Multi-Armed Bandit Approaches in Online Advertising Design when Facing Spillover Effects

Siyi Gao[1]

[1]Columbia university, Department of statistics, New York, US
[1]e-mail: sg3920@columbia.edu

ABSTRACT

With the ever-growing users of the Internet, online marketing activities have become more and more popular in recent years. Marketing companies are often faced with the problem of choosing the best online advertising strategy. The most popular experimental design method is the A/B test, but recently, with the improvement of computational power and the update of algorithms, people are beginning to pay attention to adaptive experimental design. Thompson Sampling algorithm is a heuristic and efficient method to solve this kind of multi-armed bandit problem. In practical applications, if advertising on social media, we need to consider the impact of social network, since people chat with each other on social media every day and share information very quickly. This article mainly focus on exploring the influence of social network spillover effect on Thompson Sampling method. I first use data collected by McAuley and Leskovec from Facebook to simulate social network graph and assume that the people on this social graph constitute the subject pool that I want to add different treatments. After that, I repeatedly simulated each setting 5000 times to explore the performance of TS and static design under different numbers of batches, first batch sizes, various best arm conversion rates and degrees of spillover effects. Through simulations, this paper finds that in the case of social network spillover effect, TS and static design methods have biased ATE estimates, and the accuracy of TS selecting the correct best arm is more affected than that of static design.

Key words: Multi-armed bandit, Thompson sampling, Spillover effects, Social network, Online advertising
Data repository: Please check the social network data and python codes for simulations at OSF repository `https://osf.io/wa3hz/?view_only=7ea67583bf4b418d84e8ca3f8737080b`

Introduction

In the past few years, with the popularization of the Internet, the Internet has been integrated into people's lives(Busca and Bertrandias, 2020[1]). With it comes the popularity of online advertising marketing activities. The rising popularity of digital media and the massive increase of digital media platforms also foster online marketing to become the most mainstream marketing method for many corporations today(Ko, 2019[2]). According to Mandal (2017)[3], by 2021, 73% of e-commerce sales are generated through the mobile platform. When we log into social media apps and browse messages, we sometimes can see various advertising pop-up windows.

Online advertising is favored by many companies because it is fast, convenient and relatively low-cost(Hauser et al., 2014[4]). While reviewing the literature, we can see that identifying the best choice is always a hot research topic. People pay great attention to choosing the best advertising strategy. Everyone wants to spend their money where it matters most. Advertisers in online campaigns often have to choose the most efficient one among several ads or channels(Faruk et al., 2021[5]). Thanks to the digital environments, randomization is applicable in online advertising experimental designs, and A/B test is the most common method in practise(Schwartz et al., 2017[6]). People often use this method to design and test several different advertising designs, and determine which advertising type has the largest conversion rate through carrying out experiments. In this way, the best advertisement type is selected for mass delivery. However, if we have a long list of interventions under consideration, such as various slogans, hierarchical structures, attributes of actions etc(Lambrecht and Tucker (2013)[7]; Goldfarb and Tucker (2011)[8]), it is costly and time-consuming to conduct A/B test among all the treatments. We can frame selecting the best strategy among various candidates into a multi-armed bandit problem(Robbins, 1952[9]). In the multi-armed bandit problem, people not only want to choose the optimal strategy, but also want to assign more test subjects to the treatment of the optimal strategy. Exploiting while exploring is the thought of adaptive experimental design. The most obvious difference between adaptive experimental design and static experimental design is that, we no longer equally allocate impressions to each advertising strategy throughout the trail, but instead dynamically update allocation probabilities based on observed outcomes, investing an ever-growing share of the subject pool in more promising treatment arms(Schwartz et al., 2017[6]). According to many previous studies(Keller and Oldale (2003)[10]; May et al. (2011)[11]; Scott (2010)[12]), we can find response-adaptive trials increase the experiment efficiency.

In practice applications, when allocating treatments, it is impossible to assume all participants are isolated(Aronow and Samii, 2017[13]). According to consumer-to-consumer word-of-mouth(WOM) concept, consumers' behaviors are affected by others(Chae et al., 2017[14]). Learn from this idea, we can assume that when displaying online advertisement, spillover effects through social networks do exist. Those who are treated in the experiment have chances to say things of the product to their peers, and then if their peers receive treatment in later stage, their conversion rates might be different due to the previous exposure.

In this paper, I attempt to compare the accuracy of TS and static design methods in selecting the optimal strategy in the presence of social network spillover effect, and whether the estimation of ATE is affected. In the rest of the article, I first explain the idea of Thompson Sampling and how to simulate a social network, then I introduce my simulation plans and display my simulation results.

