**Scheduling Content on Social Media: Theory, Evidence, and Application**

**Vamsi K. Kanuri, Yixing Chen, and Shrihari Sridhar**

**WEB APPENDIX**

WEB APPENDIX W1: An Overview of How Working Memory Influences Social Media Content Engagement across Dayparts

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Biological Processes Affecting Social Media Content Engagement | Day Parts within a Day | | | Implications for Social Media Content Engagement |
| Morning | Afternoon | Evening |
| Working memory capacity | High | Low | Moderate | Engagement with content will likely be highest in the morning (due to the availability of working memory), moderate in the evening, and lowest in the afternoon (H1a, H1b, H1c) |
| Activation of inhibitory processes due to constrained working memory | Low | High | Moderate |  |
| Inhibition of information that can further hinder working memory (e.g., high-arousal emotions) | Low | High | Moderate | Engagement with high-arousal emotions will likely be highest in the morning (due to the absence of inhibitory processes), moderate in the evening and least in the afternoon (H2a, H2b, H2c) |
| Distraction from the focal cognitive task due to the availability of working memory | High | Low | Moderate | Engagement with content that requires higher cognitive capabilities will be highest in the afternoon (due to minimal distraction), moderate in the evening and lowest in the morning (H3a, H3b, H3c) |
| Preferential treatment of favorable information due to external cues such as TCA | Low | High | Moderate | TCA on social media will likely be most effective in garnering engagement in the afternoon (because TCA serves as the external cue to trigger preferential processing of information), moderately effective in the evening and least effective in the morning. (H4a, H4b, H4c) |

WEB APPENDIX W2: Content Type Dictionary

|  |  |  |
| --- | --- | --- |
| **Construct** | **Notation** | **Sample Words and Word Stems** |
| High-arousal Positive Emotions  (e.g., awe, amusement) | Posemo\_arousali | * **words**: awesome, cheer/cheerful, delicious, excite/excitement/exciting, gorgeous, joy, paradise, satisfy, sexy, yummy * **word stems**: amaze, entertain, enthuse, fantasi, haha, humor, inspire, party, pleasur, succeed |
| High-arousal Negative Emotions  (e.g., anxiety, anger)  *Anxiety* | Negemo\_arousali | * **words**: afraid, anxiety/anxious/anxiously, distress, fear(s)/fearful/, guilt/guilty, horrible/horribly, nervous, scare, shaky, tense, terrify, terror, worry, worries * **word stems**: aversi, ahsam, doubt, irrita, panic, startl, struggle, vulnerab |
| *Anger* |  | * **words**: aggress(ed,es)/aggressive, angry, bastard, crap, cruel, dumb, hate, hell, jerk, madder, offense, protest, rape, temper, war, weapon, yell * **word stems**: abusi, anger, attack, battl, blam, btutal, critici, enemy, envie, hostil, insult, piss, offend, rage, rebel, threat |
| Cognitive Processing  (e.g., insight, causation)  *Insight* | Cog\_processi | * **words**: become/became closure, clue, complex/complexity, definition, sense, wisdom * **word stems**: acknowledg, analy, clarif, diagnos, disclo, relate, recogni, skeptic, |
| *Causation* |  | * **words**: affect, because, effect, factor, hence, infer, origin(s), purpose, rational * **word stems**: attribut, caus, comply, enabl, enact, stimul, motiv, originat, therefor |

