Real-time Television ROI Tracking using Mirrored Experimental Designs

Brendan Kitts[1], Dyng Au[1], Brian Burdick[1]

[1]PrecisionDemand, 821 Second Avenue, Suite 700, Seattle, WA. 98104.
{bkitts,dau,bburdick}@precisiondemand.com

Abstract. Real-time conversion tracking is the holy grail of TV advertisers. We show how to use thousands of tiny areas available via commercial cable and satellite systems to create low cost tracking cells. These areas are created as "mirrors" of a national campaign, and run in parallel with it. With properly controlled areas, it is possible to calculate national effects due to TV using statistical methods. We show performance of the method on a large-scale TV advertising campaign where it was used successfully to maintain a real-time CPA target of $60 for 179 days[1].

Keywords: Television; ROI; Conversion Tracking.

1 Introduction

Tracking ROI from Television is an unsolved problem for advertising. There are no physical mechanisms that allow for tracking a viewer from the view event to a purchase in a store, dealership, or over the web. This has led many marketers to be unable to allocate rational budgets towards TV advertising. This paper describes a method for using TV cable and satellite systems to track conversions due to TV. This method does not use panels and can be implemented using today's TV infrastructure.

2 Prior Work

There have been many attempts to track the revenue being generated from TV advertising.

---

[1] Kitts, B., Au, D., Burdick, B. (2013), "Real-time Television ROI Tracking using Mirrored Experimental Designs", Data Mining Applications in Industry and Government Workshop, Proceedings of the Seventeenth Pacific-Asia Conference on Knowledge Discovery and Data Mining, Springer.

2.1 IPTV

Many commentators have written that efforts such as IPTV will eventually enable TV conversions to be tracked via conversion tracking pixels similar to those in place today throughout the web. IPTVs obtain their TV content from the internet and use HTTP protocol for requesting content. In addition to TVs, Mobile phones such as the iPhone may be able to control the TV, and possibly click on on-screen ad content. However there are many technical challenges before this becomes a reality. Today only 8% of US TV households have IP enabled TV. Even if IPTV or web-based TV control becomes widespread, it still won't capture all of the activity such as delayed conversions, and purchasing at retail stores. In fact [18] have noted that TV disproportionately reaches low brand users, and so immediate ad response conversions may undercount TV's effectiveness.

2.2 RFI Systems

Some companies have experimented with methods for enabling existing TVs to be able to support a direct "purchase" or "Request For Information" (RFI) actions from "the lounge" using present-day Set Top Box systems and hand-held TV remote controls. The QUBE system, piloted in the 1970s, was an early example. In 2010 Backchannel Media developed an on-screen "bug" that appeared at the bottom of the screen and asked the consumer if they would want more information or a coupon. The consumer could click on their remote control to accept. Although promising, adoption of remote control RFI systems is constrained by lack of hardware support and standards. Canoe Ventures – a joint venture with 6 MSOs - was tasked with developing these standards, but in 2012 laid off 120 of 150 employees including its CEO, and closed its New York office. These systems also have the same disadvantages of IPTV, in being unable to track delayed conversions.

2.3 TV Broadcast Time Alignment

Some authors have proposed a "Broadcast Time Alignment" strategy in which TV broadcasts are aligned with website activity within a few seconds of the TV airing [7], [8], [13]. For example, [13] show close time alignment between movie Super Bowl ads and searches for the same movie name. Unfortunately the method can have difficulty attributing small TV broadcasts [8] and is unable to provide insight into delayed conversions.

2.4 Panels

One of the most common fallbacks, when faced with difficult-to-measure effects, is to use volunteer, paid panels to find out what people do after they see the ads. There are several companies that use panels to try to track TV exposures to sales. These include the Nielsen, IRI and TRA panels. One advantage of this method is real-time tracking is possible. However, in all cases, the small size of the panel (e.g. Nielsen 25,000 people) presents formidable challenges for extrapolation and difficulty finding enough transactions to reliably measure sales. Another problem with the panel approach is the cost of maintaining the panels[2].

2.5 Mix Models

If data from previous campaigns has been collected, then it may be possible to regress the historical marketing channel activity (e.g. impressions bought on TV ads, Radio ads, web ads, print, etc) against future sales [8]. Unfortunately, such an approach offers no help if the relationships change in the future – it certainly is not real-time tracking! In addition, historical factors are rarely orthogonal - for example, retailers often execute coordinated advertising across multiple channels correlated in time on purpose in order to exploit seasonal events. This can lead to a historical factors matrix that aliases interactions and even main effects. Even if there are observations in which all main effects vary orthogonally, there may be too few cases for estimation.

