11660-2016

Redesigning Control Using the GMATCH Algorithm to Isolate Impact of a Specific Marketing Intervention from Overlapping Solicitations

Mou Dutta, Arjun Natarajan - Genpact LLC (Analytics and Research)

ABSTRACT

The success of any marketing promotion is measured by the incremental response and revenue generated by the targeted population known as Test in comparison with the holdout sample known as Control. An unbiased random Test and Control sampling ensures that the incremental revenue is in fact driven by the marketing intervention. However, isolating the true incremental effect of any particular marketing intervention becomes increasingly challenging in the face of overlapping marketing solicitations. This paper demonstrates how a look-alike model can be applied using the GMATCH algorithm on a SAS® platform to design a truly comparable control group to accurately measure and isolate the impact of a specific marketing intervention.

INTRODUCTION

Marketing promotions are crucial to retain customers in a competitive business environment. Exclusive promotions drive customer delight and get businesses, incremental revenue. Once the promotion-eligible population is determined, random sampling is applied to create test and control groups. The idea of creating a hold-out control from the promotion-eligible population is to measure the success of the marketing promotion in terms of incremental revenue from the test group. Random sampling ensures every member of the population is equally likely to be selected and the sample is a true representation of the population. For example – if the average age of the population is 45, the random sample will be a true representative of the population if the average age in the sample is also around 45. PROC SURVEYSELECT is an easy and efficient way to design such random test and control groups. However, what happens if the control group from one marketing solicitation becomes a part of test group of another marketing solicitation with overlapping promotion window? This paper discusses a case study of such overlapping marketing interventions for a major US bank where a premier marketing promotion was followed by a card reissue program which had to include control group from the premier promotion. As a result, test group of the premier marketing promotion could no longer have a true comparable control group. Greedy algorithm on SAS® platform was applied instead to re-create a truly comparable test-control group to measure the success of the premier promotion.

METHODOLOGY & RESULTS

Step 1 – Design Test-Control using simple random sampling for Premier Marketing Promotion, henceforth called Premier Test and Premier Control

Proc surveyselect data=data1 method=SRS n=x out=sample1; run;

Step 2 – measure performance of Premier Test and Premier Control during premier promotion window

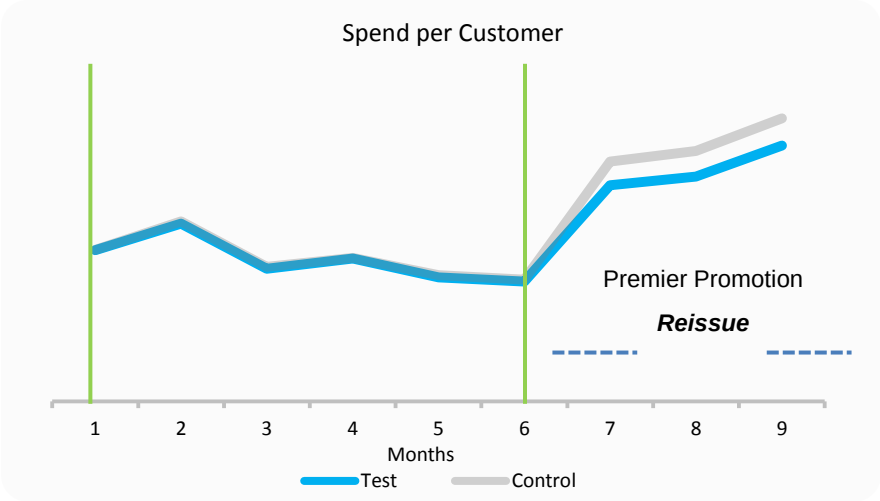


Figure 1 Spend per Customer Test versus Control by Month

Figure 1 Spend per Customer Test versus Control by MonthConfirms Premier Test and Premier Control were performing similarly during pre-promotion period. It was expected that with premier marketing intervention, Premier Test would generate lift in spending compared to Premier Control who did not receive the communication. Instead, Premier Control started outperforming Premier Test consistently as premier promotion overlapped with Reissue solicitation in which some members of Premier Control group had to be included as well. The overlapping marketing interventions inadvertently corrupted the sanctity of the test control study. Thereby not allowing us to draw any conclusion on the profitability of the campaign under consideration