WEB APPENDIX W3: Summary Statistics

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Mean | SD | Min | Max |
| Night Dummy (0:00-5:59) | .022 | - | 0 | 1 |
| Morning Dummy (6:00-11:59) | .358 | - | 0 | 1 |
| Afternoon Dummy (12:00-17:59) | .421 | - | 0 | 1 |
| Evening Dummy (18:00-23:59) | .199 | - | 0 | 1 |
| TCA (1 = yes) | .091 | - | 0 | 1 |
| High-arousal Negative Emotions (message) | .466 | 1.850 | 0 | 50 |
| High-arousal Positive Emotions (message) | 3.068 | 6.003 | 0 | 100 |
| Cognitive Processing (message) | 7.499 | 7.606 | 0 | 66.670 |
| Message Length | 21.261 | 11.962 | 0 | 130 |
| Message Readability (FOG index) | 9.685 | 5.533 | 0 | 37.086 |
| Inter-post Duration | 67.666 | 118.479 | 0 | 953 |
| Local News Dummy | .302 | - | 0 | 1 |
| Business News Dummy | .085 | - | 0 | 1 |
| Sports News Dummy | .180 | - | 0 | 1 |
| Entertainment News Dummy | .094 | - | 0 | 1 |
| Life News Dummy | .153 | - | 0 | 1 |
| Opinion Dummy | .024 | - | 0 | 1 |
| National News Dummy | .091 | - | 0 | 1 |
| Organic Reach | 18705.870 | 12031.760 | 0 | 173043 |
| Log (Organic Reach) | 9.621 | .703 | 0 | 12.061 |
| Link Clicks | 966.737 | 4413.677 | 0 | 152448 |
| Log (Link Clicks) | 5.452 | 1.553 | 0 | 11.935 |

WEB APPENDIX W4: Correlations

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  |  | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
| 1 | High-arousal Positive Emotions (message) | 1.000 |  |  |  |  |  |  |  |
| 2 | High-arousal Negative Emotions (message) | -.026\* | 1.000 |  |  |  |  |  |  |
| 3 | Cognitive Processing (message) | .001 | .013 | 1.000 |  |  |  |  |  |
| 4 | Message Length | -.152\* | .012 | -.030\* | 1.000 |  |  |  |  |
| 5 | Message Readability | -.090\* | .059\* | -.038\* | .328\* | 1.000 |  |  |  |
| 6 | Inter-post duration | -.025 | -.001 | -.005 | .083\* | .044\* | 1.000 |  |  |
| 7 | Log(Organic Reach) | -.052\* | .053\* | .024 | -.021 | -.030\* | .037\* | 1.000 |  |
| 8 | Log(Link Clicks) | -.034\* | .045\* | .036\* | -.088\* | -.049\* | .046\* | .814\* | 1.000 |

\* Denotes correlations are significant at 95% level.

WEB APPENDIX W5: Identification Strategy

|  |  |  |  |
| --- | --- | --- | --- |
| **Challenge** | **Source of Endogeneity** | **Identification Strategy** | **Rationale** |
| Strategic Allocation of Posts | Non-random selection (relevance-based) in showing stories in users’ news feed | Control for organic reach and a set of fixed effects | The selection of delivery of messages is captured by the number of unique users who see the story for free (i.e., organic reach) |
| Time of Day | Social media managers choose timing strategically based on private knowledge (e.g., expected number of clicks in each daypart), which affects the probability of clicking | Control function: average number of breaking news tweets by CNN and Associated Press as the excluded variable | Breaking news events are external exogenous shocks (e.g., terror attacks) as they are likely uncorrelated with the anticipated link clicks of a news story originally planned in the given daypart |
| Content Type: High-arousal Emotions | Social media managers draft message for a higher probability of clicking since correlated unobservables drive both content tone and social media post performance | Control function: level of high-arousal emotions in story description as the excluded variable | Story description of the original article is not written by social media managers who might have expectations when engaging their audience but exogenously given by journalists or editors to social media managers. |
| Content Type: Cognition Processing | Social media managers draft message for a higher probability of clicking since correlated unobservables drive both content tone and social media post performance | Control function: level of cognition processing required in story description as the excluded variable | Story description of the original article is not written by social media managers who might have expectations when engaging their audience but exogenously given by journalists or editors to social media managers. |
| TCA | Social media managers make TCA decisions for a higher probability of clicking | Latent instrumental  variables (LIV) approach | LIV by definition is uncorrelated with the error term in the main regression |