2.6 Market Tests

Market Tests overcome the problems of aliasing by creating orthogonal experimental designs to study the phenomena under question. They also overcome the problem of historical data by using conventional scientific testing methods to run TV in some geographic areas and not others, and then compare the sales between the two [5], [11], [16], [17]. An added benefit of Market Tests is that they are easy for independent parties to validate. Mix models are ultimately black boxes which are determining how many conversions to attribute to TV versus other channels – a contentious decision since it impacts other channels, teams and marketing budgets. Market Tests, on the other hand, can be implemented by just one team with a marketing budget. It is also easier to get acceptance from organizational stakeholders, as the results are directly observable. The answer to the question of “how do you know TV is working” is to point to the area that has had TV applied, and its sales data in comparison to its controls.

The problem with Market Tests is their inability to be used during a national campaign, since there are no longer any controls that aren’t receiving the TV signal. Therefore, after getting a nice “research binder” showing the value of TV, once the advertiser starts up their national campaign they once again lose measurement.

The method we describe will update the market testing concept to make it work with national campaigns and using current TV infrastructure. The method enables real-time tracking across all advertiser sales channels without a panel and is implementable by advertisers using their own data without any panel costs.

[2] The authors are working on another approach to television tracking that is panel-like but which uses Set Top Boxes. Further details can be found in [9].

3 Mirrored Tracking Overview

Our problem is as follows: There is a national television ad campaign underway which is injecting $I_N(N)$ impressions into all national areas, and we want to measure its sales effects $Q(N)$. Our strategy for this problem will be to use local ad insertion systems to add impressions $I_L(d)$ to some of the local areas d. We need to add the impressions carefully to ensure that the impression nationally has exactly the same viewership as an equivalent impression locally. This process is what we refer to as "mirroring" and involves careful matching of treatment areas to national, and of local impressions to national. Using each day's observed injected impressions and additional quantity, we can then calculate what impressions are causing in terms of quantity in general.

4 TV Hardware

One of the key aspects of our approach is the use of existing TV capabilities in order to create mirrored tracking cells. A TV video stream is generally compressed using MPEG-2 format with possibly some other embedded instructions. National advertisements can be inserted into national video stream by electronically or manually trafficking the ads and rotation logic to the Network (e.g. ABC, CBS, Fox, NBC, CW) or cable station (e.g. CNN) directly.

Local insertion requires local content to be spliced into a national video stream. Local insertion has been used for many years to create screen overlay "bugs" such as semi-transparent icons at the bottom of the screen that identify the station on news reports. It also allows local stations to splice in their station information to comply with FCC rules which require hourly station identification. Local insertion has also been used to inject ads; in the past this was primarily for local businesses such as local car dealers. We will co-opt the infrastructure to create a mirrored design for a major national campaign.

For Cable stations, the Cable MSO itself inserts the ad into the video stream, and has 2 minutes per hour of possible ads to insert, so approximately 13% of ad inventory. The MSO has multiple levels of signal control. This includes the Cable Interconnect – which cover Direct Marketing Association (DMA) areas, Zone – a collection of about 10,000 households (Fig. 2).

For Broadcast stations, local ad insertion is handled by the local station (e.g. KOMO). 4 minutes per hour of time are provided for local station IDs, as well as local ads, so about 26% of ad inventory is available to be purchased locally.

There are approximately 2,000 cable zone areas that can be purchased, and over 2,500 local broadcast stations, providing considerable ability to create representative treatment mirrors (Fig. 1).

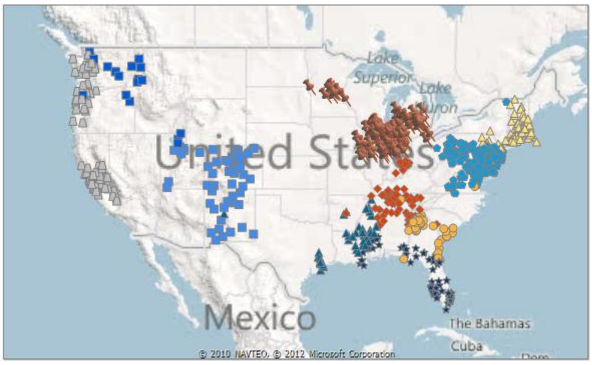

Fig. 1. Geographic distribution of Cable Zones from Comcast.

