

**Understanding Online Hotel Reviews through Automated Text Analysis**  
**Electronic Supplement**

**Table 1: Multicollinearity Assessment. Generalized variance inflation factors after adjusting for degrees of freedom<sup>a</sup>, for each of the regression models.**

	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7	Model 8
Property Dummy	1.00	1.00	1.00	1.00	1.00	1.00	1.01	1.01
log(Word Count)	1.70	1.93	1.88	1.94	1.94	1.95	2.12	2.17
Sentiment-	1.37	2.36	1.52	1.53	1.53	1.54	1.57	4.61
Sentiment+	1.47	2.30	2.30	2.30	2.30	2.31	2.44	4.66
Sentiment-^2		2.07						
Sentiment+^2		1.86						
Sentiment-^2 + Sentiment+^2			1.88	1.88	1.88	1.88	1.91	2.09
Focus				1.14	1.14	1.14	1.29	1.37
Flesch-Kincaid Grade Level					1.00	1.00	1.00	1.00
Traveler Type:						1.03	1.03	1.03
<b>Topics:</b>								
Topic 1 (Experience)							1.16	2.03
Topic 2 (Location)							1.25	2.15
Topic 3 (Transactions)							1.23	2.19
Topic 4 (Amenities)							1.18	1.98
Topic 5 (Value)							1.36	2.28
<b>Sentiment- x Topic:</b>								
Sentiment- x Topic 1								2.18
Sentiment- x Topic 2								1.74
Sentiment- x Topic 3								3.17
Sentiment- x Topic 4								2.19
Sentiment- x Topic 5								1.61
<b>Sentiment+ x Topic:</b>								
Sentiment+ x Topic 1								2.88
Sentiment+ x Topic 2								2.52
Sentiment+ x Topic 3								3.00
Sentiment+ x Topic 4								2.55
Sentiment+ x Topic 5								2.52

<sup>a</sup> Generalized variance inflation factors (GVIF) were calculated using the Fox-Monette (1992) method. Reported results in the table are after adjustment for degrees of freedom (df) using the formula  $GVIF^{1/(2 * df)}$  as suggested by Fox and Monette (1992).

**Table 2: Robustness Check: Summary of common readability measures**

	Flesch-Kincaid Readability	SMOG Readability	FOG Readability	ARI Readability
Min	1.575	3.129	1.943	-3.456
25% Quartile	7.318	9.236	9.021	7.175
Median	8.728	10.290	10.793	8.832
Mean	9.053	10.503	11.142	9.241
75% Quartile	10.297	11.472	12.733	10.651
Max	41.256	26.683	43.873	50.829

**Table 3: Robustness Check: Using the SMOG Readability Index**

	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7	Model 8
Intercept	3.240 (0.044)***	3.147 (0.041)***	3.041 (0.049)***	3.049 (0.047)***	3.049 (0.047)***	2.982 (0.051)***	3.138 (0.061)***	3.111 (0.075)***
Log(word count)	-0.203 (0.026)***	-0.223 (0.021)***	-0.270 (0.024)***	-0.231 (0.022)***	-0.232 (0.022)***	-0.239 (0.022)***	-0.116 (0.020)***	-0.086 (0.020)***
(Sentiment)-	-0.113 (0.006)***	-0.150 (0.006)***	-0.098 (0.006)***	-0.093 (0.006)***	-0.093 (0.006)***	-0.091 (0.006)***	-0.076 (0.006)***	-0.091 (0.012)***
(Sentiment)+	0.067 (0.004)***	0.091 (0.005)***	0.088 (0.005)***	0.086 (0.004)***	0.086 (0.004)***	0.085 (0.004)***	0.069 (0.004)***	0.080 (0.007)***
[(Sentiment)-] <sup>2</sup>		0.003 (0.000)***						
[(Sentiment)+] <sup>2</sup>		-0.001 (0.000)***						
[(Sentiment)-] <sup>2</sup> + [(Sentiment)+] <sup>2</sup>			-0.000 (0.000)***	-0.000 (0.000)***	-0.000 (0.000)***	-0.000 (0.000)***	-0.000 (0.000)***	-0.000 (0.000)***
Focus				-0.092 (0.013)***	-0.092 (0.013)***	-0.093 (0.013)***	-0.104 (0.013)***	-0.090 (0.015)***
SMOG Readability Index					0.015 (0.011)	0.014 (0.011)	0.013 (0.011)	0.014 (0.011)
Traveler Type:								
couples						0.153 (0.029)***	0.117 (0.026)***	0.114 (0.025)***
family						0.190 (0.046)***	0.138 (0.043)**	0.135 (0.041)***
friends						0.096 (0.038)*	0.063 (0.036)	0.068 (0.036)
solo						0.118 (0.043)**	0.106 (0.039)**	0.091 (0.038)*
unspecified						0.059 (0.034)	0.053 (0.036)	0.049 (0.036)
Topics:								
Topic 1 (amenities)							-0.068 (0.021)**	-0.006 (0.044)
Topic 2 (location)							0.121 (0.021)***	0.182 (0.035)***
Topic 3 (transactions)							-0.122 (0.026)***	-0.254 (0.037)***

