Daily Optimization Project

Personalized Facebook Analysis

10/29/15

By Kevin DeLuca

ThoughtBurner

Introduction

There are all sorts of theories out there regarding the best time to post to Facebook. Many of them are conflicting. Some recommend **posting in the morning**[i], others say **post at 5pm**[ii], and some suggest the **evening is best**[iii]. **Weekends are bad**[iv], or **maybe they're good**[v]. Frequency is often considered as well – some suggest you should only post about **once every two days**[vi], while others recommend closer to **two posts per day**[vii]. Pictures are supposed to give you **more likes**[viii] and **more comments**[ix]. So who's advice do we follow? Currently, much of this advice is aimed towards companies hoping to get more Facebook users engaged with their organization's page, and that fact is another reason to be skeptical that the recommendations would be applicable to your own personal posting.

Discerning the most effective way to personally use Facebook is not typically of concern to researchers who have the tools and skills capable of doing so. Many people, however, use Facebook daily, so effective Facebook-use strategies could be interesting to everyday people, particularly if the strategies are tailored specifically to the individual. Personalized analyses and strategies are not (as) interesting to researchers, so these types of studies have not been carried out – most analyses of Facebook posting are done with large sample sizes across a wide variety of people in hopes of finding generalizable results, usually aimed at helping the promotion of businesses and organizations[x]. These recommendations for Facebook posting are made from a business perspective with certain goals in mind. More average or recreational Facebook users have different reasons and objectives when posting, so strategies devised from a business perspective may not apply to individuals using Facebook for personal reasons.

Analyzing Facebook posts is a time consuming, data intensive procedure that most people will not undertake on their own. The expected benefits of having optimized Facebook strategies is most likely not worth doing all the work to discover them. The everyday use of Facebook combined with the high cost of research relative to expected everyday benefits of optimal Facebook posting makes analyzing Facebook posts a great project for ThoughtBurner's mission and purpose – to optimize performance on daily problems and absorb the research costs. As typical with ThoughtBurner research, the benefits of such research cannot be easily captured by a single entity because they will be spread out over many individuals and long periods of time.

Because many features of Facebook posts, such as likes and comments, are quantifiable we can use the standard statistical methods often employed in economics to gain insight about possible causal relationships. Then, depending on individually defined goals, we can use these suspected causal relationships to devise optimal Facebook usage strategies. For example, suppose we want to get the maximum number of likes on a new Facebook post. Using collected data, we could learn how including a link in a Facebook post affects its total number of likes. We could then decide whether to include a link or not. In this case, the optimal strategy would be the choice that leads to the highest number of expected likes on Facebook posts.

There are a large number of different possible individualized Facebook usage goals. The goal of this study is not to exhaustively list these optimal strategies given sets of personal preferences. Rather, it is meant to demonstrate how ThoughtBurner's Daily Optimization concept can be used with personalized data to create personalized optimization strategies.

This is the second part of a project that analyzes data collected from the researcher's Facebook friends. I now turn exclusively to my own Facebook posts. Using data collected on all of my posts since the introduction of the "like" button, I check for significant relationships between post characteristics and the number of likes each post receives. I then describe likely causes for the relationships observed. Because they are my own posts, I also reveal personal insights and describe features of my behavior that may have influenced the results. Using these results, I devise my personalized, optimal Facebook posting strategy. Last, I compare my individualized results to the results of the analysis of my collective friend group in the first part of this report. I find that the individual results differ from the group results significantly, demonstrating that optimal posting strategies for individuals (in this case, me personally) may be different than what would be considered an optimal Facebook posting strategy on average.

The results are not meant to be generalizable to all Facebook users. As mentioned before, the purpose of the paper is more to demonstrate an application of ThoughtBurner's Daily Optimization concept to a Daily Optimization problem – how to optimally post to Facebook. Users should be able to replicate the study using their own Facebook data and data from their friends if they want personal estimates of effects and individualized strategies for optimizing their own Facebook posts. Other factors suspected to influence aspects of post performance could also be added to analysis for greater insights.

Individual Analysis

In February 2009, Facebook created the "like" button[xi]. I recorded post characteristics on all of my own posts starting in February 2009 until I had the idea for this post in June 2015. Variables included date, day of the week, total likes, and dummies for pictures and links. In total, data on 450 posts was collected (after throwing out three outliers due to unusually high like/comment counts after rare life events).

