Online Appendix to: Down-to-the-Minute Effects of Super Bowl Advertising on Online Search Behavior

RANDALL A. LEWIS, Google, Inc.
DAVID H. REILEY, Google, Inc.

APPENDIX 1: CALCULATIONS

The key to our success in detecting large search lifts due to Super Bowl advertising is the concentration of ad spending against a large audience reached at a single point in time combined with temporally granular search data. We consider the statistical problem. Let C=\$3 million be the cost of the commercial; assume that the expected ad effect is proportional to the cost: ad impact = $\alpha \cdot C$ for some α. This is reasonable for an advertiser's marginal spending for which their return on investment (ROI) should be close to the cost of capital. Let σ be the baseline standard error of an estimate of a single commercial's impact. We observe that the t-statistic has both components:

$$t = \alpha \cdot \frac{C}{\sigma}$$

where $\alpha \cdot C$ and σ are in units of the outcome of interest.

Now, observe the t-statistic when we split the budget C into N commercials. The signal remains the same—we still spend the same amount of money and expect the same ROI—but the baseline variance is now scaled up by the number of commercials, or equivalently, the standard deviation is scaled up by the square-root of N:

$$t = \alpha \cdot \frac{C}{\sqrt{N} \cdot \sigma}.$$

Alternatively, if we consider the t-statistic for each commercial, we divide the expected signal of each commercial by N:

$$t = \alpha \cdot \frac{C/N}{\sigma}.$$

Intuitively, each commercial is an observation—but here we have made each observation $1/N$ less informative. As a result, to achieve the same t-statistic as before, we will need N^2 of the less informative observations.

For example, if we estimate a t-statistic of 15 for a given commercial's search lift, we should expect to find a t-statistic of $15/N$ if we split a commercial's budget into N less expensive nationwide ads. Consider a 1/20 ad buy of \$150,000. We would expect a t-statistic of roughly 15/20 = 0.75 from a single commercial. In order to be confident in detecting statistical significance, we need an expected t-statistic of 3. This would require running 4^2=16 commercials at a cost of \$150,000·16 = \$2.4 million. Even by spending the Super Bowl's ad budget of \$3 million, we only achieve an expected t-statistic of $\sqrt{20} \cdot 0.75 = 3.5$ rather than 15 under such dilution. We would need to spend 20 times the Super Bowl budget, or \$60 million, to achieve the same level of statistical certainty about the effects of that spending.

Before concluding, we consider one additional setting. Suppose our budget was instead split among mutually exclusive geographic or audience segments. Let S be the number of segments. Now consider the effect of a single commercial:

$$t = \alpha \cdot \frac{C/S}{\sigma/\sqrt{S}}.$$

Author's addresses: R. Lewis, randall@econinformatics.com, and D. Reiley, david@davidreiley.com, Google, Inc., 1600 Amphitheatre Parkway, Mountain View, CA 94043.

DOI:http://dx.doi.org/10.1145/0000000.0000000

Here we have divided both the signal and the noise among the segments. If the commercials and our outcomes can be accurately divided, we actually do not suffer as much as above where the noise was unaffected by splitting up the budget among N commercials. The simplified expression for segmentation follows:

$$t = \alpha \cdot \frac{C}{\sqrt{S} \cdot \sigma}.$$

This expression for a single segmented ad is identical to the expression for spending the whole budget C on N nationwide ads. Therefore, we can achieve the same level of statistical precision afforded by a Super Bowl ad by segmenting. This is quite intuitive: if we ran a Super Bowl ad and observed segment-level and nationwide data, we should expect the same level of statistical significance from both outcomes.

Super Bowl commercials are perfect examples of market-level concentrated exposure in TV—economically significant ad expenditure that produces a detectable effect against a large share of individuals for whom we can observe behaviors [Lewis 2012]. We can use the equations above in conjunction with t-statistics from the Super Bowl to extrapolate the statistical power of measuring brand-related search lift for other TV commercials. We already answered the question, "How many $150,000 commercials does it take to achieve a Super-Bowl-sized t-statistic?" Now we ask, "What is the smallest commercial for which we should expect a statistically significant search lift?" We again consider nationwide and segmented commercials.

For nationwide commercials, we still face the same baseline variability in searches. We again use t=3 as our requirement for reliably expecting a significant search lift and t=15 as our expectation for the Super Bowl commercial's statistical significance. We compute the relative costs using the ratio of t-statistics where their numerators are proportional to cost and the denominators are the same for the single nationwide commercial (denoted with the subscript 1NC) and the Super Bowl commercial (denoted with the subscript SB):

$$\frac{t_{1NC}}{t_{SB}} = \frac{C_{1NC}}{C}$$

$$C_{1NC} = \$3 \text{ million} \cdot \frac{3}{15} = \$600{,}000.$$

Thus, a $600,000 nationwide commercial is the least expensive commercial for which we can reliably detect a search lift for the typical Super Bowl advertiser.