Step 3 – Re-design Premier Test and Premier Control to remove Reissue Effect Figure 4Step 3.a – Exclude Reissued members from Premier Control Group, henceforth called Premier New Control

Step 3.b – Apply look-alike modeling to pair match Premier Test and Premier New Control

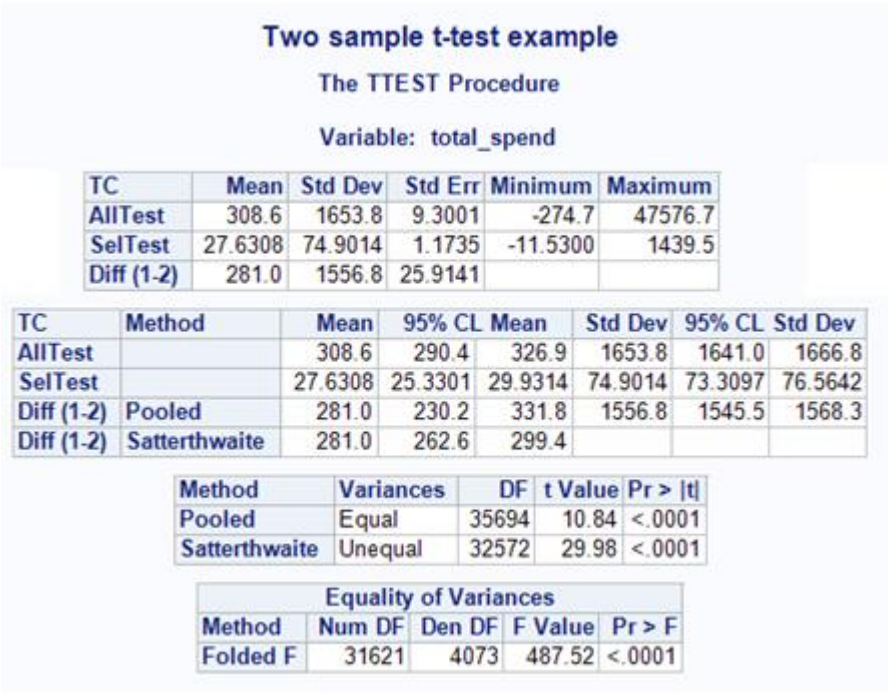
Look-alike Model Variables	Description
Spend_6M	Total spend amount during pre 6 months of premier promotion
Trans_6M	Total number of transactions pre 6 months of premier promotion
MOB	Months on books
Active_6M	Number of months active pre 6 months of premier promotion

Table 1 Variables considered for the look-alike model

Gmatch macro code instance:

```
%gmatch(
              data=Stdze_base1,
/*SAS Data set with the test and potential pool of control*/
              group=test_control,
/*Variable distinguishing test from control*/
              id=Customer_id,
/*ID variable*/
              mvars=
/*List of variables used to select the control */
              mob
/*Variable1: Months on Books*/
              Spend_6M
/*Variable2: Total spend amount during pre 6 months of premier promotion*/
              Trans_6M
/*Variable3: Total number of transactions pre 6 months of premier promotion*/
              Active_6M ,
/*Variable4: Number of months active pre 6 months of premier promotion*/
              wts= 1 1 1 1,
/*weights corresponding to each variable*/
              dmaxk= 0 5 1 0,
/*largest possible absolute difference acceptable for a match*/
              transf=0,
/*No standardization required*/
              time=,
/*No time variable used*/
              dist=2,
/*Euclidean distance used to calculate the proximity of test and control*/
              ncontls=1,
/*indicates the number of control accounts to be selected for every test*/
              seedca=234098,
/*seed value used to randomly sort the test prior to matching*/
              seedco=0489,
/*seed value used to randomly sort the control prior to matching*/
              out=Greedy_Output,
/*The output data set with the paired accounts*/
              outnmco=Non_Matched4,
/*Test accounts for which a corresponding pair was not found under the specified
conditions*/
print=n);
```

Step 3.c – Use t-test to check if the new test is true representation of the Premier Test.