My Facebook posting habits changed fairly dramatically over the period of observation. For example, in 2010 and 2011 combined, I only posted 26 times total, an average of slightly more than 1 post per month. Later, in 2013 and 2014 I posted 282 times – almost 12 times per month. The most dramatic single change between years occurred between 2011 and 2012, where my posting frequency increased by over 500%. The average number of likes on my posts also spiked up between 2011 and 2012, from about 3.8 likes per post in 2011 to about 8.5 likes per post in 2012, a 220% increase. Figure 1 shows the increase in the frequency of my posting, while Table 1 shows all summary statistics for my posts' characteristics by year.

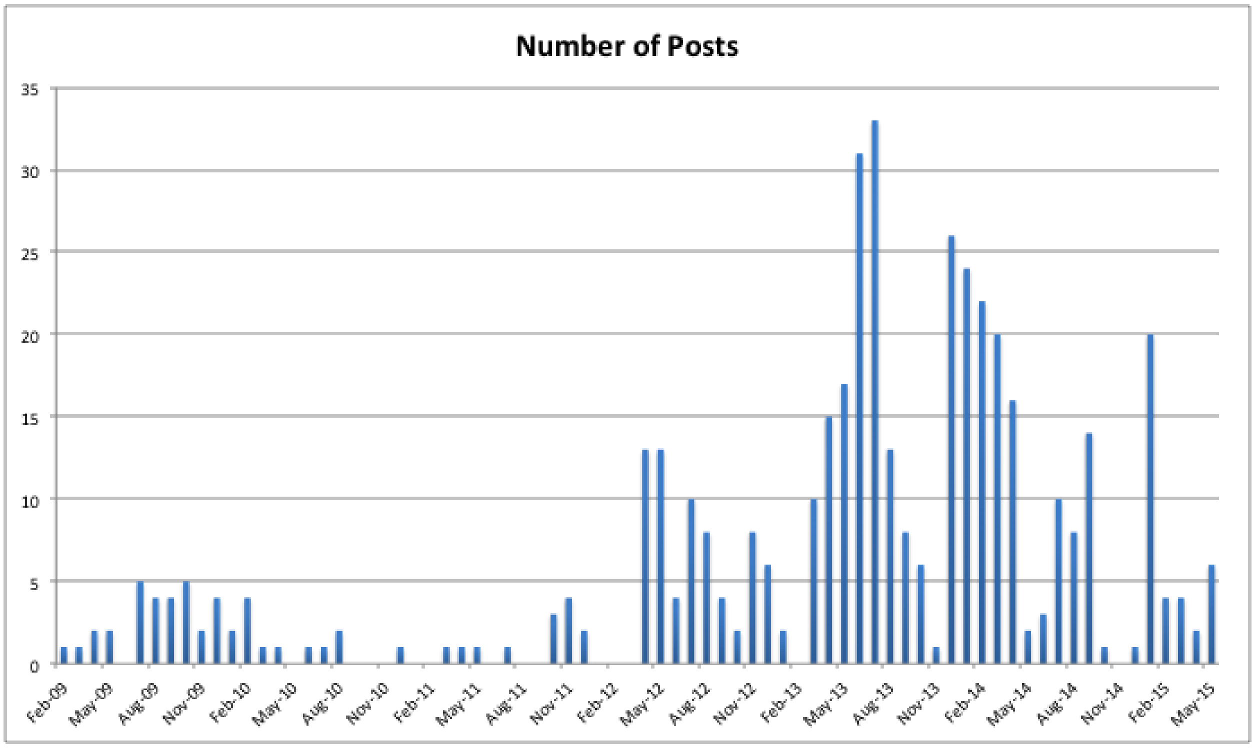

Figure 1: Number of posts over time

Because of what seems to be a characteristically different pattern of posting in the post-2011 years, and because the fraction of my total posts that occurred before 2012 is fairly small, I decided to only use posts that occurred from January 2012 until June 2015. Also, because the objective of the analysis is to determine how to optimize Facebook posting *now,* I suspect that using more recent posts in the analysis will provide better estimates of the current effects of post characteristics on likes and comments. Table 2 shows how the full data set differs from the post-2011 and onward data.