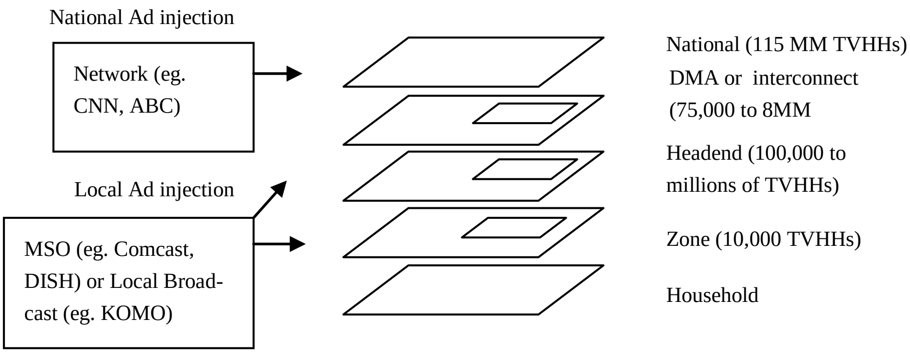

Fig. 2. Local and National ad Insertion for Cable Systems

5 Treatment Area Selection

There are 2,000 cable and broadcast areas d available. Which should we select for local ad injection? The key objective for selecting good treatment areas is to find areas that match national well enough so that they allow for accurate extrapolation to national. Ideally the local areas and national are homogenous populations, and so ads displayed locally have the same effect as is occurring nationally. In order to maximize the chances of homogeneity in the local areas to national, the area needs to meet several criteria:

5.1 Treatment Fitness Criteria

Low Census Disparity from US Average: The mean absolute difference between the i-th US population census demographic $x_i(N)$, and the demographic reading $x_i(d)$ of a particular region d needs to be as low as possible. A lower value indicates that the area is not greatly different from the US average. Zip-code-level demographics are publicly available from the US Census Bureau and these can be aggregated to the same level as the cable and broadcast systems. In the formula below w_i is a weight applied to each demographic.

$$m_1(d) = \sum_i w_i \cdot |x_i(d) - x_i(N)|$$

Average Sales per capita: If a candidate area has sales per capita that are higher than the national average, then it is possible that the area in question might have advertising elasticities which are also different. In order to introduce fewer assumptions or differences into the design, we will therefore favor areas which have sales per capita close to the national average. *Q(d)* = *q*(*d*) / TVHH(*d*) = conversions per capita in area *d*. *q*(*d*) is the quantity of conversions generated in area *d*. TVHH(*d*) are the number of TV Households in area *d*. TVHH(*N*) = 112,000,000 are the number of TV Households nationally.

$$m_2(d) = |Q(d) - Q(N)|$$

Matched Targeting: The targeting of the TV media via the local injection systems needs to match the media being purchased nationally. Targeting is measured by the demographic viewership match between media and the product demographics *r*(*d*) as discussed in Kitts (2013b).

$$m_3(d) = |r(d) - r(N)|$$

Low Volatility: If the treatment area is already subjected to high levels of noise, then the signal we are trying to measure may not be detectible against the area's organic background noise. We measure this as the variance of sales per capita per day in the area.

$$m_4(d) = \frac{1}{T}\sum_t (Q(d,t) - E[Q(d,t)])^2$$

High Geographic dispersion from other experimental areas: It is important to avoid areas which are too close together. Multiple test cells all in the same general geographic area increases the threat that some unique factor in this particular region is influencing sales and elasticities. By spreading out the test cells over a wider area, this threat can be reduced. In addition, increasing the dispersion of tracking cells also even helps avoid spillover of TV broadcasts into neighboring areas, avoiding contamination of other treatment cells. Let the set of possible geographic areas be *G*, and already selected areas $S \subseteq G$. We use the Great Circle method [23] to find the closest already-selected treatment area in Earth Surface distance kilometers, and report this as dispersion from previously selected areas. In the definition below, latitude and longitude are both converted from Cartesian to radians; $d_{lat} = \frac{d_{lat}}{180/\pi}$, and $K = 6378$ is the Earth radius in kilometers.

$$m_0(d_j) = \min(ESD(d_j, e) : \forall e \in d_{1..j-1});$$

$$ESD(d, e) = \text{acos}[\sin(d_{lat}) \sin(e_{lat}) + \cos(d_{lat}) \cos(e_{lat}) \cos(e_{lon-}d_{lon})] \cdot K$$

Low Cost: Cheaper areas allow for more media to be run for the same price. Prices of areas are available from companies which monitor the clearing price of all ad buys on TV. Smaller geographic areas tend to be less in demand and have lower prices, and so are favored for testing over areas such as New York.

$$m_6(d) = TVHH(d) \cdot CPM(d)/1000$$

Average Cable and Satellite penetration: Some areas of the country have lower numbers of cable TVs. We try to avoid selecting areas with unusually low cable adoption rates.

$$m_7(d) = |pen(d) - pen(N)| < PEN$$

Minimum Number of Insertible networks: Insertible networks are stations that can have ads inserted to them. If the number of insertible networks becomes too low, then local inventory may not be able to match national.

$$m_8(d) = \text{sgn}(insert(d) \geq INS)$$

5.2 Treatment Selection Algorithm

Using the factors above, a weighted fitness score is calculated. Other researchers have discussed either "caliper" matching in which only subjects that meet particular significance testing criteria are selected, and "nearest neighbor matching" where subjects that have the closest match to ideal are selected [2], [15], [20]. We favor the nearest neighbor method, and we effectively pick the best *N* areas for mirrored cells.