Two sample t-test example

The TTEST Procedure

Variable: total_spend

TC	Mean	Std Dev	Std Err	Minimum	Maximum
AllTest	308.6	1653.8	9.3001	-274.7	47576.7
SelTest	27.6308	74.9014	1.1735	-11.5300	1439.5
Diff (1-2)	281.0	1556.8	25.9141		

TC	Method	Mean	95% CL Mean		Std Dev	95% CL Std Dev	
AllTest		308.6	290.4	326.9	1653.8	1641.0	1666.8
SelTest		27.6308	25.3301	29.9314	74.9014	73.3097	76.5642
Diff (1-2)	Pooled	281.0	230.2	331.8	1556.8	1545.5	1568.3
Diff (1-2)	Satterthwaite	281.0	262.6	299.4			

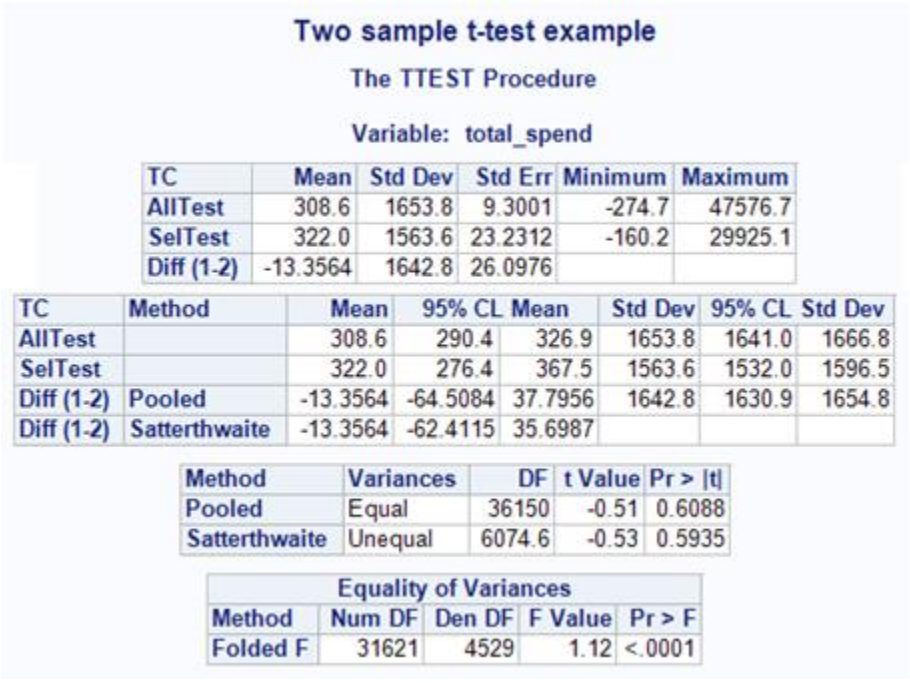
Method	Variances	DF	t Value	Pr > |t|
Pooled	Equal	35694	10.84	<.0001
Satterthwaite	Unequal	32572	29.98	<.0001

Equality of Variances				
Method	Num DF	Den DF	F Value	Pr > F
Folded F	31621	4073	487.52	<.0001

Figure 2 Two Sample t-test output Two sample t-test results with P<.001 from which we can infer that the selected test differs from the Premier test significantly

Figure 2 Two Sample t-test output before applying look-alike modeling

Step 3.d – If the selected test is not a true representation of the population, apply targeted GMATCH to offset the biasedness in selected test and repeat the process till we ensure that the selected test and the Premier test are statistically identical.