	Summary Stats - By Year						
	N	Total Likes	Total Comments	Average Likes	Average Comments	Percent with Picture	Percent with Link
All Years	450	4453	809	9.896 (9.409)	1.798 (2.557)	0.67	0.07
2015	44 9.8%	637 14.3%	75 9.3%	14.477 (11.629)	1.705 (2.086)	0.77	0.20
2014	121 26.9%	1450 32.6%	171 21.1%	11.983 (9.191)	1.413 (1.999)	0.96	0.01
2013	161 35.8%	1618 36.3%	280 34.6%	10.050 (9.189)	1.739 (2.448)	0.72	0.12
2012	68 15.1%	581 13.0%	137 16.9%	8.544 (8.391)	2.015 (3.107)	0.43	0.03
2011	13 2.9%	49 1.1%	68 8.4%	3.769 (5.630)	1.385 (1.981)	0.46	0.08
2010	13 2.9%	56 1.3%	151 18.6%	4.308 (7.005)	2.692 (3.838)	0.00	0.00
2009	30 6.7%	62 1.4%	93 11.5%	2.067 (3.183)	3.100 (3.497)	0.00	0.00

Table 1: Selected post characteristics over time

	Summary Stats - Comparison						
	N	Total Likes	Total Comments	Average Likes	Average Comments	Percent with Picture	Percent with Link
All Years	450	4453	809	9.896 (9.409)	1.798 (2.557)	0.669	0.071
2012-2015	385 85.6%	4123 92.6%	650 80.3%	10.709 (9.225)	1.688 (2.428)	0.745	0.081

Table 2: Full dataset relative to post-2011 dataset

A variety of post characteristics were recorded for all of my posts: date, number of likes, number of comments, a dummy for if a picture was in the post, and a dummy for if a link was in the post. Other controls were calculated for each friend after data collection, such as total number of friends, total number of likes, total number of posts, average number of likes, average number of comments, and days since the last post. For a more detailed and interactive look at my Facebook posting data, see **my visualization of likes** and **my visualization of friends** on my Tableau Public profile[xii].

I performed an analysis on my own individual posts in a similar fashion to the analysis done in **part one**[xiii] (with data from my Facebook friends' posts). The end purpose of this new analysis was to devise an optimal Facebook posting strategy – namely, a strategy to maximize average likes. Table 3 shows the results of a regression of *likes* on selected post characteristics. There were no significant effects of any post characteristics on the number of comments my posts received, so I do not report these results. I slowly add in variables as an informal test of robustness. Most of the effects are similar across specifications. A more detailed discussion of the variables and the results follows.

Table 3: Personal Data Like Effects

	Number of Likes a Post Receives						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Year	**1.725**	-1.660	-1.532	-1.599	-1.537	-1.679	-1.411
	(0.536)	(1.488)	(1.447)	(1.468)	(1.465)	(1.468)	(1.454)
Total Friends		**0.024**	**0.033**	**0.033**	**0.032**	**0.033**	**0.031**
		(0.009)	(0.009)	(0.009)	(0.009)	(0.009)	(0.009)
Link			**-9.495**	**-9.525**	**-9.488**	**-9.736**	**-9.446**
			(1.971)	(1.973)	(1.998)	(1.998)	(2.011)
Picture			**-7.021**	**-6.896**	**-6.718**	**-6.374**	**-6.260**
			(1.573)	(1.572)	(1.651)	(1.687)	(1.707)
Days Since Last Post					0.061	**0.323**	0.551*
					(0.081)	(0.147)	(0.325)
Days Squared							-0.027
							(0.021)
Days Cubed							0.00029
							(0.00028)
Constant	**8.433**	-0.945	-0.230	-1.352	-1.193	-2.236	-1.818
	(0.816)	(3.538)	(3.464)	(3.829)	(3.862)	(3.813)	(3.827)
Day of the Week Controls				X	X	X	X
Only Posts Within a Month						X	
R-squared	0.026	0.044	0.119	0.126	0.128	0.146	0.135
Observations	385	385	385	385	384	380	384

Results in **bold** are significant at the 5% level. No day of the week controls had significant effects, so their coefficients are not included. *Significant at the 10% level.

Discussion of Results: Individual Posts

First, I wanted to see if I had been getting better at posting to Facebook over time (in terms of average number of likes). Plotting the average number of likes over time, it seems pretty clear that my posts were getting better over time (see Figure 2). A simple regression of likes on year showed a positive relationship between the average number of likes my post received and year – as time went on, my posts were scoring more likes on average (specification 1).