In order to select areas we use an interactive procedure similar to Leader Clustering [3] in which the first area with best fitness is selected, and then subsequent areas are selected until all treatment cells are selected. Iterative recalculation is needed because the GeoDispersion metric is dependent upon areas already selected. In the formula below, R converts the raw number into a percentile, m_k is a criteria and M_k the weight of that criteria.

$$d_j : \min\left(\sum_k M_k \cdot R\left(m_k(d_j)\right) : d_j \notin d_{1..j-1}\right)$$

6 Control Area Selection

We will be measuring treatment change in quantity per capita versus control change in quantity per capita over the same period of time. In order for this comparison to show differences due to TV (and not other factors), it is critical to ensure that the control area purchase behavior, demographics, and responsiveness to advertising are all as close as possible to the treatment areas [19]. The only difference between the areas should be the application of additional media. Various authors have referred to this as minimizing “threats to validity” or “matching” controls to treatments [2], [15], [20]. We use the following criteria to attempt to ensure homogeneity across multiple dimensions between the control and treatment areas.

6.1 Control Fitness Criteria

Demographic similarity: Controls should have similar demographics to their treatment group. The D_Jth area to be selected has the following match difference:

$$u_1(D_J, d) = \sum_i \left| \left(\frac{1}{J} \sum_{j=1}^{J} x_i(D_j) \right) - x_i(d) \right|$$

Geographic proximity: Whereas treatments were ideally geographically dispersed, the controls should be geographically close to their treatment areas. This helps to ensure that treatment and control areas have the same climactic factors (temperature, precipitation), economic characteristics, population attributes, and so on.

$$u_2(D_J, d) = ESD(D_J, d)$$

Matched movement: The control and treatment areas should both show coordinated movement in sales for an extended period prior to the start of the experiment. When the experimental area has high sales, the control area should have high sales, and vica versa. Systematic variation is a strong test for relatedness since it suggests that the two areas are responding in the same way to changes in environmental conditions, promotions, and other events that can affect sales. In the definition below, the error is proportional to the absolute difference between treatment and the sum of control areas by day. The difference of difference method will also scale the error by national sales, and so we also multiply the difference by national $Q(N, t)^{\eta}$.

$$u_3(D_J, d) = \sum_t Q(N, t)^{\eta} \cdot \left[Q(d, t) - \frac{\sum_j q(D_j, t)}{\sum_j TVHH(D_j, t)} \right]^{\delta}$$

6.2 Control Selection Algorithm

Treatment areas need to utilize local ad injection systems, which restricts feasible treatments to the 2,000 broadcast and cable zone service areas discussed above. However controls – which don't require any media - are under no such obligations. Any controls in the country can be selected. This is useful because it means that controls can be selected at a finer-grain than the treatments. Treatments utilized 2,000 zones, averaging about 55,000 TV households each. Controls can be built from over 30,000 zipcodes, averaging just 3,800 TV Households. Our objective is therefore to assemble a set of controls that match very precisely the demographics of the treatments that we are running in.

The algorithm for selecting controls is iterative, similar to the treatment selection. However, one important difference is that multiple controls are selected for each treatment, and the set of controls are "assembled" to collectively match the treatment. The method starts by selecting the best matching control. Let's say that this control matches well, but has too few African Americans. When selecting the next control area, the error function is the match between the total controls, including the new candidate control, and the treatment. As a result, if one of the candidate controls causes the African American quota to move closer to the treatment, then this control will be favored. As a result, the iterative procedure "self-corrects" and successively selects areas which together have demographics and sales which match the treatment area.

This produces very useful behavior in practice. Figure 3 shows areas chosen after 4 iterations. After 100 iterations, the system has "self-corrected". We found that this kind of population matching behavior has also been critical when faced with real-world events. During one TV ROI Tracking campaign in February 2012 for a music product, we noticed a significant increase in sales in many US geographic areas. Areas lifting appeared to have high African American populations. It turned out that Whitney Houston died on February 11, 2012. Without good matching this would have led to significant distortion in our lift estimates.