Two sample t-test example

The TTEST Procedure

Variable: total_spend

TC	Mean	Std Dev	Std Err	Minimum	Maximum
AllTest	308.6	1653.8	9.3001	-274.7	47576.7
SelTest	322.0	1563.6	23.2312	-160.2	29925.1
Diff (1-2)	-13.3564	1642.8	26.0976		

TC	Method	Mean	95% CL Mean		Std Dev	95% CL Std Dev	
AllTest		308.6	290.4	326.9	1653.8	1641.0	1666.8
SelTest		322.0	276.4	367.5	1563.6	1532.0	1596.5
Diff (1-2)	Pooled	-13.3564	-64.5084	37.7956	1642.8	1630.9	1654.8
Diff (1-2)	Satterthwaite	-13.3564	-62.4115	35.6987			

Method	Variances	DF	t Value	Pr > |t|
Pooled	Equal	36150	-0.51	0.6088
Satterthwaite	Unequal	6074.6	-0.53	0.5935

Equality of Variances				
Method	Num DF	Den DF	F Value	Pr > F
Folded F	31621	4529	1.12	<.0001

Figure 3 Two Sample t-test output after applying look-alike modelingTwo sample t-test results after targeted GMATCH with P>0.001 from which we can infer that the selected test and Premier test are statistically similar

