
Mid LLM
Low LLM
Low CI
Mid CI
High CI
4.39% (n=71)
9.10% (n=60)
12.84% (n=88)
3.51% (n=51)
1.63% (n=101)
0.84% (n=67)
3.58% (n=97)
0.00% (n=58)
0.00% (n=64)
2.78% (n=162)
0.05% (n=81)
11.51% (n=53)
0.47% (n=65)
1.09% (n=120)
1.40% (n=111)
11.08% (n=70)
-4.28% (n=95)
-27.48% (n=132)
0.33% (n=106)
4.33% (n=80)
1.80% (n=108)
7.42% (n=108)
2.04% (n=79)
0.43% (n=106)
3.00% (n=80)
0.15% (n=134)
0.29% (n=80)
Mean Δcpdag_sid_avg (%)
20
10
0
−10
−20

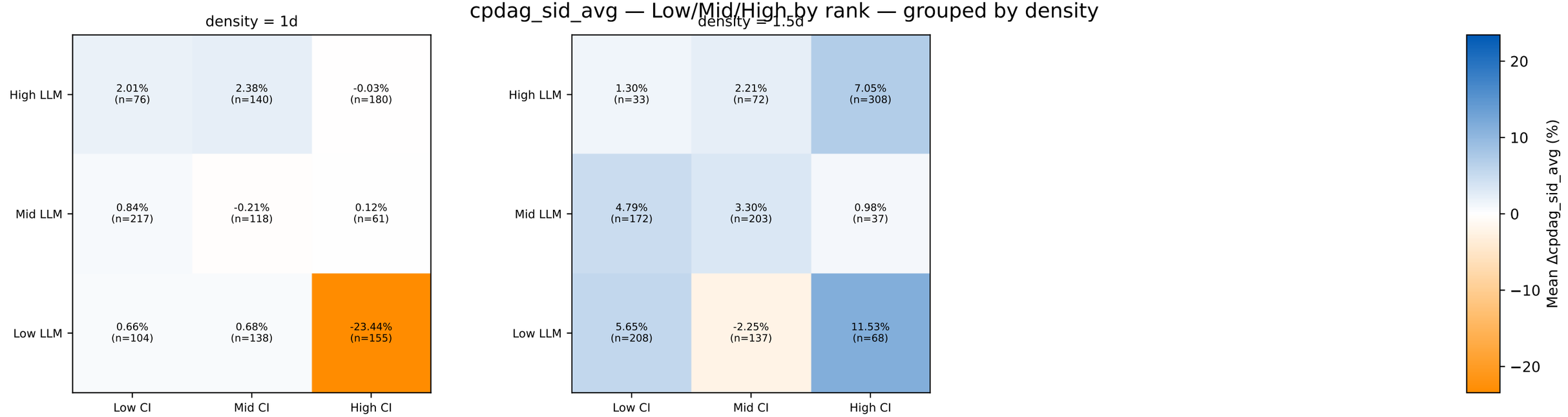

cpdag_F1 — Low/Mid/High by rank — grouped by type
type = ER
type = SF
type = random
High LLM
Mid LLM
Low LLM
Low CI
Mid CI
High CI
0.04 (n=93)
0.01 (n=60)
0.06 (n=119)
0.03 (n=122)
0.02 (n=129)
0.04 (n=21)
0.00 (n=57)
-0.00 (n=83)
-0.07 (n=132)
0.00 (n=34)
0.00 (n=72)
0.00 (n=142)
0.01 (n=95)
0.00 (n=128)
-0.01 (n=24)
0.02 (n=119)
0.01 (n=47)
0.00 (n=82)
0.03 (n=20)
0.07 (n=23)
-0.01 (n=242)
0.04 (n=137)
0.02 (n=128)
0.02 (n=19)
0.02 (n=128)
-0.02 (n=133)
-0.00 (n=24)
Mean Δcpdag_F1
0.06
0.04
0.02
0.00
−0.02
−0.04
−0.06