The best controls *D(d)* for each treatment group *d* are selected based on the score below. We tend to refer to the control construction process as "annealing".

$$D_J : \min\left(\sum_k U_k \cdot R\left(u_k\left(D_j, d\right)\right) : D_J \notin d_{1..j} \wedge D_J \notin D_{1..J-1}\right)$$

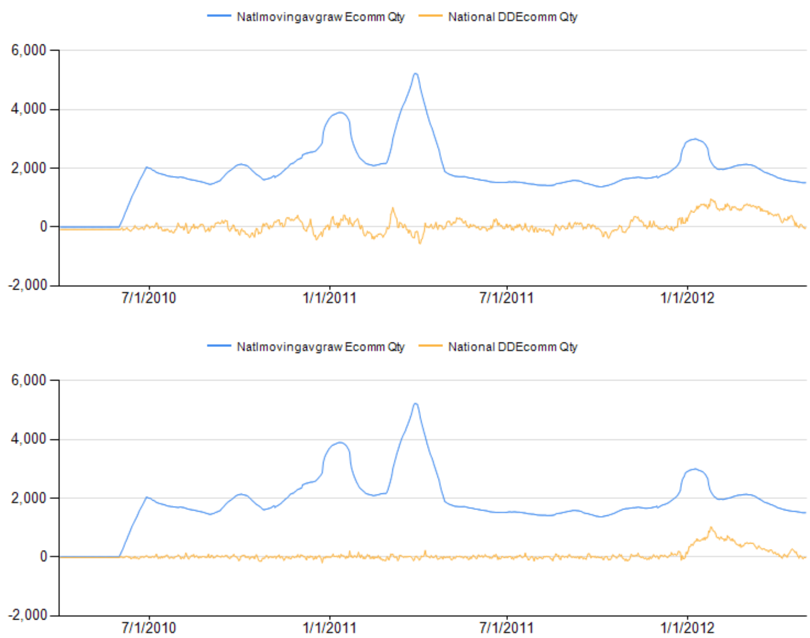

Fig. 3. (top) Estimated conversions due to TV for 4 best matching zip control areas (based on closeness to control). (bottom) 100 zip control areas. Top line is national conversion count, bottom is the attributed conversions due to TV. Tracking cells above begin on 12/12/2012.

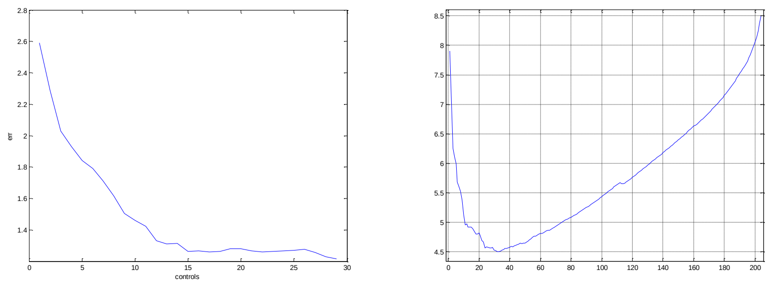

Fig. 4. Error as more DMA controls are added. Optimal number of controls is 29 and then error starts to increase. x-axis is number of controls and y-axis is error.

7 Real-Time Mirrored Estimation

We have now selected treatment areas d_i and multiple control areas $D_j(d_i)$ for each treatment. We will now run additional advertising I_L in the local area d_i. Our objective is to create a detectible increase in the sales per capita in area d, compared to the control areas. Based on the size of the increase, we can then measure how TV advertising is driving sales, and then estimate the unknown (but simultaneously executing) national effects.

Let $\Delta Q_{\mathrm{N}}(d, t_1, t_2)$ be the quantity per capita per week that is being generated in area d between time t_1 and t_2 due to I_N impressions of national TV. This is what we want to estimate. Let E be the quantity per capita per week that is occurring in a local area d without TV; possibly due to other marketing programs and also due to noise. The total quantity that we observe in area d is therefore $Q_{\mathrm{N+}}(d, t_1, t_2) = E + \Delta Q_{\mathrm{N}}$. The quantity per capita per week produced by I_N is an unknown function that we will refer to as f. Our objective is to report on ΔQ_{N} – the quantity due to TV - each week due to running media.

$$\Delta Q_{\mathrm{N+}}(d, t_1, t_2) = \Delta Q_{\mathrm{N}}(d, t_1, t_2) + E = f(I_N) + E$$

In order to make the quantity measurable, we will inject an additional amount of local impressions per capita per week I_L into a local area d using the local ad insertion systems which produce local revenue per capita per week of $\Delta Q_{\mathrm{L}}(d, t_1, t_2)$. Let $\Delta Q_{\mathrm{L+}}(d, t_1, t_2)$ be the total revenue now observed in the local area inclusive of local and national ads. $Q_{\mathrm{L+}}(d, t_1, t_2) = \Delta Q_{\mathrm{L}}(d, t_1, t_2) + \Delta Q_{\mathrm{N}}(d, t_1, t_2) + E$ We now have:

$$\Delta Q_{\mathrm{L}}(d, t_1, t_2) = \Delta Q_{\mathrm{L+}}(d, t_1, t_2) - \Delta Q_{\mathrm{N+}}(d, t_1, t_2)$$