Theory Background

Introduction of Thompson Sampling

The core idea of response-adaptive experimental design is the trade off between exploration and exploitation, or in other words, earning through learning. Such kinds of exploration and exploitation problems are first posed by Thompson (1933)[15]. In the framework of this multi-armed bandit problem, experimenters are supposed to sequentially allocate finite units across multiple treatment arms. At each trial period, the results of each treatment can be observed, and the known results can be used to estimate possible conversion rate, and based on this inference, the next stage of experimental sample allocation is made. One common algorithm to achieve this purpose is Thompson Sampling.

Thompson Sampling method has great advantages. The most obvious feature of Thompson sampling is that it is intuitive, easy to understand and easy to operate. Thompson sampling is a heuristic but effective and efficient approach to navigate the exploration–exploitation trade-off. Secondly, its advantage is that in each stage, more than one unit is selected, but n individuals are selected and assigned to each arm according to a predetermined probability.

In this paper, I use binary rewards, since the observation from advertising campaigns is either a success or a failure(people purchase or not), denoted as $x \in \{0,1\}$, to illustrate how Thompson Sampling algorithm works. Suppose I have K arms with unknown set of conversion rates $\theta_1,...,\theta_k$. The result(how many success cases) of each arm is generated by binomial distribution with parameters n denoted as assigned samples and p denoted as conversion rate. Conversion rates $\theta_1,...,\theta_k$ follow Beta distribution Beta(α,β). I set the prior Beta distribution as Beta$(1,1)$. In every stage(batch), there are n samples waiting to be allocated into these K arms. Before assigning subjects at period $t+1$, I first observe the results from period t. I observe number of assigned samples into each arm in period t, denoted as $n_{k,t}$ and number of successes of each arm in period t, denoted as $x_{k,t}$, then update the posterior Beta distribution as Beta$(1+\sum_1^t x_{k,t}, 1+\sum_1^t n_{k,t} - x_{k,t})$ for each arm's conversion rate. Then in the next period $t+1$, samples are randomly assigned according to the posterior probabilities of θ_k being the largest.

Finite Sample Unbiased Estimator

According to Offer-Westort et al.,(2021)[16], I use S_i to represent the history of treatment assignments and outcomes observed before assigning treatment to subject i. In each batch, the history should be same for all observations. For treatment K_i, I have an assignment probability denoted as $\pi_i(k;S_i) = Pr[K_i = k|S_i]$. Suppose there are N samples in each batch, then the unbiased estimator of conversion rate is

$$\hat{\mu}_k^{TS} = \frac{1}{N}\sum_{i=1}^{N} Y_i \frac{I[K_i = k]}{\pi_i(k; S_i)}$$

.

Spillover Effects through Social Networks

In this article, I want to study the case that the spillover effect between treatment groups and control groups is realized through social networks. In other words, it is not a direct effect between the treatment measures of the experimental groups, but rather an interaction between subjects who participated in the experiment and those who did not. If the experiment is only one round, there will be no social network-based spillover effects since subjects don't affect each other in the same round. However, if the experiment is conducted in multiple rounds, those who have participated in the experiment before may pass on the impact of the experiment to their online friends through chat or other means. If these friends are selected to participate in one of the next rounds, their previous influence from friends will affect the results of the experiment.

For example, we want to select the best advertisement for a new product and display it on social media. When conducting experiments on social media platforms, we randomly select users on social media and tag each of them with a user ID and then randomly assign them to different experimental groups, and observe the experimental results. No matter how many rounds the experiment takes, we preselect all the subjects. Suppose we have 1000 subjects and 10 experimental groups. If the experiment takes only one round, we assign 100 participants to each experimental group. If the experiment takes 10 rounds, according to the idea of static design, in each round, we assign 10 subjects to each experimental group. Each subject will be only selected in one experimental group and attend only one round of the experiment. There is a time interval between each round of the experiment, which provides the possibility for communication between subjects. Assuming user 1 and user 10 are good friends and both are selected into the experiment, user 1 has participated in a certain experimental group in the first round, and user 10 will participate in a certain experimental group in the fourth round. After participating in the experiment, user 1 may share with user 10 something about the new product like views or feelings. In this way, user 10 has received some treatment influences on the product before participating in the experiment. Then, the spillover effect between the experimental groups exists.

Since the spillover effects based on social networks may be positive or negative, strong or weak, and exist between direct friends or indirect friends, there are too many situations that need to be simulated. The computing power of the computer does not support running all the simulations, so in this paper I only chose the positive spillover effect between direct friends for research. I assume that a positive spillover effect exists between each pair of direct friends, and I use the name "augmentation value" to represent the spillover effect in the simulation studies in this paper.