Topic 4 (value)							-0.177 (0.025)***	-0.268 (0.040)***
Topic 5 (experience)							0.306 (0.024)***	0.571 (0.046)***
<b>(Sentiment)- x Topics:</b>								
(Sentiment)- x Topic 1								0.027 (0.006)***
(Sentiment)- x Topic 2								0.030 (0.011)**
(Sentiment)- x Topic 3								-0.013 (0.007)*
(Sentiment)- x Topic 4								-0.009 (0.009)
(Sentiment)- x Topic 5								-0.008 (0.009)
<b>(Sentiment)+ x Topics:</b>								
(Sentiment)+ x Topic 1								-0.013 (0.004)**
(Sentiment)+ x Topic 2								-0.016 (0.004)***
(Sentiment)+ x Topic 3								0.016 (0.004)***
(Sentiment)+ x Topic 4								0.010 (0.003)**
(Sentiment)+ x Topic 5								-0.024 (0.004)***
<b>R<sup>2</sup></b>	0.435	0.454	0.442	0.45	0.45	0.455	0.489	0.509
<b>Adjusted R<sup>2</sup></b>	0.429	0.448	0.437	0.444	0.444	0.448	0.483	0.502

Note: \*\*\* : p-value < 0.001, \*\*: p-value < 0.01, \*: p-value < 0.05.

**Table 4: Robustness Check: Using the FOG Readability Index**

	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7	Model 8
Intercept	3.240 (0.044)***	3.147 (0.041)***	3.041 (0.049)***	3.049 (0.047)***	3.049 (0.047)***	2.982 (0.051)***	3.138 (0.062)***	3.110 (0.075)***
Log(word count)	-0.203 (0.026)***	-0.223 (0.021)***	-0.270 (0.024)***	-0.231 (0.022)***	-0.232 (0.022)***	-0.239 (0.022)***	-0.116 (0.020)***	-0.086 (0.020)***
(Sentiment) <sup>-</sup>	-0.113 (0.006)***	-0.150 (0.006)***	-0.098 (0.006)***	-0.093 (0.006)***	-0.093 (0.006)***	-0.091 (0.006)***	-0.076 (0.006)***	-0.091 (0.012)***
(Sentiment) <sup>+</sup>	0.067 (0.004)***	0.091 (0.005)***	0.088 (0.005)***	0.086 (0.004)***	0.086 (0.004)***	0.085 (0.004)***	0.069 (0.004)***	0.080 (0.007)***
[(Sentiment) <sup>-</sup> ] <sup>2</sup>		0.003 (0.000)***						
[(Sentiment) <sup>+</sup> ] <sup>2</sup>		-0.001 (0.000)***						
[(Sentiment) <sup>-</sup> ] <sup>2</sup> + [(Sentiment) <sup>+</sup> ] <sup>2</sup>			-0.000 (0.000)***	-0.000 (0.000)***	-0.000 (0.000)***	-0.000 (0.000)***	-0.000 (0.000)***	-0.000 (0.000)***
Focus				-0.092 (0.013)***	-0.092 (0.013)***	-0.093 (0.013)***	-0.104 (0.013)***	-0.090 (0.015)***
Gunning Fog Index					0.013 (0.011)	0.012 (0.011)	0.011 (0.011)	0.012 (0.011)
Traveler Type:								
couples						0.153 (0.029)***	0.117 (0.026)***	0.114 (0.026)***
family						0.190 (0.046)***	0.138 (0.043)**	0.135 (0.041)***
friends						0.096 (0.038)*	0.063 (0.036)	0.068 (0.036)
solo						0.118 (0.043)**	0.106 (0.039)**	0.092 (0.038)*
unspecified						0.059 (0.034)	0.053 (0.036)	0.049 (0.036)
Topics:								
Topic 1 (amenities)							-0.068 (0.021)**	-0.006 (0.044)
Topic 2 (location)							0.121 (0.021)***	0.182 (0.035)***

Topic 3 (transactions)							-0.122 (0.026)***	-0.253 (0.037)***
Topic 4 (value)							-0.177 (0.025)***	-0.268 (0.040)***
Topic 5 (experience)							0.306 (0.024)***	0.571 (0.046)***
<b>(Sentiment)- x Topics:</b>								
(Sentiment)- x Topic 1								0.027 (0.006)***
(Sentiment)- x Topic 2								0.030 (0.011)**
(Sentiment)- x Topic 3								-0.013 (0.007)*
(Sentiment)- x Topic 4								-0.009 (0.009)
(Sentiment)- x Topic 5								-0.008 (0.009)
<b>(Sentiment)+ x Topics:</b>								
(Sentiment)+ x Topic 1								-0.013 (0.004)**
(Sentiment)+ x Topic 2								-0.016 (0.004)***
(Sentiment)+ x Topic 3								0.016 (0.004)***
(Sentiment)+ x Topic 4								0.010 (0.003)**
(Sentiment)+ x Topic 5								-0.024 (0.004)***
<b>R<sup>2</sup></b>	0.435	0.454	0.442	0.45	0.45	0.454	0.489	0.509
<b>Adjusted R<sup>2</sup></b>	0.429	0.448	0.437	0.444	0.444	0.448	0.483	0.502

Note: \*\*\* : p-value < 0.001, \*\*: p-value < 0.01, \*: p-value < 0.05.