Figure 3 Two Sample t-test output after applying look-alike modeling

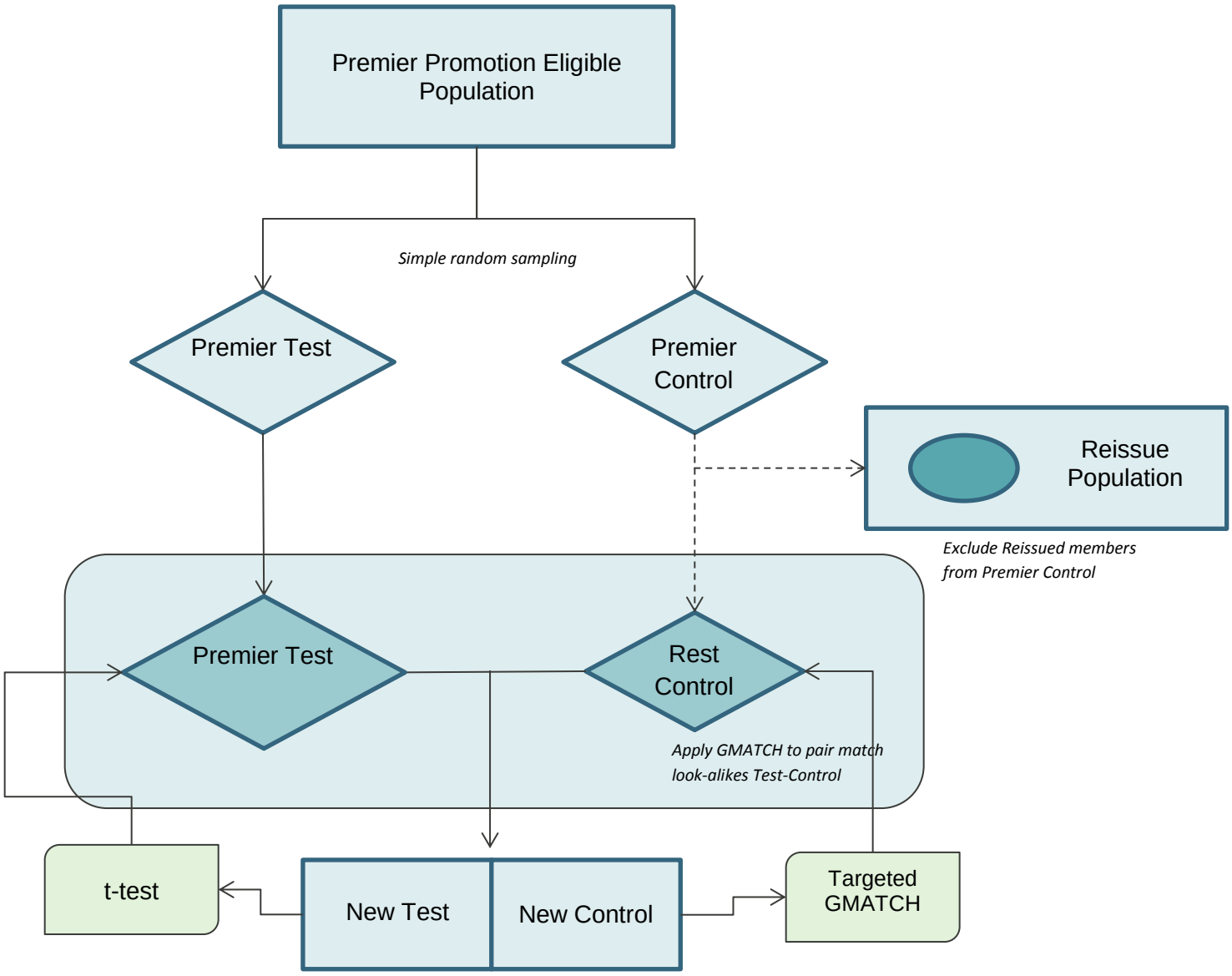


Figure 4 Redesigned Process Flow

Step 4 – re-measure performance of Premier New Test and Premier New Control during Premier Promotion window

Figure 5 Premier New Test and Premier New Control spent similar during pre 6 months of premier promotion. The New Control doesn't include the solicitation effect of Reissue anymore and the corresponding New Test generated basis look-alike modeling ensures both the groups are comparable. The difference in spending between New Test and New Control observed during premier promotion window can purely be attributed to this specific marketing touch point and hence helps the business to assess the incremental spend and also making a decision on possibility to continue this type of marketing intervention in future

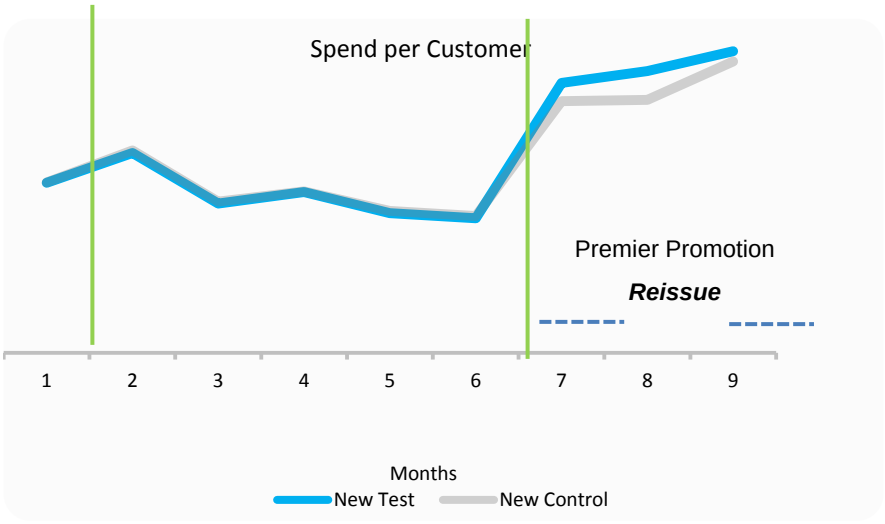


Figure 5 Spend per Customer Test versus Control by Month post new control design

DISCUSSIONS & FUTURE RESEARCH

Application of randomized test control design is standard practice in marketing solicitations. Matching algorithm is used in context of observational studies in the field of epidemiology, social sciences and clinical trials where specific treatments cannot be decided basis randomized control trials. This paper discusses a unique scenario in the world of marketing where marketing interventions inadvertently overlapped each other due to regulatory mandate thereby disturbing the randomness of test and control groups. Look-alike modeling based on GMATCH algorithm was applied on SAS platform to create comparable test and control groups in order to draw inference on the effectiveness of the marketing solicitation.