In the above formula, $\Delta Q_{\mathrm{L+}}(d, t_1, t_2)$, the quantity per capita per week in the local area, is observable. However $\Delta Q_{\mathrm{N+}}(d, t_1, t_2)$, the quantity that would have occurred, is not directly observable since we over-concentrated in d. However, if we now find matched areas $D(d)$ that are homogenous with d, then we can use their performance (which is only based on national ad insertion) for $\Delta Q_{\mathrm{N+}}$ since it is observable. We now have:

$$\Delta Q_{\mathrm{L}}(d, t_1, t_2) = \Delta Q_{\mathrm{L+}}(d, t_1, t_2) - \Delta Q_{\mathrm{N+}}(D(d), t_1, t_2)$$

$$\Delta Q_{\mathrm{L}}(d, t_1, t_2) = \Delta Q_{\mathrm{L+}}(d, t_2) - \Delta Q_{\mathrm{L+}}(d, t_1) - \Delta Q_{\mathrm{N+}}(D(d), t_2) - \Delta Q_{\mathrm{N+}}(D(d), t_1)$$

Where $Q_{\mathrm{L+}}$ and $Q_{\mathrm{N+}}$ are both observable and we injected I_{L} impression concentration. We therefore have an observation between impressions and quantity a point "higher" than the national impressions. We want to try to infer the same relationship, but at a point "lower". In order to infer ΔQ_N , which is running with I_{N}, we need to know something about the shape of the TV impression to quantity function f. One property that we can determine *a priori* is that $f(0) = 0$. If 0 impressions are injected into an area, then the revenue due to those 0 ad injections will also be zero. We can also assert that $f(I_N) \leq Q_{\mathrm{N+}}(d, t_1, t_2)$ since it is not possible to produce more quantity than what were observed.

What shape should we assume for f? There is a large body of advertising research which shows diminishing returns at higher levels of advertising [4], [6], [22], [24]. Ordinarily, a linear assumption for an advertising response function would lead to unrealistically optimistic estimates. However, in this case, we are actually "extrapolating downwards" and the diminishing returns observation works to our advantage.

Assuming diminishing returns as advertising impressions increase, a linear fit to the observed data at the higher concentration level, actually becomes a lower bound on the lift produced by the national ad impressions. Therefore, the simple linear model results in conservative – and in fact lower bound - estimates of TV effects. This is a convenient result that allows us to be more confident about our TV effects.

In accordance with the observations above, a linear function should be intercept-less and so we have a function form of $f(I) = cI$, we can calculate an estimate for the function as follows:

$$\Delta Q_{\mathrm{N}}(t_1, t_2) = c \cdot I_N \text{ where } c = \frac{\Delta Q_{\mathrm{L}}(d_j, t_1, t_2)}{I_L}$$

$$\text{if } \Delta Q_N \leq 0 \text{ then } \Delta Q_N = 0 \text{ ; if } \Delta Q_N \geq \Delta Q_{N+} \text{ then } \Delta Q_N = \Delta Q_{N+}$$

8 Experiment

We used the technique to achieve a television-drive-to-web Cost Per Acquisition (CPA) goal from September 10, 2012 to March 1, 2013: a period of 172 days.

The CPA objective was to spend $45 per signup to their website. This kind of goal is typical for online advertising with cookie based conversion tracking systems, but for TV, it is virtually unheard of to be able to actively target and achieve a CPA.

Our tracking markets consisted of three areas in sequence (a) Lacrosse,WI-MN, (b) Tulsa, OK and Charleston, SC, (c) Harrisburg, PA and Milwaukee, WI. We used Lacrosse from 9/10/2012 to 10/31/2012, Tulsa and Charleston from 11/12/2012 to 1/6/2013, and Harrisburg and Milwaukee from 1/7/2013 to 3/1/2013. All areas were near the 50$^{\text{th}}$ percentile for sales per capita and better than 30$^{\text{th}}$ percentile for census disparity to US. Each area was matched to 300 zip controls.

The reason for using three areas was because we found that over time the lift in the areas began to decline. In order to ensure our lift readings remained accurate for national, we rotated the tracking areas every couple of months.

Local ads were targeted to match the tratios of the national campaign, by selecting relevant programming to run the ads during. The disparity between tratio of local and national was negligible – both were about tratio = 0.265. National media was maintained at approximately 345 impressions per thousand households and approximately $8,500 per day, although we increased this significantly over Christmas. Local markets were executed at approximately 400 impressions per thousand households and matched tratio.

Our initial market Lacrosse performed relatively poorly. It quickly lost lift after two weeks, and we brought Tulsa and Charleston online starting in November.