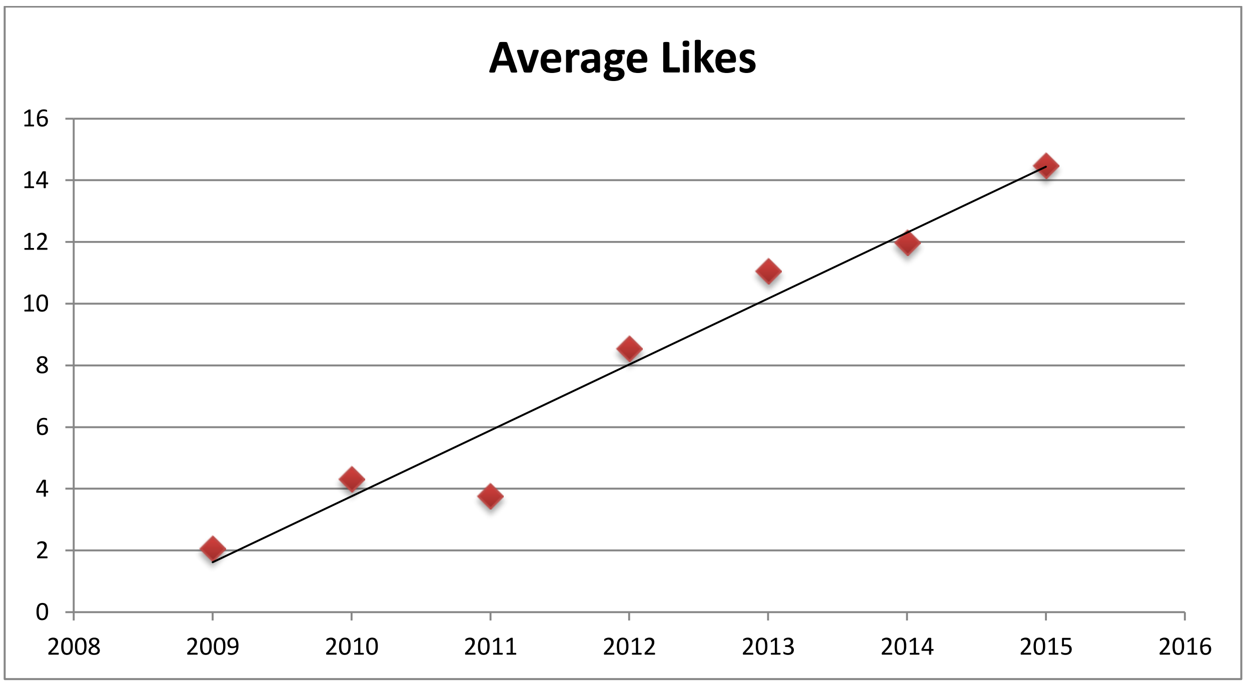

Figure 2: Average number of likes over time

However, I thought that this effect might be driven by more and more of my friends using Facebook over time. Adding in a control for the total number of friends I had at the end of each month (2), the positive significant coefficient for year goes away and instead the coefficient for total friends becomes positive and significant. Unfortunately for me, it seems that the increased number of likes I've been getting on my posts has been mostly driven by the fact that I have more friends on Facebook, rather than because of an increase in the quality of my posts. The total friends coefficient estimate is relatively stable across all additional specifications.

Posts of mine that had a link had about 9.5 less likes than posts of mine without links. This effect was highly significant. Posts with pictures also had significantly less likes than posts without pictures – between 6 to 7 less likes. While the effect of pictures may seem surprising to those who subscribe to the idea that pictures should increase the number of likes, it did not surprise me personally – this is because I know about my past posting habits.

For a while, I was on a "one-a-day" posting habit when it came to instagram pictures. I thought it would be cool to look back at my instagram photo album on Facebook and see a picture of something that happened each day. While I wasn't totally consistent in following this habit, Figure 3 shows that starting around 2012 you can see a large increase in the fraction of my posts with pictures in them. Darker green means a higher percentage of my posts were pictures, while the height of the bar shows the average number of likes (per month).