WEB APPENDIX W6: Twitter Data

We collect the Twitter data from the Associated Press (AP; <https://twitter.com/ap>) and CNN Breaking News (CNNBN; <https://twitter.com/cnnbrk>). Using Twitter Advanced Search, we restrict our data to tweets posted by AP and CNNBN, excluding retweets, from December 31, 2014 to December 31, 2015. We collect tweet id, link, time stamp, engagement metrics (i.e., replies, retweets, likes), and Twitter message. In total, we collect 25,861 tweets from AP and 6,502 tweets from CNNBN. While we consider all tweets posted by CNNBN as breaking news, we identified AP breaking news tweets by looking for the keyword “BREAKING” (see Figure W6-1 below). Indeed, “breaking” tweets on average garner a higher number of replies, retweets, and likes than standard tweets (*p* < .01, see Table W6-1). The distribution of breaking news tweets from AP and CNNBN at the hourly and day-part level can be seen in Figure W6-2.

TABLE W6-1: Engagement Metrics across Two Groups of AP Tweets

FIGURE W6-1: AP Breaking News Tweet AP



|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Reply | Retweet | Like | # of Posts |
| AP Standard Tweets | 11.0  (14.8) | 86.9  (84.7) | 47.5  (49.2) | 22,406 |
| AP Breaking Tweets | 24.6  (35.0) | 337.2  (693.3) | 141.0  (403.5) | 3,455 |

Note. Standard deviations in parentheses.

FIGURE W6-2: Number of Breaking Tweets by CNNBN and AP

WEB APPENDIX W7: Illustration of Breaking News Tweets and Social Media Schedule

**Typical Schedule of Collaborating Newspaper**

**(January 23, 2015, no breaking news)**

**5:45 am, Story from Life News**

**(typically seen in the Night Daypart)**

8:15 am, Story from Sports News

1:30 pm, Story from the Local News

9:46 am, Story from Local News

11:10 am, Story from Business News

12:29 pm, Story from Local News

2:30 pm, Story from Life News

7:41 am, Story from Sports News

**Schedule of Collaborating Newspaper**

**(January 2, 2015)**

**Breaking News Tweets**



**6:00 am, Story from Breaking News**

**8:53 am, Story from Life News**

**(was pushed down to Morning)**

1:30 pm, Story from Local News

9:16 am, Story from Sports News

10:34 am, Story from Local News

11:37 am, Story from Business News

12:04 pm, Story from Sports News

3:00 pm, Story from Entertainment news

WEB APPENDIX W8: First Stage Results via Control Function Approach

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  | Night  (probit) | Afternoon  (probit) | Evening  (probit) | Negative emotions (message) | Positive emotions (message) | Cognitive processing (message) | Inter-post duration |
| Breaking Tweets (daypart) | -.0763\*\*\* | .0112\*\* | -.0220\*\*\* |  |  |  |  |
|  | (.0281) | (.00517) | (.00787) |  |  |  |  |
|  |  |  |  |  |  |  |  |
| Negative Emotions (description) |  |  |  | .288\*\*\* |  |  |  |
|  |  |  |  | (.0167) |  |  |  |
|  |  |  |  |  |  |  |  |
| Positive Emotions (description) |  |  |  |  | .211\*\*\* |  |  |
|  |  |  |  |  | (.0223) |  |  |
|  |  |  |  |  |  |  |  |
| Cognitive Processing (description) |  |  |  |  |  | .201\*\*\* |  |
|  |  |  |  |  |  | (.0176) |  |
|  |  |  |  |  |  |  |  |
| Breaking Tweets (lag hourly) |  |  |  |  |  |  | -3.839\*\*\* |
|  |  |  |  |  |  |  | (1.256) |
|  |  |  |  |  |  |  |  |
| Log (Organic Reach) | .217\*\*\* | -.0543\*\* | .133\*\*\* | .112\*\*\* | -.241\*\* | .117 | 3.768 |
|  | (.0735) | (.0256) | (.0304) | (.0353) | (.115) | (.147) | (2.328) |
|  |  |  |  |  |  |  |  |
| Inter-post Duration | .0204\*\*\* | -.00608\*\*\* | .0145\*\*\* | -.000263 | .00295 | .0000798 |  |
|  | (.00136) | (.000691) | (.00104) | (.000795) | (.00258) | (.00331) |  |
|  |  |  |  |  |  |  |  |
| Inter-post Duration2 | -.0000322\*\*\* | .00000384\*\*\* | -.0000362\*\*\* | .000000334 | -.00000496 | -.000000791 |  |
|  | (.00000246) | (.00000116) | (.00000372) | (.00000119) | (.00000388) | (.00000497) |  |
|  |  |  |  |  |  |  |  |
| Message Length | -.0122\*\*\* | -.00716\*\*\* | .00822\*\*\* | -.00182 | -.0587\*\*\* | .00325 | .784\*\*\* |
|  | (.00454) | (.00166) | (.00184) | (.00223) | (.00725) | (.00929) | (.146) |
|  |  |  |  |  |  |  |  |
| Message Readability | -.00177 | .000552 | .0119\*\*\* | .0175\*\*\* | -.0260\* | -.0645\*\*\* | .176 |
|  | (.00901) | (.00341) | (.00388) | (.00466) | (.0151) | (.0194) | (.306) |
|  |  |  |  |  |  |  |  |
| Intercept | -4.704\*\*\* | .964\*\*\* | -2.853\*\*\* | -.812\*\* | 6.487\*\*\* | 3.994\*\*\* | 13.20 |
|  | (.764) | (.268) | (.312) | (.358) | (1.163) | (1.490) | (24.23) |
| Day-of-week, news topic, month | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| N | 5706 | 5706 | 5706 | 5706 | 5706 | 5706 | 5705 |