**Table 5: Robustness Check: Using the Automated Readability Index**

	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7	Model 8
Intercept	3.240 (0.044)***	3.147 (0.041)***	3.041 (0.049)***	3.049 (0.047)***	3.048 (0.047)***	2.981 (0.051)***	3.136 (0.062)***	3.109 (0.076)***
Log(word count)	-0.203 (0.026)***	-0.223 (0.021)***	-0.270 (0.024)***	-0.231 (0.022)***	-0.232 (0.022)***	-0.239 (0.022)***	-0.116 (0.020)***	-0.086 (0.020)***
(Sentiment)-	-0.113 (0.006)***	-0.150 (0.006)***	-0.098 (0.006)***	-0.093 (0.006)***	-0.093 (0.006)***	-0.091 (0.006)***	-0.076 (0.006)***	-0.091 (0.012)***
(Sentiment)+	0.067 (0.004)***	0.091 (0.005)***	0.088 (0.005)***	0.086 (0.004)***	0.086 (0.004)***	0.085 (0.004)***	0.069 (0.004)***	0.080 (0.007)***
[(Sentiment)-] <sup>2</sup>		0.003 (0.000)***						
[(Sentiment)+] <sup>2</sup>		-0.001 (0.000)***						
[(Sentiment)-] <sup>2</sup> + [(Sentiment)+] <sup>2</sup>			-0.000 (0.000)***	-0.000 (0.000)***	-0.000 (0.000)***	-0.000 (0.000)***	-0.000 (0.000)***	-0.000 (0.000)***
Focus				-0.092 (0.013)***	-0.092 (0.013)***	-0.093 (0.013)***	-0.104 (0.013)***	-0.090 (0.015)***
Automated Readability Index					0.009 (0.011)	0.009 (0.011)	0.010 (0.011)	0.011 (0.011)
Traveler Type:								
couples						0.154 (0.029)***	0.117 (0.026)***	0.114 (0.025)***
family						0.191 (0.046)***	0.139 (0.043)**	0.136 (0.041)***
friends						0.096 (0.038)*	0.063 (0.036)	0.068 (0.036)
solo						0.117 (0.043)**	0.106 (0.039)**	0.091 (0.038)*
unspecified						0.060 (0.034)	0.054 (0.036)	0.049 (0.036)
Topics:								
Topic 1 (amenities)							-0.067 (0.021)**	-0.006 (0.044)
Topic 2 (location)							0.121 (0.021)***	0.183 (0.035)***
Topic 3 (transactions)							-0.122 (0.026)***	-0.253 (0.037)***

Topic 4 (value)							-0.176 (0.025) <sup>***</sup>	-0.268 (0.040) <sup>***</sup>
Topic 5 (experience)							0.307 (0.025) <sup>***</sup>	0.572 (0.046) <sup>***</sup>
<b>(Sentiment)- x Topics:</b>								
(Sentiment)- x Topic 1								0.027 (0.006) <sup>***</sup>
(Sentiment)- x Topic 2								0.030 (0.011) <sup>**</sup>
(Sentiment)- x Topic 3								-0.013 (0.007) <sup>*</sup>
(Sentiment)- x Topic 4								-0.009 (0.009)
(Sentiment)- x Topic 5								-0.008 (0.009)
<b>(Sentiment)+ x Topics:</b>								
(Sentiment)+ x Topic 1								-0.013 (0.004) <sup>**</sup>
(Sentiment)+ x Topic 2								-0.016 (0.004) <sup>***</sup>
(Sentiment)+ x Topic 3								0.016 (0.004) <sup>***</sup>
(Sentiment)+ x Topic 4								0.010 (0.003) <sup>**</sup>
(Sentiment)+ x Topic 5								-0.024 (0.004) <sup>***</sup>
<b>R<sup>2</sup></b>	0.435	0.454	0.442	0.45	0.45	0.454	0.489	0.509
<b>Adjusted R<sup>2</sup></b>	0.429	0.448	0.437	0.444	0.444	0.448	0.483	0.502

Note: \*\*\* : p-value < 0.001, \*\*: p-value < 0.01, \*: p-value < 0.05.

## References:

Fox J., & Monette, G. 1992. Generalized Collinearity Diagnostics. *Journal of the American Statistical Association*, 87(417), 178-183.