Experiment Designs

Social Network Simulation

I use the python NetworkX package to generate the social network graph based on the Facebook data collected by McAuley and Leskovec(2012)[17]. There are 4039 nodes and 88234 edges in the graph. In this paper, I suppose that spillover effect only exists between the directly connected nodes. The plot below is a simple representation of the generated social network.

Simulation Plans

In this paper, for all the studies, I use a fixed sample size of 4000 and 9 treatments for all experiments. The 4000 samples are divided into varying numbers of batches. In each batch, the samples are assigned to one of the 9 treatments. In reality, the batches can be seen as rounds of advertising, and various treatments can be seen as different versions of advertisements. The treatment assignments follow either an adaptive or a static design. The adaptive design assigns the samples to treatment groups using the Thompson Sampling technique, which updates the

Figure 1. Social Network Visualization

probability distribution of assignments after observing the outcomes of the previous batch. The static design assigns treatments to each sample with equal probabilities.

The treatments yield binary outcomes: success or failure. Each treatment group has a predetermined conversion (success) rate. The number of successes of each group then follows a binomial distribution with parameters n as assigned number of observations and p as the conversion rate. After observing the treatment results of each batch, the posterior beta distribution of conversion rates is updated accordingly. The best treatment is selected from the posterior distribution after all the rounds are done.

In the first study, I vary the number of batches. The 4000 samples are equally divided into each batch. I set 3 cases of conversion probability distribution among the 9 arms: clear winner case, no clear winner case, and competing second-best case. In the clear winner case, the best arm has conversion probability of 0.2, while all other arms has conversion probability of 0.1. In the no clear winner case, the best arm has conversion probability of 0.11, while all other arms has conversion probability of 0.1. In the competing second-best case, the best arm has conversion probability of 0.2, the second best arm has conversion probability of 0.18, and all other arms has conversion probability of 0.1. I conducted simulation for all 3 cases with different number of batches. For the following studies, I only consider the clear winner case. In the second study, I use 10 batches in total, set the augmentation value to 0.2, and vary only the size of the first batch. For each experiment, a varying number of samples enter the first batch, and the rest of the samples are equally divided into the next 9 batches. In the third study, I vary the true conversion rate of the best arm while holding the conversion rate of the other 8 arms constant. The number of batches is fixed at 10 so each batch is of size 400. In the third study, I use 10 batches in total, and vary only the size of the first batch. For each experiment, a varying number of samples enter the first batch, and the rest of the samples are equally divided into the next 9 batches. In the forth study, I vary the social network spillover effect augmentation value. While the augmentation is held constant at 0.2 for all other studies, in this study I used an array of different values. The number of batches is fixed at 10 and each batch is of size 400.

For each study, the whole process is simulated 5000 times for each experiment. In each simulation, the algorithm chooses a best arm and estimates the ATE of the true best arm. I then calculate the accuracy of the algorithm choosing the correct best arm, and the root mean square error (RMSE) of the ATE of the true best arm relative to the true conversion rate. I also calculate the confidence interval of the ATE. Note that as I assume the conversion rate of no treatment is 0, ATE is the estimator of the true conversion rate.

Simulation of Spillover Effects

I use the nodes of the social network mentioned above to represent people and give them unique IDs. And in this social network, two directly connected nodes are regarded as direct friends. Everyone has the same probability of entering each round of experiments and entering each experimental group. In order to simulate the spillover effect between direct friends, before each round of the experiment, the code will determine whether there are direct friends who have participated in the previous rounds among the participants in this round, and if so, change this person's success probability into treatment probability of success plus the augmentation value.

Simulation Results

Study 1: Varying Number of Batches

In this part, I investigate if increasing the number of batches could improve the algorithm's probability of choosing the correct best arm. I use the following number of batches: 1, 2, 5, 10, 20, 50. The number of samples in each batch is the same, so for the experiment with 10 batches there would be 400 samples each batch, and for the experiment with 50 batches there would be 80 samples each batch. I compare the results of 4 different settings: Thompson Sampling or static design with or without social spillover effect. In table 1, the static design result is equivalent to the result when Thompson Sampling only has 1 batch.

Without social network spillover effect, Thompson Sampling and the static design has similar performance in all 3 cases. In the presence of a clear winner, both methods can choose the best arm with high accuracy. While in the case where there is no clear winner or in the presence of competing second best arm, Thompson Sampling method slightly outperforms the static design in its ability of choosing the best arm, but both methods have indistinguishable, high accuracy in estimating the ATE. Increasing the number of batches does not seem to improve the ability of Thompson Sampling for selecting the true best arm.

Social network spillover effect negatively impacts the ability of choosing the true best arm of both methods. There is a significant drop in Thompson Sampling's accuracy, whereas the performance of static design is more stable. The estimate of ATE from both methods becomes inflated as it suffers from the impact of the spillover effect. The more batches in each experiment, the more inflated the estimate of ATE, however the effect becomes gradually less evident as the ATE converges to the sum of the true conversion rate and the augmentation value. This could be a result of increasingly more treated samples in the former rounds incurring augmentation in more samples in the latter rounds.