Standard errors in parentheses; \* *p <*.1, \*\* *p <* .05, \*\*\* *p <* .01

WEB APPENDIX W9: Robustness Checks

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | (1) | (2) | (3) | (4) |
|  | Alternative daypart definitions | Correct for lag error | Correct for inter-post duration | Use impression instead of reach |
| Night | -.138\*\*\* | -.117 | -.106 | -.0566 |
|  | (.0265) | (.105) | (.106) | (.0949) |
|  |  |  |  |  |
| Afternoon | -.100\*\*\* | -.103\*\*\* | -.103\*\*\* | -.126\*\*\* |
|  | (.0149) | (.0139) | (.0146) | (.0136) |
|  |  |  |  |  |
| Evening | -.148\*\*\* | -.150\*\*\* | -.150\*\*\* | -.167\*\*\* |
|  | (.0377) | (.0365) | (.0361) | (.0374) |
|  |  |  |  |  |
| TCA | 2.096\*\*\* | 2.119\*\*\* | 2.103\*\*\* | 2.557\*\*\* |
|  | (.198) | (.199) | (.197) | (.195) |
|  |  |  |  |  |
| Negative Emotions (Message) | .0351 | .0332 | .0351 | .0345 |
|  | (.0218) | (.0220) | (.0217) | (.0198) |
|  |  |  |  |  |
| Positive Emotions (Message) | .0277\*\* | .0285\*\* | .0287\*\* | .0289\*\* |
|  | (.00984) | (.0102) | (.0104) | (.0111) |
|  |  |  |  |  |
| Cognitive Processing Message) | .0131 | .0135 | .0137 | .0132 |
|  | (.0107) | (.0104) | (.0107) | (.0110) |
|  |  |  |  |  |
| Night × TCA | -.367 | -.232\* | -.219\* | -.496\* |
|  | (.323) | (.0999) | (.102) | (.223) |
|  |  |  |  |  |
| Afternoon × TCA | .371\*\*\* | .376\*\*\* | .370\*\*\* | .350\*\*\* |
|  | (.0809) | (.0806) | (.0813) | (.0708) |
|  |  |  |  |  |
| Evening × TCA | -.146 | -.175 | -.187 | -.185 |
|  | (.228) | (.248) | (.260) | (.262) |
|  |  |  |  |  |
| Night × negative emotions (message) | .0134 | -.0807\*\* | -.0821\*\* | -.0702\*\* |
|  | (.0314) | (.0243) | (.0244) | (.0226) |
|  |  |  |  |  |
| Afternoon × negative emotions (message) | -.0150\* | -.0151\*\* | -.0149\* | -.0120\* |
|  | (.00642) | (.00599) | (.00640) | (.00612) |
|  |  |  |  |  |
| Evening × negative emotions (message) | -.0350\* | -.0244 | -.0237 | -.0229 |
|  | (.0154) | (.0145) | (.0150) | (.0144) |
|  |  |  |  |  |
| Night × positive emotions (message) | .00154 | .00294 | .00290 | .00202 |
|  | (.000834) | (.00241) | (.00233) | (.00197) |
|  |  |  |  |  |
| Afternoon × positive emotions (message) | -.00116 | -.00152 | -.00136 | -.00128 |
|  | (.00298) | (.00321) | (.00309) | (.00294) |
|  |  |  |  |  |
| Evening × positive emotions (message) | .00107 | -.000344 | -.000269 | -.000279 |
|  | (.00378) | (.00401) | (.00385) | (.00391) |
|  |  |  |  |  |