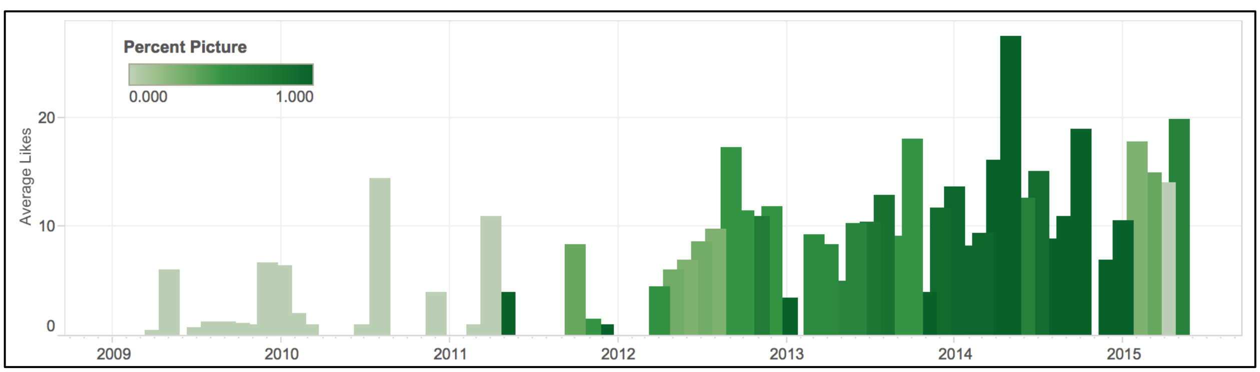

Figure 3: Average number of likes over time. Darker green indicates a higher fraction of picture posts.

My theory is this: when I was 'forced' to post a picture everyday, the average quality of my posts decreased. (The increasing average likes over time is more of an effect of my total number of friends, as mentioned before.) The quality of my picture posts decreased because I now had to post a picture every day, even if there was nothing that I normally would have deemed 'post-worthy.' Meanwhile, my non-picture posts remained high(er) quality, because I was not ‘forced’ to post non-picture posts when I felt they were not ‘post-worthy’. If this were true, we would see exactly what the data shows – a higher percentage of picture posts but a negative effect of picture on number of likes.

Next, I wanted to see if the day of the week affected the number of likes my posts receive. None of the results were significant. However, I include them in the later versions of the analysis as controls. Their inclusion or exclusion does not significantly alter any of the results. I also did not collect time stamp data – while I was recording characteristics of my posts, I did not realize I could record the exact time my posts were posted. For now, I cannot say if there are any personal time-of-day effects for me or, if they do exist, whether they are the same as the time-of-day effects from the group analysis.

In the group analysis, I was limited in my ability to test for a time-between-posting effect because I only collected data on posts within a two-week observation period. Using my own data, however, I have information about posts that were posted more than two weeks apart and can use this variation to learn more about the effects of being ‘patient’ (waiting longer between posts).

Simply including a variable for ‘days since last post’ suggests that there are no significant effects on the number of likes for being patient (5). While I was collecting data on my posts, however, I noticed that there were a few extended periods of time where I did not post at all – sometimes for months. If it were the case that the effect of waiting an additional day ‘maxed out’ after a certain amount of time, then these long periods of no-posting could be messing up the results (they are outliers, in a sense). For example, if you have already waited a month since your last post, waiting an additional day might not have much of an effect. On the other hand, if you posted earlier that day then maybe waiting a day before your next post will stop you from ‘being annoying’ to other people (possibly resulting in slightly more likes).

The first way I tested this idea was to limit my sample to posts that were posted within 1 month (4 weeks) of the last posts. This means that the first posts I posted after the long period of no-posting (and their very high 'days-since-last-post' value) were excluded. The results of this regression (6) show that there is a significant and positive relationship between the number of likes one of my posts receives and the number of days since my last post. The effect is about +0.323, which would imply that for every 3 days I waited between posts I got about 1 additional like.

I decided to add in days-since-last-post squared and days-since-last-post cubed terms to my regression to allow for a non-linear effect of being patient. This allowed me to model the idea that the effect might change depending on how long a person had already waited.

The results from this last specification (7) show that 'days-since-last-post' is no longer significant at the 5% level. However, the results are significant at the 10% level, and the coefficients on days, days squared, and days cubed are all in the direction we would expect. While not definitive, there is slight evidence that waiting between posts increased the number of likes my posts received, but the magnitude of this effect does not increase indefinitely. Using the coefficients provided in the estimation, the optimal time to wait between posts is 13 days, and the effect would be 3.24 additional likes. Figure 3 shows a visual representation of the estimation provided by specification (7).