Overall, inclusive of LaCrosse, we produced 29,811 signups from $1.797 million dollars in media spending at a CPA of $60.29. The 29,911 signups boosted sales by nearly 11%. We also recorded a CPA of $45 after November.

In addition, because of our use of treatment-control tracking markets, after leaving the tracking markets we were able to monitor delayed conversions in the form of elevated lift after switching off TV. By summing incremental conversions versus controls, we were able to report that conversions added 3.3 times the number of initially injected conversion by the end of year.

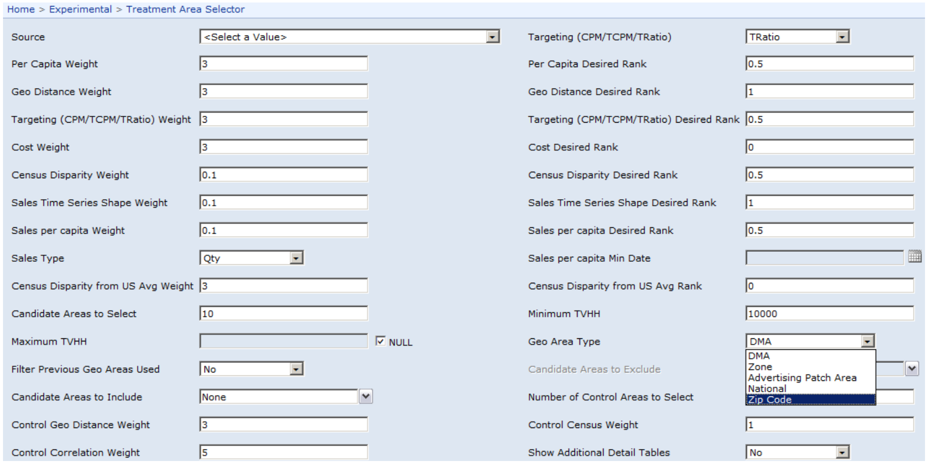

Fig. 5. Screenshot of Treatment Area Selector. Left-hand-side comprise fitness criteria. Right-hand-side show the user-defined percentile target for each criterion. For example “Sales Per Capita Weight” is the weight to use for matching “Sales Per capita” between treatment and national. “Sales Per Capita Desired Rank” is the ideal rank (0.5 or 50th percentile). “Source” refers to the customer set that is being targeted.

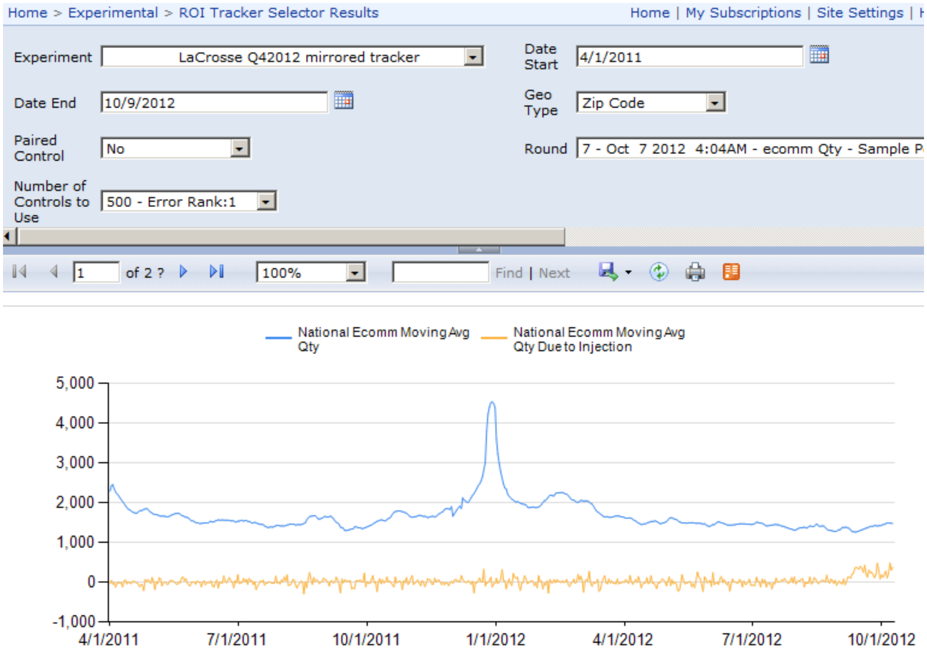

Fig. 6. Screenshot of ROI Tracking GUI. Top line is national conversion count. Bottom line is the attributed, real-time conversion amount occurring due to a running TV campaign.