| Night × cognitive processing (message) | .00530 | .00268 | .00262 | .00354 |
|  | (.00334) | (.00384) | (.00399) | (.00309) |
|  |  |  |  |  |
| Afternoon × cognitive processing (message) | .00510\*\* | .00535\*\* | .00531\*\* | .00526\*\* |
|  | (.00155) | (.00164) | (.00165) | (.00158) |
|  |  |  |  |  |
| Evening × cognitive processing (message) | .00686\* | .00704\* | .00702\* | .00768\* |
|  | (.00329) | (.00313) | (.00309) | (.00323) |
|  |  |  |  |  |
| Log (Organic Reach) | 1.740\*\*\* | 1.743\*\*\* | 1.749\*\*\* |  |
|  | (.0630) | (.0641) | (.0649) |  |
|  |  |  |  |  |
| Lag Error Term |  | .0447\* |  |  |
|  |  | (.0183) |  |  |
|  |  |  |  |  |
| Inter-post Duration (residuals) |  |  | .00191\* |  |
|  |  |  | (.000835) |  |
|  |  |  |  |  |
| Log (Impression) |  |  |  | 1.727\*\*\* |
|  |  |  |  | (.0696) |
|  |  |  |  |  |
| Message Length | -.00317\* | -.00313\* | -.00153 | -.00395\* |
|  | (.00144) | (.00151) | (.00160) | (.00173) |
|  |  |  |  |  |
| Message Readability | -.00521\*\* | -.00538\* | -.00529\* | -.00518\* |
|  | (.00194) | (.00223) | (.00231) | (.00234) |
|  |  |  |  |  |
| Inter-post Duration | -.00152 | -.00125 | -.00299 | -.00168 |
|  | (.00113) | (.00106) | (.00168) | (.00106) |
|  |  |  |  |  |
| Inter-post Duration2 | .0000011 | .0000008 | .0000008 | .000001 |
|  | (.0000007) | (.0000006) | (.0000007) | (.0000007) |
|  |  |  |  |  |
| TCA (LIV\_error term) | .110 | .0951 | .0976 | -.221 |
|  | (.137) | (.133) | (.131) | (.146) |
|  |  |  |  |  |
| λnight (inverse Mills Ratio) | .175 | .585 | .586 | .534 |
|  | (.619) | (.614) | (.609) | (.605) |
|  |  |  |  |  |
| λafternoon (inverse Mills Ratio) | 1.793 | 1.706 | 1.445 | 2.291 |
|  | (1.766) | (1.621) | (1.658) | (1.616) |
|  |  |  |  |  |
| λevening (inverse Mills Ratio) | .0178 | -.0600 | -.199 | .146 |
|  | (.351) | (.380) | (.366) | (.428) |
|  |  |  |  |  |
| Negative Emotions (residuals) | -.0208 | -.0190 | -.0211 | -.0228 |
|  | (.0202) | (.0208) | (.0201) | (.0188) |
|  |  |  |  |  |
| Cognitive Processing (residuals) | -.0168 | -.0172 | -.0174 | -.0175 |
|  | (.0109) | (.0107) | (.0110) | (.0115) |
|  |  |  |  |  |
| Positive Emotions (residuals) | -.0302\*\* | -.0308\*\* | -.0311\*\* | -.0305\*\* |
|  | (.0103) | (.0106) | (.0108) | (.0111) |
| Intercept | -13.55\*\*\* | -13.82\*\*\* | -13.58\*\*\* | -15.48\*\*\* |
|  | (1.999) | (1.877) | (1.873) | (1.949) |
| Month, dayofweek, topic effects | Yes | Yes | Yes | Yes |
| N | 5706 | 5705 | 5705 | 5706 |