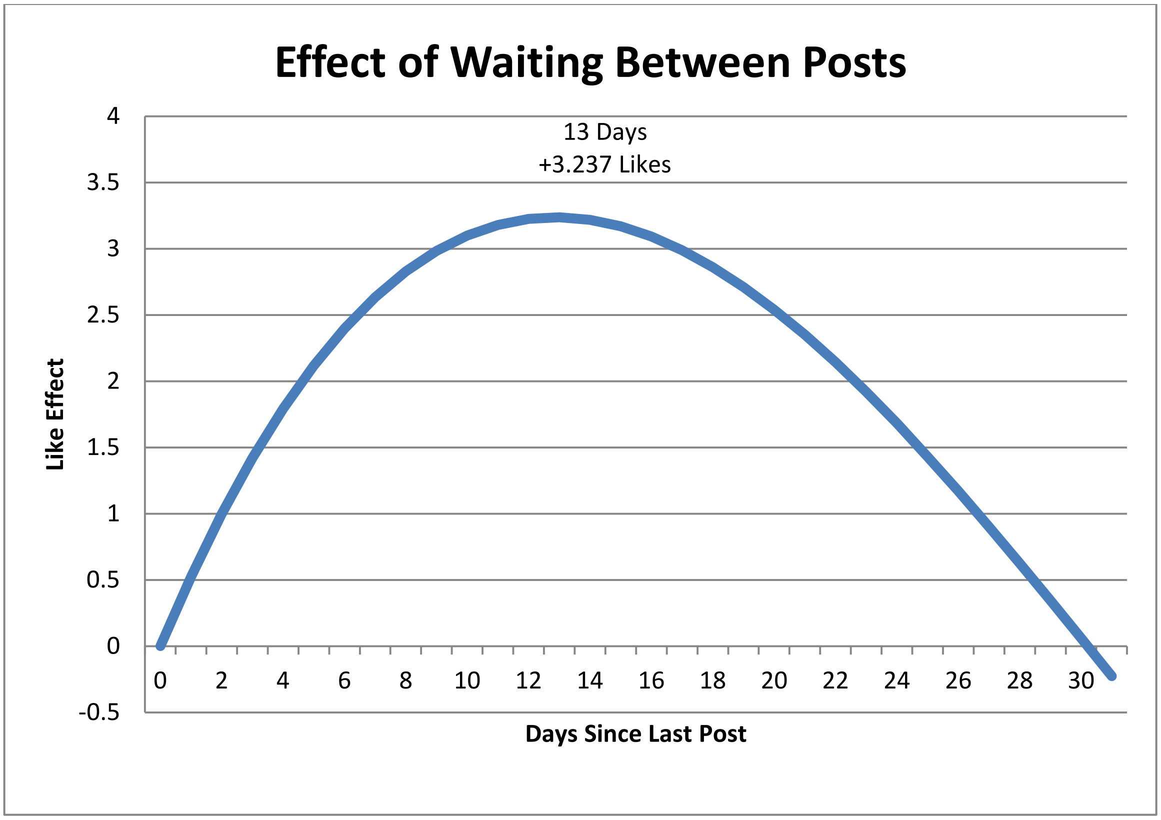

Figure 3: Effects of waiting between posts, based on specification (7)

Overall, the results show that there are a number of interesting factors that have significant effects on the number of likes my posts receive. For every 33 more friends I have, my posts got about 1 additional like on average. Posts with links in them had about 9.5 fewer likes than posts without links, and posts with pictures in them had about 6.25 fewer likes than posts without pictures. There was also some evidence that waiting between posts increases the number of likes my next post received, and my best estimation is that these effects ‘max out’ at an additional 3.24 likes for waiting 13 days. There were no significant day-of-the-week effects and there were no effects on comments that were significant at the 5% level.

Comparison to Friend Group Analysis

The results from my personalized Facebook analysis are different from the results of the group analysis. The differences, however, were only whether factors significantly affected likes or the magnitude of the effects; never were effects both significant and in the opposite direction. In terms of comments, none of the (few) significant comment effects found in the group analysis were found in the personal analysis.

The total friends effect was significant and positive in both the friend group analysis and my individualized analysis, though the effects were bigger for the individual analysis. In the group analysis, the size of the effect suggested that for every 91 more friends somebody had, they got about 1 more like on their posts on average. In the individual analysis, the effect sizes were bigger: I received about 1 more like on average for every 31-42 (depending on specification) more friends I had. This could be because the specific group of people who became my Facebook friends over this period were more likely to like my posts compared to the average friend of one of my other Facebook friends.

