

Supplementary Table 1. Model diagnostic metrics for topic number selection

K	Held-out likelihood	Residuals	Semantic coherence
13	-6.91	58.99	-167.75
14	-6.88	59.68	-165.03
15	-6.88	55.84	-173.39
16	-6.87	58.04	-175.19
17	-6.88	56.69	-181.14