This study aims to explore, learn and seek successful application of alternative approaches on test control design from other fields of research. This also leads to develop thoughts for academics and corporate researches in the field of marketing around how to best manage targeting, selection and randomized trials. In a real world scenario, businesses have to constantly communicate to their consumers in an innovative and intelligent way to keep up with the intense competition by the other market players in the industry. Hence it is of utmost importance to invest wisely and be able to measure marketing promotions accurately to decide future course of actions. On the other hand, frequent and large scale campaign execution requires a simplified approach with an ease of implementation and quick turn-around time. Pair matching algorithm ensures the robustness of test and control similarity and facilitates improved decision making, however, at the cost of increased execution time. We have considered looking in to these potential future research areas to seek applicability of pair matching algorithm in mainstream campaign management.

ACKNOWLEDGMENT

GMATCH Algorithm SAS Code Reference - SAS paper 214-26 Reducing Bias in a Propensity Score Matched-Pair Sample Using Greedy Matching Techniques

REFERENCES

Kawabata Hugh,Tran Michelle and Hines Patricia May 2004. “Using SASÒ to Match Cases for Case Control Studies, Paper 173-29.” *Proceedings of the SUGI 29 2004*, Montreal, Canada SAS Institute Inc. Available at http://www2.sas.com/proceedings/sugi29/173-29.pdf

Parsons S. Lori May 2004. “Performing a 1:N Case-Control Match on Propensity Score, Paper 088-2012.” *Proceedings of the SUGI 29 2004*, Montreal, Canada SAS Institute Inc. Available at http://www2.sas.com/proceedings/sugi29/165-29.pdf

Wang Zhiwei April 2012. “Optimized 1 : N Case-Control Match using SAS, Paper 165-29.” *Proceedings of the SAS Global Forum 2012*, Orlando FL SAS Institute Inc. Available at http://support.sas.com/resources/papers/proceedings12/088-2012.pdf

Li David April 2012. “Building Match Code Using SAS®, Paper 081-2011.” *Proceedings of the SAS Global Forum 2011*, Las Vegas NV SAS Institute Inc. Available at http://support.sas.com/resources/papers/proceedings11/081-2011.pdf

Parsons S. Lori. “Reducing Bias in a Propensity Score Matched-Pair Sample Using Greedy Matching Techniques, Paper 214-26.” *Proceedings of the SUGI 26*, SAS Institute Inc. Available at http://www2.sas.com/proceedings/sugi26/p214-26.pdf

Grandits Greg, Neuhaus Jacqueline April 2010. “Using SAS® to Perform Individual Matching in Design of Case-Control Studies, Paper 061-2010.” *Proceedings of the SAS Global Forum 2010*, Seattle, Washington SAS Institute Inc. Available at http://support.sas.com/resources/papers/proceedings10/061-2010.pdf

Mandrekar N. Jayawant and Mandrekar J. Sumithra April 2010. “An Introduction to Matching and its Application using SAS®, Paper 208-29.” *Proceedings of the SUGI 29 2004*, Montreal, Canada SAS Institute Inc. SAS Institute Inc. Available at http://www2.sas.com/proceedings/sugi29/208-29.pdf

Mayo Clinic Biomedical Statistics and Informatics http://www.mayo.edu/research/departments-divisions/department-health-sciences-research/division-biomedical-statistics-informatics/software/locally-written-sas-macros

ACKNOWLEDGMENTS

GMATCH Algorithm SAS Code Reference - SAS paper 214-26 Reducing Bias in a Propensity Score Matched-Pair Sample Using Greedy Matching Techniques

CONTACT INFORMATION

Your comments and questions are valued and encouraged. Contact the author at:

Mou Dutta
Genpact LLC, Analytics & Research
42, Old Ridgebury Rd.
Danbury, CT – 06810
Phone: +1 203 730 5130

Arjun Natarajan
Genpact LLC, Analytics & Research
42, Old Ridgebury Rd.
Danbury, CT – 06810
Phone: +1 203 730 5130