Note. Standard errors in parentheses; \* *p <* .1, \*\* *p <* .05, \*\*\* *p <* .01. Daypart definitions in (2)-(4) are consistent with Table 1.

WEB APPENDIX W10: Intuition behind Execution of Genetic Algorithm

Rooted in Charles Darwin’s evolutionary theory, genetic algorithms (GA) primarily involve three operations: reproduction, crossover, and mutation. Reproduction is a process in which certain characteristics from a generation (e.g., parent) are copied to the next (e.g., child) based on a specified objective function. Crossover refers to the crossing over of randomly paired off-springs. Specifically, children originating from the reproduction operation are randomly paired and then combined using a pre-defined criterion (e.g., random number generations, partial string exchanges) to create a new generation of solutions. Mutation is the process of randomly altering a solution to ensure the algorithm is not stuck in a local minima. For example, constantly performing reproduction and crossover among solutions that represent local minima will simply result in more generations of the same. Mutation can break this trend by altering the gene and creating a new species.

The three operations are sequentially executed in the following manner:

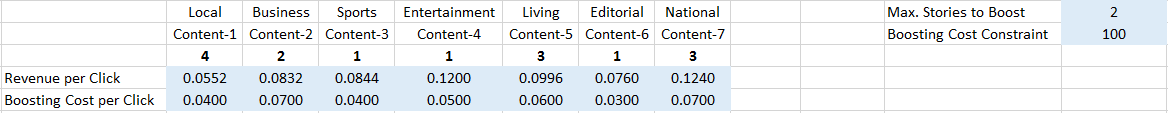
|  |  |
| --- | --- |
| Step 1: | GA starts with a population of N random solutions. |
| Step 2: | The “fittest” members of this initial population that maximize the objective function survive and move on to produce the next generation of solutions. |
| Step 3: | A new generation of solutions replace the population through the process of crossover and mutation. |
| Step 4: | The fittest members among the newest generation of solutions survive and move on to the next level. |
| Step 5: | Steps 2 through 4 are repeated until a stopping condition (e.g., convergence criterion) has been reached. |

To ensure near optimality, we repeat steps 1 through 5 for different mutation rates, initial population sizes, and stopping criterion. The key strength of GA over other nonlinear search techniques, such as coordinate ascent, is that it adopts a non-sequential search through random mutation[[1]](#footnote-1). Consequently, it not only covers more area within the problem space efficiently and quickly, but it also overcomes the issue of becoming stuck in a local minima. Consequently, it has a better chance of reaching the global optimal.