The following charts are the selection accuracy of each of the 3 cases. The lines are respectively Thompson Sampling without social network impact, Thompson Sampling with social network impact, and static design with social network impact.

Figure 2. Best Arm Selection Accuracy

	Design		Probability	RMSE	Confidence interval		Mean
Assignment Algorithm	Case	Batches	Best arm selected	ATE	Low	High	ATE
TS	1: Clear winner	1	1.0000	0.0190	0.1994	0.2004	0.2000
		2	1.0000	0.0157	0.1994	0.2003	0.2000
		5	0.9980	0.0121	0.1994	0.2000	0.2000
		10	0.9980	0.0112	0.1997	0.2003	0.2000
		20	1.0000	0.0107	0.1996	0.2002	0.2000
		50	0.9994	0.0111	0.1995	0.2002	0.2000
	2: No clear winner	1	0.2704	0.0147	0.1096	0.1104	0.1100
		2	0.2648	0.0346	0.1090	0.1110	0.1100
		5	0.2902	0.0374	0.1091	0.1112	0.1100
		10	0.2978	0.0272	0.1094	0.1109	0.1100
		20	0.2956	0.0257	0.1091	0.1105	0.1100
		50	0.3084	0.0236	0.1089	0.1102	0.1100
	3: Competing second best	1	0.7652	0.0189	0.1993	0.2004	0.2000
		2	0.8612	0.0173	0.1997	0.2007	0.2000
		5	0.8976	0.0443	0.1991	0.2016	0.2000
		10	0.9092	0.0160	0.1999	0.2008	0.2000
		20	0.9054	0.0150	0.1994	0.2003	0.2000
		50	0.8996	0.0154	0.1994	0.2003	0.2000
TS with social network	1: Clear winner	1	-	-	-	-	-
		2	0.9996	0.0990	0.2973	0.2982	0.2980
		5	0.9912	0.1587	0.3533	0.3553	0.3540
		10	0.9762	0.1802	0.3693	0.3725	0.3710
		20	0.9770	0.1868	0.3773	0.3803	0.3790
		50	0.9860	0.1871	0.3824	0.3844	0.3830
	2: No clear winner	1	-	-	-	-	-
		2	0.2460	0.1168	0.2063	0.2098	0.2080
		5	0.2514	0.2621	0.2932	0.2319	0.2445
		10	0.2324	0.3130	0.2518	0.2670	0.2594
		20	0.2320	0.2489	0.2757	0.2857	0.2807
		50	0.2360	0.2047	0.2901	0.2952	0.2927
	3: Competing second best	1	-	-	-	-	-
		2	0.7792	0.1006	0.2981	0.2992	0.2986
		5	0.7814	0.3507	0.1791	0.3481	0.3534
		10	0.7830	0.2012	0.3614	0.3678	0.3646
		20	0.7900	0.2192	0.3730	0.3802	0.3766
		50	0.8016	0.1943	0.3792	0.3831	0.3811
Static with social network	1: Clear winner	1	-	-	-	-	-
		2	0.9970	0.1014	0.2974	0.2988	0.2980
		5	0.9974	0.1567	0.3535	0.3551	0.3550
		10	0.9954	0.1739	0.3708	0.3724	0.3720
		20	0.9946	0.1813	0.3782	0.3798	0.3790
		50	0.9950	0.1861	0.3831	0.3847	0.3839
	2: No clear winner	1	-	-	-	-	-
		2	0.2100	0.1816	0.2892	0.2906	0.2899
		5	0.2068	0.1565	0.2640	0.2653	0.2646
		10	0.2022	0.1731	0.2807	0.2821	0.2814
		20	0.2100	0.1816	0.2892	0.2906	0.2899
		50	0.1996	0.1854	0.2929	0.2943	0.2936
	3: Competing second best	1	-	-	-	-	-
		2	0.7432	0.1013	0.2973	0.2987	0.2980
		5	0.7416	0.1566	0.3534	0.3549	0.3541
		10	0.7356	0.3715	0.3707	0.3723	0.3715
		20	0.7354	0.1824	0.3793	0.3809	0.3801
		50	0.7356	0.1857	0.3826	0.3842	0.3834

Figure 3. Simulation Statistics, Varying Number of Batches

Study 2: Varying First Batch Size

In this part of the experiment, I fix the number of batches to 10, set the augmentation value to 0.2, and vary the size of the first batch. The size of the first batch is set to 100, 200, 500, 1000, 2000. I only use the clear winner case for this part. Without social network spillover effect, changing the size of the first batch does not have noticeable impact on the accuracy or the point estimation of ATE for either methods. I do observe a decreasing variance in the estimation with growing first batch size. Intuitively, the larger the batch size, the less likely that batch yields a unlikely situation in which the best arm produces less successes than the sub-optimal arms. Since the posterior allocation probability is sequentially based on the results of former batches, a more stable first batch is likely to lead to better performance in latter batches, and therefore smaller variance in estimating the ATE.