Posts with links had 9.5 fewer likes than posts without links in my personalized analysis, which is a noticeably smaller effect size compared to the effect in the friend group analysis (which was close to 16 fewer likes). Posts with pictures, on the other hand, had around 6 to 7 fewer likes than posts without pictures in my personalized analysis, whereas in the friend group analysis there was no significant effect of including pictures in Facebook posts.

There was no evidence of day of the week effects in the personalized analysis, while there was slight evidence for significant effects in the friend group analysis (see the first report, Table 3 specification (5)). In the friend group analysis, there was also evidence that Monday posts received significantly more comments (about 1 more on average), but there was no evidence of any significant comment effects in the personalized analysis.

Last, the friend group analysis revealed that posting more than once in a two-week period resulted in significantly less likes (about 0.5 fewer likes per additional post). Using the greater variation in days between posts in the personalized analysis, I found suggestive evidence that waiting a short time between posts increases likes, but that waiting too long might negate these effects. The best guess for the average-like-maximizing time for me to wait between posts is 13 days.

The fact that there are differences in the effects and their sizes suggests that the optimal Facebook posting strategy on average may be different than a personalized Facebook posting strategy. The last section explores the implications of these differences and, more specifically, how they affect the solution to the Facebook posting Daily Optimization Problem.

Personalized Posting Optimization

Given that the results for my personalized analysis of Facebook posts were different than the friend group analysis, it makes sense to try to re-optimize. What works as a strategy 'on average' is not very useful to the individual for whom the strategy doesn't work. In order to truly optimize your Facebook posting, you need to know specifically what works for you.

Using the results from my personalized analysis, the simple advice I can give (to myself) would be to not include links or pictures in my posts and to wait almost two weeks (13 days) between each post. Comparing this 'best' strategy to the 'worst' strategy, my expected average likes would be *12.68 likes higher*. If I assume that the time-of-day effects are the same for me as they were for my group of friends, then I could also make sure to post in the evening and this new expected average would be *19.02 likes higher*. This is about 4 likes less than the maximum effect predicted by optimizing based on the friend group analysis. Qualitatively, however, the optimal strategies are very similar. But since there are minor differences, the most obvious difference being the interpretations/effects of the time-since-last-post factor, we should take a closer look at the implications of each result in terms of determining an optimal Facebook posting strategy.

For example, both the friend group optimized strategy and my personalized optimization strategy prescribe specifically not to include links in posts if the objective is to maximize average likes. A small difference, however, is that the personalized results also advise against posting pictures, while the group analysis would say it was fine (no significant negative effects). Unfortunately I can't compare time-of-day effects at this time so we don't know if this advice would differ based on the separate analyses. There were no day-of-the-week effects in either, so in that sense the advice between the two analyses would be the same.

Another difference is in the frequency or time-between-posts effects. In the friend analysis, the most I could say is that within a two-week period, people who posted more often received fewer likes on average (about 0.5 fewer likes per additional post). Using the personalized analysis, we were able to estimate the specific effects of posting one day after, posting two days after, etc. (see Figure 3). Our best guess on this, then, is that the like-maximizing time is to wait 13 days between posts. This is in contradiction with the friend group analysis advice, which cautions against posting more than once within a two-week period.

To give you a better idea of how the optimal strategies differ, and how they would affect expected likes on Facebook posts, Table 4(a) and 4(b) shows the differences in expected likes based on different assumptions. The columns indicate which dataset is being optimized, whereas the rows indicate which strategy is being used. The cells display the expected difference between the like-maximizing strategy and a strategy-constrained like-minimizing strategy – that is, they show the difference between the maximum like effect possible and the minimum like effect possible on the column-determined data while following the row-determined strategy. For example, the friend group optimizing strategy suggests that pictures have no effect on likes – but including a picture while optimizing my personal data would reduce the expected average number of likes by about 7. The cells where the data and strategy are mismatched (white cells) show 'how wrong' things can go by following the wrong advice (which gives you an idea of the range of the total effect on likes). Table 4(a) assumes that there are no time-of-day effects in the personal data, while Table 4(b) assumes that the time-of-day effects are identical.

Table 4(a)

	Friend Group Data	Personal Data
Friend Group Optimization Strategy	*+23.61*	*+6.35*
Personal Optimization Strategy	*+23.06*	*+12.68*

Table 4(b)

	Friend Group Data	Personal Data
Friend Group Optimization Strategy	*+23.61*	*+12.69*
Personal Optimization Strategy	*+23.06*	*+19.02*

Tables 4 (a) and (b) reveal something interesting. While an average person in my friend group would have been fine with my personalized optimization strategy, I may not have been fine using the strategy based on group data. Assuming that the friend group is representative of your posting, the difference between using the friend group strategy and the personalized strategy is only 0.55 likes on average – not very large. This small difference is a result of posting a little too often; you would post every 13 days, which is more than once every two weeks, and your expected likes would decrease by 0.55 likes by doing this. You'd be better off using the friend group strategy, but only by a little.