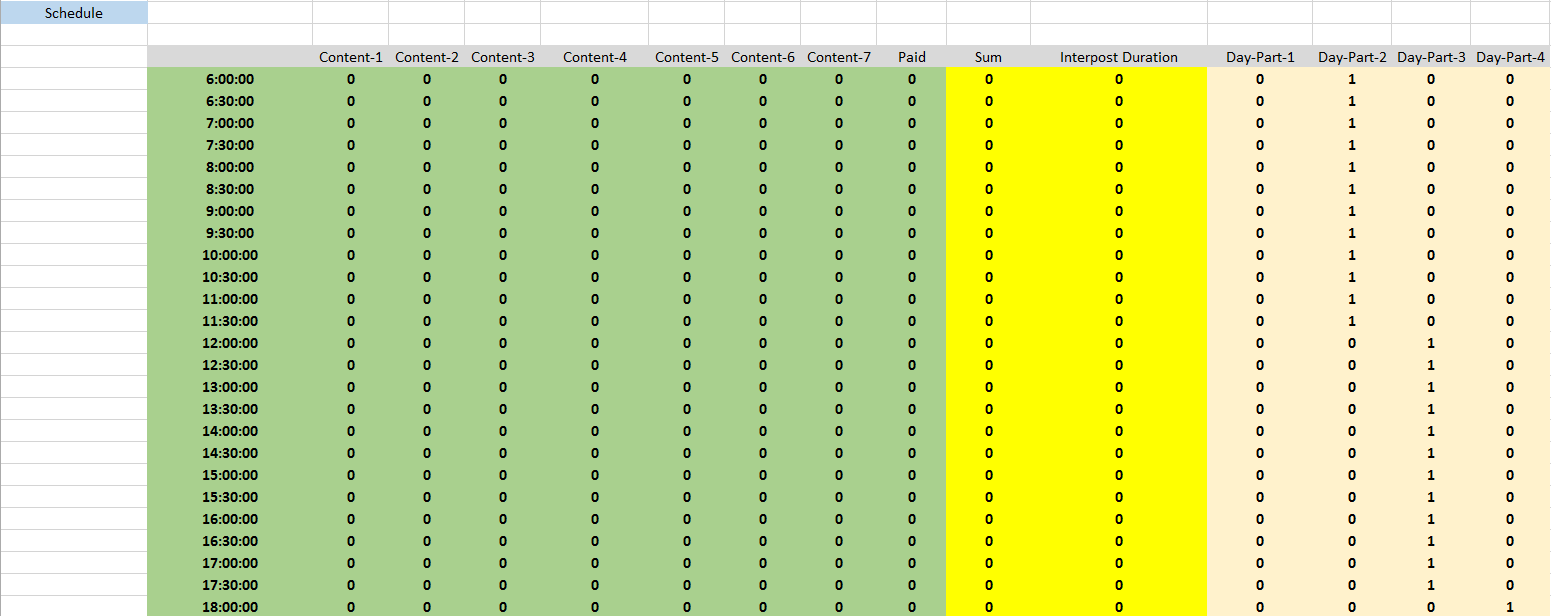
WEB APPENDIX W11: A Practitioner Guide to Scheduling Content on Social Media

*Step 1*: Calibrate the predictive model shown in Equation 13 and obtain the coefficients for all the parameters featured in the optimizer

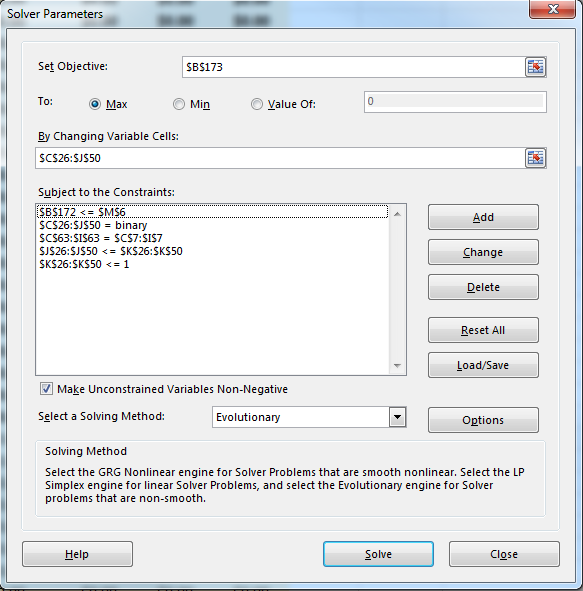
*Step 2*: Set the initial values for the optimizer (i.e., number of stories within each content category, number of stories to be boosted or the total boosting budget, revenues per click in each content category and average cost of boosting per link click within each content category)