Segmentation is defined as the ability to filter the search queries to the particular TV audience. For segmented commercials, the baseline variability in searches is reduced by the square-root of the segmentation factor, S. This is a result of the impact of the TV commercial's impact being focused on a smaller audience, which naturally generates a smaller cumulative variance. Again, we take the ratio of t-statistics (denoting the segmented commercial with 1SC):

$$\frac{t_{1SC}}{t_{SB}} = \frac{C_{1SC}/\sqrt{S}}{C}.$$

We know that $C_{1SC} = C\,/\,S$ which simplifies the expression to

$$\frac{3}{15} = \frac{(C/S)/\sqrt{S}}{C} \Longrightarrow S = 25.$$

Therefore, for a segmented buy, we should be able to detect a TV commercial that has a Super Bowl level of per-person spending (2-3 cents) as long as it costs at least $3 million / 25 = $120,000. However, few other commercials achieve a Super-Bowl-sized local or segment reach of 1/3 for a single commercial. Adjusting for reach, r, introduces a variance scaling factor in the denominator:

$$\sigma_S = \sigma/\sqrt{S \cdot r/⅓}$$

which adjusts the minimum-detectable cost by a factor of 1/3 / r: \$120,000 / 3r. This is still a large spend for a single commercial, but this minimum-detectable cost can be further reduced if we can identify who is or is not watching the show that the commercial airs on. The improvement gained by focusing on a narrower segment versus measuring the outcome at the national level can be extended by performing the analysis only on those individuals who were likely to have seen the ad. Given that a typical commercial reaches less than 10% of the population, other ways to exclude the 90% of non-viewers from the analysis can theoretically reduce the minimum-detectable cost by a factor of at least 3.

Understanding the practical structure and theoretical limitations of detecting the impact of TV advertising on search behavior not only illustrates the difficulty of the problem, as Super Bowl ads are atypical, but also outlines the feasibility set and opportunities to improve the signal. Online search behavior as a signal of TV commercial effectiveness can be further enhanced by advertisers and technologists.

Advertisers can design commercials to stimulate a viewer's motivation to search online. Orders of magnitude of difference in detectability are observed across products: for example, compare Doritos' much larger search lift with Pepsi's. This could easily be done by highlighting the Pepsi's broadly appealing web presence. This way of strengthening the search signal could be especially useful for advertisers who want to measure attentiveness to the TV commercial across placements.

Technologists can construct better aggregations of commercial-related queries by efficiently extracting only the data for impacted queries. This can both increase the signal by capturing affected queries that were missed in this research and decrease the noise by omitting unaffected queries that were erroneously captured. Along these same lines, only including very significantly affected queries and ignoring less significantly affected queries can also improve detectability as not every signal is worth the noise it carries along when included.

Finally, while the number of incremental searches impacts the detectability of the search lifts, the baseline search variability is the other half of the equation. Large, well-known Super Bowl advertisers may tend to have greater baseline search variability—perhaps in proportion to their size. Thus, in line with Lewis and Rao [2013], there are both affordability and detectability limitations on detecting the effects of TV commercials on search behavior. Smaller advertisers who can afford less expensive commercials may be able to detect meaningful effects from those whereas the large advertisers cannot, due to their more volatile search baseline. Thus, we note that the bounds computed here are for Super-Bowl-scale advertisers, not for all advertisers. Lewis and Rao [2013] show that the detectability of ad effects increases in the absolute cost of advertising media but decreases in the percentage of firm revenues. Future research can investigate how detectability changes with firm size.

APPENDIX 2: DEFINITION OF *RELATED QUERIES*

A search page view is defined as related if either the query or any search or ad link's URL matches one or more of the following regular expressions:

.*\.audiusa\.com.*
.*\.audi\.com.*
.*\.bestbuy\.com.*
.*\.bmwusa\.com.*
.*\.bmw\.com.*
.*\.bridgestone\.com.*
.*\.bridgestonetire\.com.*
.*\.captainamerica\.marvel\.com.*
.*\.carmax\.com.*
.*\.cars\.com.*
.*\.thecoca-colacompany\.com.*
.*\.coca-cola\.com.*
.*\.cowboysandaliensmovie\.com.*
.*\.doritos\.com.*
.*\.fritolay\.com.*
.*\.doritoslatenight\.com.*
.*\.doritoschangethegame\.com.*
.*\.crashthesuperbowl\.com.*
.*\.etrade\.com.*
.*\.godaddy\.com.*
.*\.groupon\.com.*
.*\.homeaway\.com.*
.*\.hyundaiusa\.com.*
.*\.kia\.com.*
.*\.mbusa\.com.*
.*\.mercedes-benz\.com.*
.*\.motorola\.com.*
.*\.pepsi\.com.*
.*\.salesforce\.com.*
.*\.skechers\.com.*
.*\.mars\.com.*
.*\.teleflora\.com.*
\.thor\.marvel\.com.
.*\.volkswagen\.com.*
.*\.vw\.com.*
.*\.transformersmovie\.com.*
.*\.rangomovie\.com.*
.*\.rio-themovie\.com.*
.*\.super8-movie\.com.*
.*captain america.*
.*cowboys and aliens.*
.*cowboys & aliens.*
.*cowboys-and-aliens.*
.*limitless.*
.*pirates of the.*
.*rango movie.*
.*rio movie.*
.*rio the movie.*
.*super 8 movie.*
.*super8 movie.*
.*thor movie.*
.*thor 2011.*
.*packers.*
.*steelers.*
.*christina aguilera.*
.*black eyed peas.*
.*usher.*

APPENDIX 3: EXAMPLES OF *RELATED QUERIES*

Below we find some examples of related queries on January 30, 2011 for Captain America listed in decreasing frequency of appearance:

captain america
captain america trailer 2011
captain america movie
captain america trailer
captain america movie trailer
captain america the first avenger
…
when will we see captain america appear in thor
new captain america movie
captain america super bowl
who will play captain america
the first avenger captain america 2011 trailers
captain america in thor
…
iron man finds captain america
wii games captain america
captain america teaser trailer
captain america cycling jersey
is there a female eivalant to captain america
captain america poster 2011
captain america cmoic
…
captain america kids halloween costume
captain america of vietanam
green lantern "captain america"
who is playing captain america
play captain america games
captain america arrest florida