With social network spillover effect, the static design remains high accuracy regardless of the size of the first batch; while for Thompson Sampling, as stated before, the spillover effect has a more significant negative impact on its performance. However, the larger the first batch, the higher the accuracy of Thompson Sampling; with 2000 samples in the first batch it even slightly outperforms the static design. This could be because the first batch is not affected by the spillover effect, so the overall spillover effect on the whole population is less significant as less samples participate in the latter batches, thus more stability in the algorithm's choice.

Assignment Algorithm	Design Case	First batch size	Probability Best arm selected	RMSE ATE	Confidence interval Low	High	Expectation ATE
TS	1: Clear winner	100	1.0000	0.0177	0.1997	0.2007	0.2000
		200	0.9994	0.0157	0.1997	0.2006	0.2000
		500	0.9998	0.0102	0.1997	0.2002	0.2000
		1000	1.0000	0.0083	0.1998	0.2002	0.2000
		2000	1.0000	0.0086	0.2000	0.2004	0.2000
Static	1: Clear winner	100	1.0000	0.0229	0.1991	0.2003	0.2000
		200	1.0000	0.0217	0.1992	0.2004	0.2000
		500	1.0000	0.0210	0.1998	0.2010	0.2000
		1000	1.0000	0.0220	0.1998	0.2010	0.2000
		2000	1.0000	0.2675	0.1992	0.2006	0.2000
TS with social network	1: Clear winner	100	0.9880	0.1673	0.3633	0.3651	0.3640
		200	0.9794	0.1729	0.3655	0.3680	0.3670
		500	0.9798	0.1761	0.3670	0.3723	0.3710
		1000	0.9904	0.1766	0.3743	0.3756	0.3750
		2000	0.9984	0.1785	0.3777	0.3784	0.3780
Static with social network	1: Clear winner	100	0.9942	0.1662	0.3628	0.3644	0.3640
		200	0.9952	0.1695	0.3664	0.3680	0.3670
		500	0.9946	0.1753	0.3722	0.3738	0.3730
		1000	0.9974	0.1780	0.3744	0.3761	0.3750
		2000	0.9982	0.1831	0.3782	0.3803	0.3790

Figure 4. Simulation Statistics, Varying First Batch Size

Study 3: Varying Best Arm Conversion Rate

In this part of the experiment, I fix the conversion rate of 8 out of the 9 arms, and only vary the conversion rate of the best arm. The 8 sub-optimal arms all have a conversion rate of 0.1, and that of the other arm varies from 0.1 to 0.22, with a step size of 0.2.

Without social network effect, Thompson Sampling is more sensitive to small increases in the conversion rate of the best arm. Both methods' accuracy increases as the conversion rate of the best arm goes up. As the best arm's conversion rate reaches 0.14, the accuracy of Thompson Sampling is already above 90%, while that of the static design is only 85%.

Assignment Algorithm	Design Best arm value	Probability Best arm selected	RMSE ATE	Confidence interval Low	High	Mean ATE
TS	0.12	0.5634	0.0248	0.1192	0.1206	0.1200
	0.14	0.9336	0.0174	0.1400	0.1409	0.1400
	0.16	0.9936	0.0135	0.1598	0.1605	0.1600
	0.18	0.9988	0.0131	0.1796	0.1803	0.1800
	0.22	1.0000	0.0107	0.2195	0.2201	0.2198
Static	0.12	0.4884	0.5125	1.8995	0.4610	0.5641
	0.14	0.8492	0.8816	0.1643	0.2131	0.1800
	0.16	0.9776	0.0189	0.1592	0.1602	0.1600
	0.18	0.9976	0.0200	0.1797	0.1808	0.1800
	0.22	1.0000	0.0220	0.2192	0.2204	0.2198
TS with social network	0.12	0.4186	0.2816	0.2700	0.2830	0.2765
	0.14	0.7184	0.2565	0.2980	0.3090	0.3040
	0.16	0.8678	0.2114	0.3231	0.3303	0.3270
	0.18	0.9502	0.3477	0.1877	0.3453	0.3500
	0.22	0.9930	0.1742	0.3903	0.3921	0.3912
Static with social network	0.12	0.3262	0.1727	0.2901	0.2915	0.2910
	0.14	0.6490	0.1734	0.3107	0.3121	0.3110
	0.16	0.8710	0.1734	0.3306	0.3321	0.3313
	0.18	0.9722	0.1733	0.3503	0.3519	0.3511
	0.22	0.9994	0.1749	0.3917	0.3933	0.3925

Figure 5. Simulation Statistics, Varying Best Values

Study 4: Varying Augmentation Value

In this part of the experiment, I use the clear winner case with a fixed number of batches of 10. Under the social network effect, I tried different augmentation values: 0.01, 0.05, 0.1, 0.3, 0.5, 0.7. The two methods make very similar estimations to ATE regardless of the augmentation values. However, their performance in accuracy differs significantly.