On the other hand, if I had personally followed the friend group strategy I could have ended up with an expected like bonus of only 6.35 more likes. Compare this to my personalized strategy, which had a max of 12.68 more likes. The difference between the two strategies is 6.33 likes, meaning that I could be pretty seriously sub-optimizing. The discrepancy comes from two factors; the fact that in the friend group strategy, pictures were found to have no significant effects and the difference between the optimal wait time between posts.

If I were to post pictures – something totally fine according to the friend group results – then I would end up suffering the personal negative consequences of including pictures in my posts (6.26 fewer likes on average). And if I were to post on average once every two-weeks – again, the frequency prescribed by the friend group results – I would be slightly off from the personal optimal timing of posting 13 days. Specifically, the effect of waiting to post would be +3.17 rather than +3.23, so I would miss out on an extra 0.06 likes. Cumulatively, it would result in 6.33 fewer likes than if I had followed my personalized optimal strategy.

Conclusion

The purpose of this project was to find an optimal Facebook posting strategy. Using two different datasets, possible effects of commonly suspected post characteristics were investigated. A different, best-guess average-like-maximizing Facebook posting strategy was created using the significant relationships found in the analyses, though the results could also be used with other objectives in mind. The strategies were not identical, suggesting that in order to truly solve the Facebook posting Daily Optimization Problem, an approach using aggregated data on Facebook posting may not be appropriate.

Of course, the main limitation of both of these analyses is the restricted sample. There is no guarantee that any of the results or strategies found are generalizable outside of the friend group analyzed (plus myself). The weakness of a restricted sample, however, is also a strength – the point of the project is more to demonstrate how ThoughtBurner's Daily Optimization concept could be used to develop personalized optimization strategies with data available to the everyday person. This cannot be accomplished without using a restricted sample (namely, restricting the sample to the data directly related to your optimization).

Another limitation is that some of my older posts were liked by friends I made *years after* I originally posted them (you know who you are), meaning that 'long term' effects of factors may not be accounted for in the more recent data. Basically, your future friends might affect your current post 'performance' as measured over a longer period of time. For example, there may some of my Facebook posts that will be liked in a year from now by people who I am not friends with yet, and these likes were not taken into account in the current analysis. These effects are probably small though.

Despite limitations in interpreting the data, the results also lend weight to certain theories about optimal Facebook posting. The results will be less interesting to those not in the sample, but for individuals who comprise the analyzed group the effects and interpretations are more likely to be relevant and true.

It also demonstrates the fact that what works 'on average' may not help individual users solve their Daily Optimization problems. While there were not major differences in the friend group optimization strategy and the personal data optimization strategy, the possible effects of sub-optimizing (i.e. using the wrong strategy given some specific data) were not insubstantial. In order to truly solve Daily Optimization problems, it is essential to analyze, understand, and somehow take into account these differences.

---

[i] http://blogs.constantcontact.com/best-time-post-facebook/
[ii] https://blog.optimizely.com/2015/07/08/how-to-find-the-best-time-to-post-on-facebook/
[iii] https://www.quicksprout.com/2015/01/02/what-are-the-best-times-to-post-on-social-media/
[iv] http://versus.com/en/2014/09/23/want-more-likes-the-best-time-of-the-day-to-post-on-social-media
[v] http://www.exacttarget.com/blog/why-you-should-be-posting-to-facebook-on-the-weekends/
[vi] https://blog.kissmetrics.com/science-of-social-timing-1/
[vii] https://blog.bufferapp.com/social-media-frequency-guide
[viii] https://blog.kissmetrics.com/more-likes-on-facebook/
[ix] https://blog.bufferapp.com/7-facebook-stats-you-should-know-for-a-more-engaging-page

[x] https://blog.bufferapp.com/social-media-stats-studies
[xi] https://www.facebook.com/notes/facebook/i-like-this/53024537130
[xii] https://public.tableau.com/profile/kevin.deluca#!/
[xiii] http://thoughtburner.org/2015/10/15/thought-optimizing-facebook-posts/