9 Conclusion

We have described a method for tracking Television ROI in real-time. The method does not rely upon panels, and can be implemented by an advertiser using existing television infrastructure and without their own sales data. The technique also measures delayed conversions, which is a key aspect of television campaigns. We believe that this is useful technique for tracking real-time and delayed ROI in what is often considered an untrackable medium.

References

1. Angrist, J. and Pischke, J: Mostly Harmless Econometrics, Princeton University Press (2010)
2. D'Agostino, R.: Tutorial In Biostatistics: Propensity Score Methods for Bias Reduction in the Comparison of a Treatment to a Non-Randomized Control Group, Statistics in Medicine, Vol 17, pp. 2265-2281. (1998)
3. Duda, R., Hart, P., Stork, D.: Pattern Classification (second edition), Wiley (2000)
4. Hanssens, D., Parsons, L., Schultz, R.: Market Response Models: Econometric and Time Series Analysis, Kluwer Academic Press, Boston (2001)

5. Hu, Y., Lodish, L., Krieger, M.: An Analysis of Real World TV Advertising Tests: a 15 year update, Journal of Advertising Research, Vol. 47, No. 3, pp. 341-353. (2007)
6. Johansson, J. K.: Advertising and the S-Curve: A New Approach, Journal of Marketing Research, Vol. 16, No. 3, pp. 346-354. American Marketing Association (1979)
7. Joo, M., Wilbur, K., Zhu, Y.: Television Advertising and Online Search, SSRN Working paper, http://ssrn.com/abstract=1720713 (2012)
8. Kitts, B., Wei, L., Au, D., Powter, A., Burdick, B.: Attribution of Conversion Events to Multi-Channel Media, In: Proceedings of the Tenth IEEE International Conference on Data Mining, December 14-17, (2010).
9. Kitts, B., Au, D., Burdick, B.: Television Conversion Tracking using Anonymized Set Top Box Data, unpublished. (2013)
10. Kokernak, M.: What's Television's Next Business Model? Media Post Daily News, March 17, (2010)
11. Lambert, D. and Pregibon, D.: Online effects of Offline Ads, In: Proceedings of the Second International Workshop on Dataa Mining and Audience Intelligence for Advertising, ACM Press, New York (2008)
12. Leaders: In Praise of Television: The great survivor, The Economist, April (2010)
13. Lewis, R. and Reiley, D.: Down-to-the-Minute Effects of Super Bowl Advertising on Online Search Behavior, Working Paper. http://davidreiley.com/papers/DownToTheMinute.pdf (2010)
14. Lewis, R. and Rao, J.: On the Near Impossibility of Measuring Advertising Effectiveness, Yahoo! Research working paper, http://justinmrao.com/lewis_rao_nearimpossibility.pdf (2011)
15. Little, R. and Rubin, D.: Causal Effects in Clinical and Epidemiological Studies Via Potential Outcomes: Concepts and Analytical Approaches, Annual Review of Public Health, Vol. 21, pp. 121-145, (2000).
16. Lodish, L., Abraham, M., Kalmenson, S., Livelsberger, J., Lubetkin, B., Richardson, B., Stevens, M.: How T.V. Advertising Works: A Meta-Analysis of 389 Real World Split Cable T.V. Advertising Experiments, Journal of Marketing Research, Vol. 32, No. 2, pp. 125-139. (1995a)
17. Lodish, L., Abraham, M., Kalmenson, S., Livelsberger, J., Lubetkin, B., Richardson, B., Stevens, M.: A Summary of Fifty-five In-Market Experimetns on the Long-term Effect of TV Advertising, Marketing Science, Vol. 14, No. 3, pp. 133-140. (1995b)
18. Nelson-Field, K, Riebe, E and Sharp, B.: What's Not To Like?; Can a Facebook Fan Base give a Brand the Advertising Reach it needs?, Journal of Advertising Research (2012)
19. Rosenbaum, P. and Rubin, D.: Constructing a Control Group Using Multivariate Matched Sampling Methods That Incorporate the Propensity Score, The American Statistician, Vol. 39, No. 1. (1985)
20. Rubin, D.: Estimating Causal Effects of Treatments in Randomized and Nonrandomized Studies, Journal of Educational Psychology, Vol. 66, No. 5, pp. 688-701, (1974)
21. Simon, H.: Price Management, North-Holland Publishing Company, Amsterdam. (1989)
22. Simon, J.L. and Arndt, J.: The Shape of the Advertising Response Function, Journal of Advertising Research, Volume 20, Number 4, pp. 767 (2002)
23. Sinnott, R.: Virtues of the Haversine, Sky and Telescope, Vol. 68, No. 2, p. 159 (1984)
24. Vakratsas, D., Feinberg, F., Bass, F., and Kalyanaram, G.: Advertising Response Functions Revisited, Marketing Science, 23(1), pp. 109–119, (2004)