cpdag_shd — Low/Mid/High by rank — grouped by type
type = ER
type = SF
type = random
High LLM
Mid LLM
Low LLM
Low CI
Mid CI
High CI
0.56%
(n=93)
0.11%
(n=60)
-2.51%
(n=119)
0.93%
(n=122)
1.17%
(n=129)
2.54%
(n=21)
0.08%
(n=57)
0.05%
(n=83)
-8.11%
(n=132)
0.17%
(n=35)
0.03%
(n=71)
0.15%
(n=144)
-0.19%
(n=95)
-0.09%
(n=130)
-1.36%
(n=25)
1.92%
(n=120)
-0.04%
(n=49)
0.00%
(n=81)
0.62%
(n=18)
1.32%
(n=23)
2.68%
(n=246)
0.60%
(n=138)
0.38%
(n=132)
0.45%
(n=17)
0.31%
(n=131)
-1.59%
(n=132)
0.06%
(n=24)
Mean Δcpdag_shd (%)
8
6
4
2
0
−2
−4
−6
−8


cpdag_sid_avg — Low/Mid/High by rank — grouped by type
type = ER
type = SF
type = random
High LLM
Mid LLM
Low LLM
Low CI
Mid CI
High CI
5.99% (n=93)
4.19% (n=60)
8.58% (n=119)
5.07% (n=122)
6.93% (n=129)
13.76% (n=21)
0.12% (n=57)
-0.05% (n=83)
-27.56% (n=132)
0.02% (n=35)
0.03% (n=71)
0.24% (n=144)
0.78% (n=95)
-0.94% (n=130)
-4.74% (n=25)
7.41% (n=120)
0.98% (n=49)
0.00% (n=81)
3.41% (n=18)
7.87% (n=23)
4.42% (n=246)
5.55% (n=138)
2.43% (n=132)
2.18% (n=17)
-0.17% (n=131)
-3.93% (n=132)
0.37% (n=24)
Mean Δcpdag_sid_avg (%)
20
10
0
−10
−20


cpdag_F1 — Low/Mid/High by rank — grouped by density
density = 1d
density = 1.5d
High LLM
Mid LLM
Low LLM
Low CI
Mid CI
High CI
0.01 (n=76)
0.00 (n=140)
-0.00 (n=180)
0.02 (n=217)
0.00 (n=118)
0.00 (n=61)
0.01 (n=103)
0.00 (n=138)
-0.06 (n=155)
0.02 (n=39)
0.04 (n=72)
0.02 (n=297)
0.03 (n=166)
0.03 (n=201)
0.01 (n=41)
0.02 (n=204)
-0.02 (n=135)
0.04 (n=70)
Mean Δcpdag_F1
0.04
0.02
0.00
−0.02
−0.04


cpdag_shd — Low/Mid/High by rank — grouped by density
density = 1d
density = 1.5d
High LLM
Mid LLM
Low LLM
Low CI
Mid CI
High CI
0.36%
(n=76)
0.13%
(n=140)
-0.06%
(n=180)
0.35%
(n=217)
-0.06%
(n=118)
0.01%
(n=61)
0.32%
(n=104)
0.12%
(n=138)
-6.90%
(n=155)
0.56%
(n=33)
0.57%
(n=72)
1.33%
(n=308)
0.53%
(n=172)
0.38%
(n=203)
0.09%
(n=37)
1.20%
(n=208)
-1.65%
(n=137)
1.94%
(n=68)
Mean Δcpdag_shd (%)
6
4
2
0
−2
−4
−6


cpdag_sid_avg — Low/Mid/High by rank — grouped by density
density = 1d
density = 1.5d
High LLM
Mid LLM
Low LLM
Low CI
Mid CI
High CI
2.01%
(n=76)
2.38%
(n=140)
-0.03%
(n=180)
0.84%
(n=217)
-0.21%
(n=118)
0.12%
(n=61)
0.66%
(n=104)
0.68%
(n=138)
-23.44%
(n=155)
1.30%
(n=33)
2.21%
(n=72)
7.05%
(n=308)
4.79%
(n=172)
3.30%
(n=203)
0.98%
(n=37)
5.65%
(n=208)
-2.25%
(n=137)
11.53%
(n=68)
Mean Δcpdag_sid_avg (%)
20
10
0
−10
−20