When the social network effect is small (under 0.1) both methods have over 99% accuracy. As the spillover effect escalates, Thompson Sampling's accuracy drops dramatically, with only 84% accuracy for an augmentation value of 0.7; the static design, on the other hand, remains stable, with 99.76% accuracy for an augmentation value of 0.7.

Figure 7 shows the best arm selection accuracy with differing extent of social spillover effect using the 2 methods.

Change augmentation	Clear winner	Probability	RMSE	Confidence interval		Mean
TS with social network	augmentation	Best arm selected	ATE	Low	High	ATE
	0.01	0.9996	0.0140	0.2083	0.2089	0.2090
	0.05	0.9986	0.0447	0.2425	0.2432	0.2430
	0.1	0.9958	0.0870	0.2849	0.2859	0.2850
	0.3	0.9520	0.2891	0.4505	0.4581	0.4540
	0.5	0.8876	0.4649	0.6044	0.6165	0.6100
	0.7	0.8414	0.6265	0.7578	0.7728	0.7650
Static with social network	0.01	0.9998	0.0230	0.2083	0.2095	0.2090
	0.05	0.9998	0.0484	0.2419	0.2432	0.2430
	0.1	0.9982	0.0887	0.2845	0.2859	0.2850
	0.3	0.9928	0.2586	0.4559	0.4576	0.4570
	0.5	0.9922	0.4310	0.6286	0.6306	0.6300
	0.7	0.9976	0.6015	0.7990	0.8012	0.8000

Figure 6. Simulation Statistics, Varying Augmentation Values

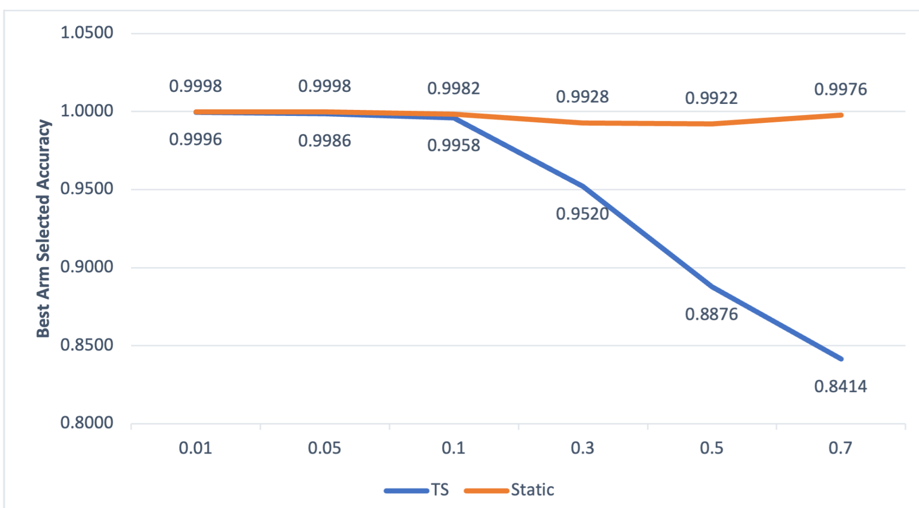

Figure 7. Best Arm Selection Accuracy, Varying Augmentation Values

Next, in order to explore why the Thompson Sampling method is more inaccurate when the social network effect is more obvious, I obtained two sets of special cases by setting two different random seeds, and drew graphs showing the trajectory of each arm judged to be the optimal one in each round of the experiment. In order to make a better comparison, in each case I took two examples where the augmentation value is equal to 0.01 and equal to 0.7, and both cases were 10-round experiments. Figure 8 shows the trajectory of the probability of each arm being judged as the optimal solution in each round of experiments under Thompson Sampling or static method when augmentation value is set to 0.01 and 0.7 in Case 1. In all the four panels in Case 1, the true best arm obtained the highest probability of being the optimal solution in the first round. In the following rounds, although there was an occasional decrease in the static method case, the probability that the true best arm being the optimal solution showed an overall upward trend in all the four panels. In addition, it can also be observed that when the augmentation value is the same, TS method converges faster than the static method. The TS method converged at most 4 rounds, but the static method needed at least 7 rounds to converge. With augmentation value equal to 0.7, TS converged faster than with augmentation value equal to 0.01. It took only 2 rounds for the TS method to converge when augmentation value is equal to 0.7.