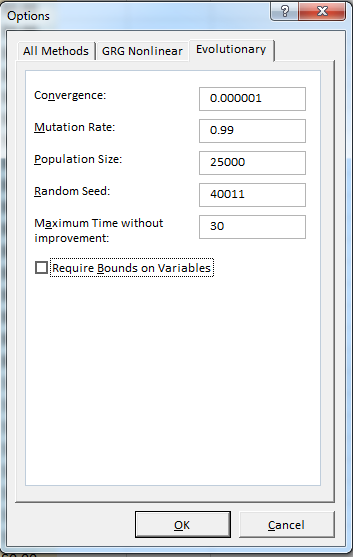
*Step 3*: Set initial values of the schedule to be 0



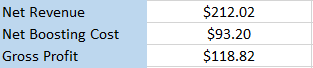
*Step 4*: Specify the objective function and constraints in Solver based on Equations 7-15

****

*Step 5*: Set the starting values for the Genetic Algorithm within Solver

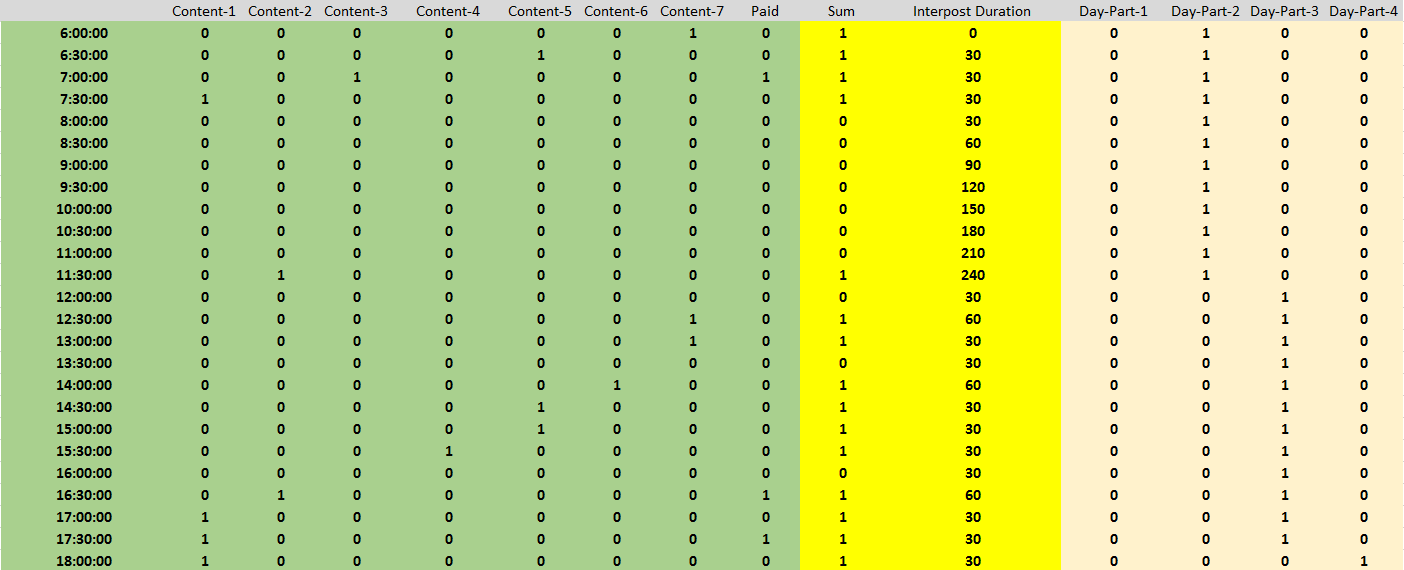
****

*Step 6*: Run the Solver and take note of optimal schedules and gross profit

****

*Step 7*: In order to minimize the likelihood of finding a local optima, repeat steps 5 and 6 with different starting values for the Genetic Algorithm until the incremental change in gross profit is negligible

*Step 8*: Implement the final schedule to maximize gross profits

****

WEB APPENDIX W12: Gross Profit at Various Levels of TCA Budget

(December 29, 2015 and December 30, 2015)

1. It is important to note that heuristic techniques do not always guarantee a global maxima and could predict an interior solution. This limitation does not undermine our algorithm because its goal is to simply derive a profit maximizing solution and not a global optima. In that sense, we claim our gross profit predictions to be conservative. [↑](#footnote-ref-1)