

Supplementary Materials

Exploring Serial Patterns in Negative Hotel Reviews

The supplementary materials present results from an alternative specification of the Structural Topic Model, based on $K = 32$ individual topics. This analysis complements the main text's findings, which were based on $K = 27$ topics.

Table S1 shows the list of $K=32$ generated topics (ranked according to their importance) together with the most frequent words. The most highly weighted topics are 'Room size', 'Food and dining (1)', 'Check-in waiting times', and 'Communication problems'. Table S1, column four, presents the results of the model selection procedure with respect to lags p , q and l described in the main text. Columns five and six report the results of the Chi-square tests for serial correlation across seasons. Specifically, we test the joint significance of all selected lags in low and, separately, high seasons. Individual coefficients and their significance levels are omitted for brevity. GMM specifications include the same set of explanatory variables as reported in Table 3 (main text). The results suggest that the following nine topics show statistically significant levels of serial correlation both across low and high seasons: 1 (Room size), 9 (Reservation and payment), 11 (Location 1), 16 (Coffee and tea-making facilities), 22 (Wi-fi and internet), 23 (Pool and spa), 27 (Location 2), 31 (Mismanaged orders) and 32 (Misc. facilities). In addition, further five topics show statistically significant results in high seasons only: 3 (Check-in waiting times), 5 (Air quality and temperature), 8 (Spatial aspects and renovation), 12 (Unhelpful and unfriendly staff), and 18 (Parking). Finally, Topic 6 (Bed quality) displays negative serial correlation, or over-reaction to complaints, in low seasons. For the remaining $32 - 15 = 17$ topics however, we find no statistical evidence of serial correlation. These findings are consistent with the interpretation that hotel management effectively addresses the root causes of emergent customer complaints across 17 transient (not serially correlated) topics.

We first note that, similar to the analysis in the main text, the list of serially correlated topics above is heavily influenced by a range of capacity constraints (Topics 1, 5, 8, 11, 18, 22, 23, 27, 32). Arguably, these topics might not be easily remediable in the short term. Perhaps the most valuable insights from this specification concern the other remaining serially correlated topics in the list: 'Check-in waiting times' (high seasons), 'Reservation and payment' (all seasons), 'Unhelpful and unfriendly staff', (high seasons), 'Mismanaged orders' (all seasons), where we find indications of potential inefficiencies.

Overall, the findings reported in the main body of the paper for $K = 27$ should be viewed in conjunction with the findings reported for $K = 32$. The two alternative specifications produce complementary results. In particular, both specifications highlight potential inefficiencies with respect to 'Unhelpful and unfriendly staff'. On the other hand, the two specifications differ on their findings with respect to complaints about reservation and payment. More specifically, while specification $K = 27$ defines 'Reservation' and 'Payment' as two separate topics and finds them individually transient, specification $K = 32$ merges the two into one 'Reservation and Payment' topic, which it then reports as serially correlated. Finally, the serially correlated topic 'Mismanaged orders' ($K = 32$) might complement the findings of serially correlated 'Unresolved requests over multiple days' and 'Slow and overpriced services' ($K = 27$).

Table S1. Topics and tests for serial correlation in monthly average Topic Relevance.

Topic	Label	Top words	p, q, l – autoregressive lags selected based on the Model Selection Criterion using the Hannan-Quinn Information Criterion	Low season Chi-square	High season Chi-square
1	Room size	Room, small, size, tini, standard	4, 1, 0	89.93***	113.44***
2	Food and dining (1)	Breakfast, food, include, tabl, limit	1, 1, 0	0.23	1.31
3	Check-in waiting times	Noth, time, realli, checkin, wait	1, 3, 0	0.76	12.54***
4	Communication problems	Told, said, arriv, ask, card	1, 1, 0	0.12	0.75
5	Air quality & temperature	Work, window, air, open, hot	1, 1, 0	3.79*	6.44**
6	Bed quality	Bed, two, double, hard, pillow	1, 3, 0	6.66***	1.61
7	Bathroom experience, cleanliness	Bathroom, clean, smell, dirti, bedroom	2, 1, 0	0.41	4.91*
8	Spatial aspects & renovation	Need, view, look, old, light	1, 6, 0	0.04	13.59**
9	Reservation and payment	Book, service, even, charge, pay	4, 1, 0	30.36***	36.04***
10	General experience (1)	Hotel, star, guest, travel, rate	1, 1, 1	0.88	0.94
11	Location (1)	Can, walk, around, minut, take	4, 1, 0	20.32***	53.16***
12	Unhelpful or unfriendly staff	Staff, recept, help, friend, rude	5, 1, 0	8.20	14.23**
13	Noise (1)	Door, floor, next, morn, sleep	1, 2, 0	0.50	0.65

14	Pricey drinks and busy lounge areas	Price, bar, area, drink, busy	2, 1, 0	0.60	0.67
15	Noise (2)	Room, night, first, noise, made	1, 1, 0	0.00	3.64*
16	Coffee, tea-making and complimentary items	Use, coffee, make, facility, tea,	3, 1, 0	14.53***	21.42***
17	Shower & toilet, maintenance	Shower, water, toilet, bath, broken	3, 1, 0	1.46	0.63
18	Parking	Park, noise, lot, place, find	1, 1, 0	0.42	4.07**
19	Value and expectations	Bit, good, bad, quit, cost	1, 1, 0	1.31	0.50
20	Food and dining (2)	Little, expensive, restaurant, much, money	1, 2, 0	0.07	0.01
21	Misc. services	One, back, lift, come, front	1, 2, 0	0.00	5.74*
22	Wi-fi & internet	Poor, wifi, slow, free, quality	5, 1, 0	27.14***	32.40***
23	Pool and spa	Pool, thing, felt, spa, swim	6, 1, 0	202.14***	466.55***
24	NEUTRAL/POSITIVE	Like, just, expect, well, feel	1, 1, 0	0.06	0.71
25	Unmet requests	Get, ask, want, took, never	1, 1, 0	0.05	0.40
26	General experience (2)	Stay, great, perfect, long, thought	1, 4, 0	4.60**	8.50*
27	Location (2)	Location, far, city, main, road	5, 1, 0	67.17***	84.45***
28	Misc	Better, everything, change, elevator, especially	2, 1, 0	1.42	4.88*
29	Housekeeping services	Day, although, keep, housekeeper, done	1, 4, 0	0.56	7.79*
30	Misc	Also, seem, way, found, mani	1, 1, 0	0.10	1.44

31	Mismanaged orders	Got, people, order, however, given	3, 1, 0	19.93***	32.66***
32	Misc. facilities	Outsid, offer, lobbi, complain, right	8, 3, 0	32.58***	65.30***

Notes: a) We implemented the **Structural Topic Model** uses stemming for word reduction in text processing. Words like "locate" and "location" share the same root "locat". Stemming ensures that such variations are treated as the same word, improving the accuracy and focus of topics. b) * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$