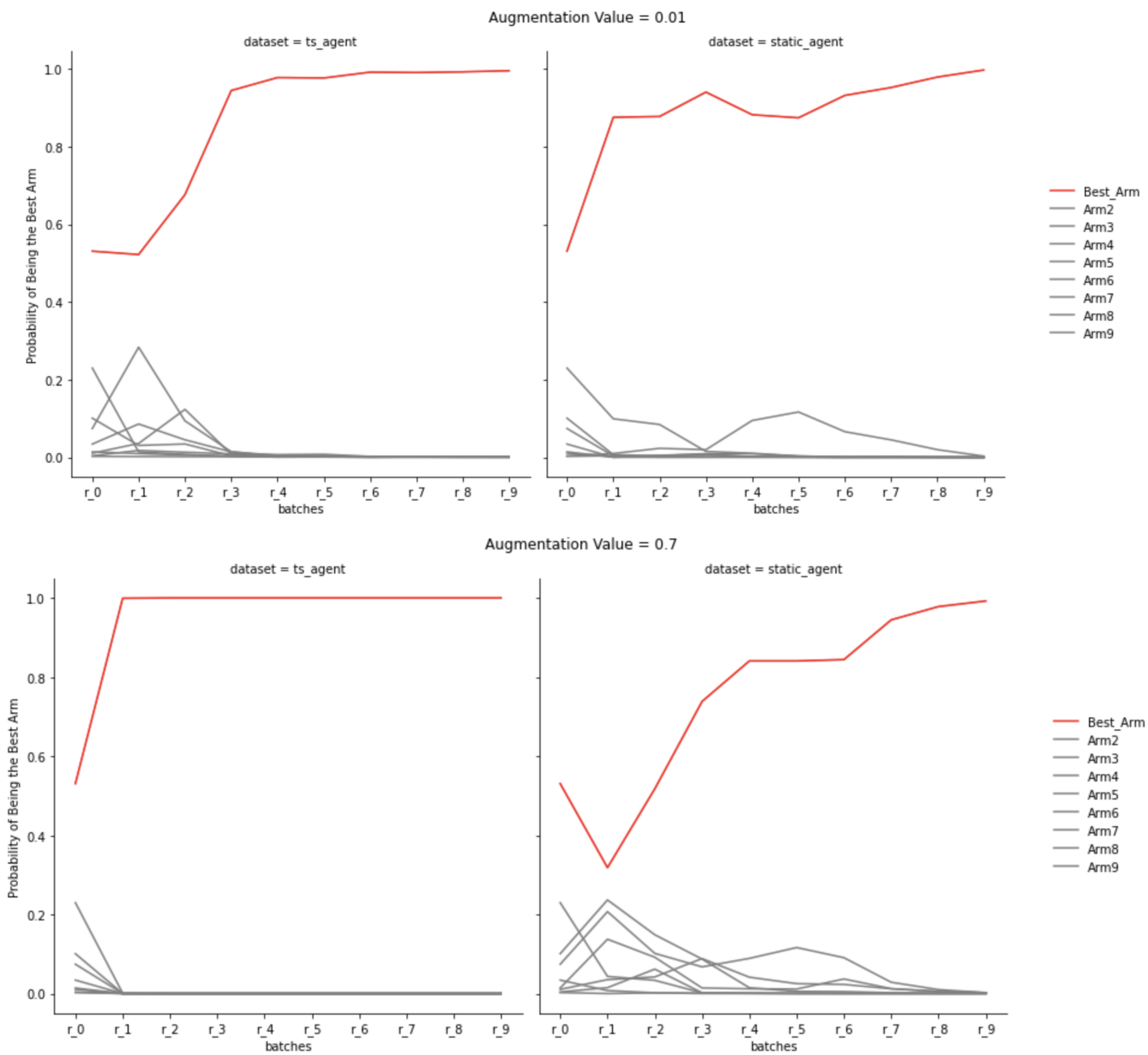

Figure 8. Case 1: True best arm is judged as the optimal one in the first round.

Although in Case 1, the TS method seems to perform better when the augmentation value is larger, in Case 2, when the true best arm is not considered to be the optimal solution in the first round, the performance of Thompson

Sampling is not satisfactory. Figure 9 is a 4-panel graph showing the trajectory of the probability of each arm being judged as the optimal solution in each round of experiments under Thompson Sampling or static method under two different augmentation values in Case 2. It can be seen from the figure that after the first round of experiments, in all 4 panels, the true best arm has lower probability of being the optimal solution than other arms. But in the situation when augmentation value is equal to 0.01, both TS and static method corrected the error and correctly identified the true best arm in the next few rounds of experiments. Still, Thompson Sampling converged faster than static method. However, in the case of augmentation value equal to 0.7, the results of TS and static method are very different. TS converged very fast and kept getting wrong, while after 2 rounds, static method continued to increase the possibility that the true best arm is the optimal solution, and finally recognized it correctly.

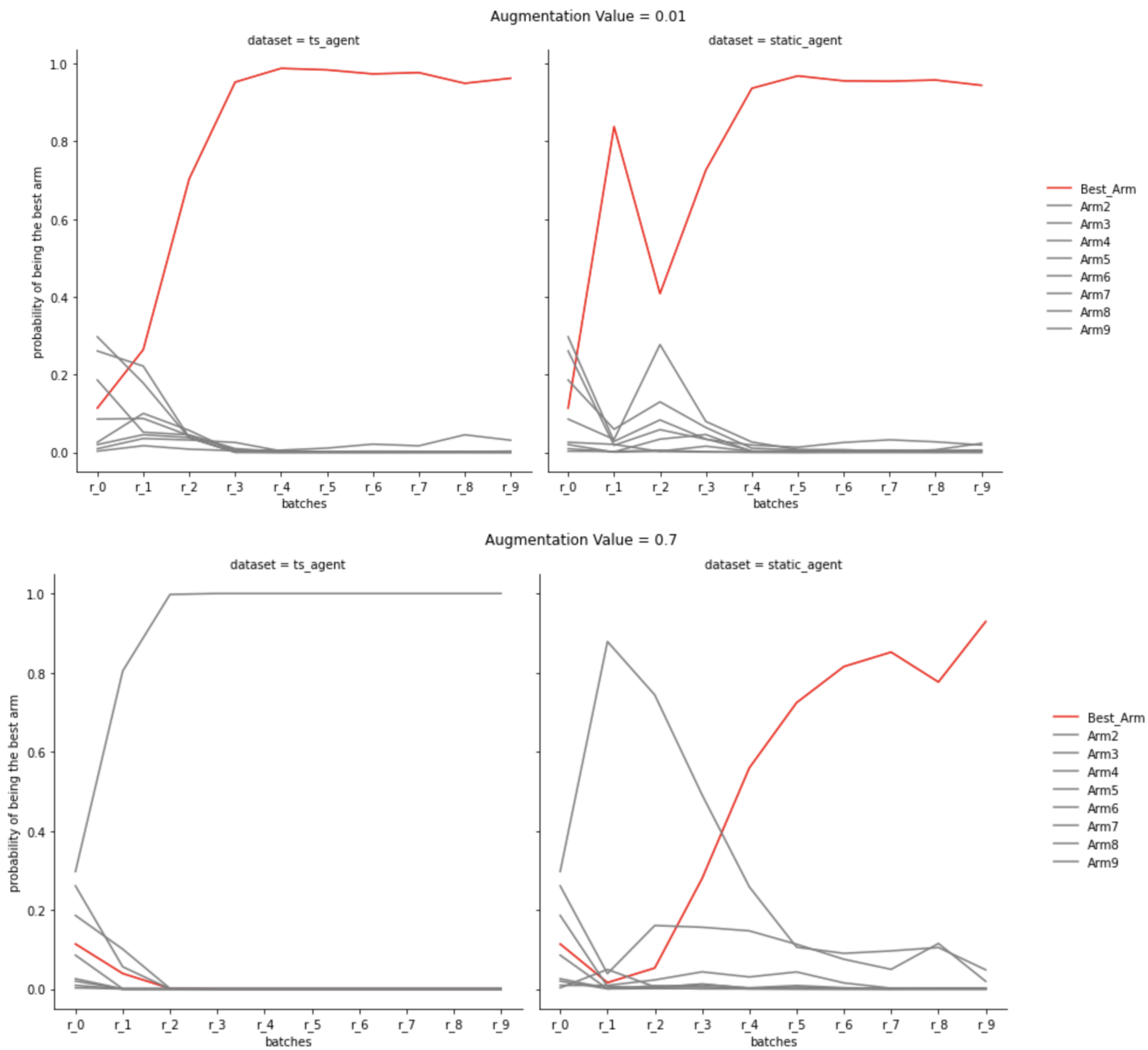

Figure 9. Case 2: True best arm is not judged as the optimal one in the first round.

Through the above two cases, we can find that when the augmentation value is large, that is, the social network effect is large, the results of the first round of experiments have a great impact on the accuracy of the TS method. It is difficult for TS method to correct initial false judgment in subsequent rounds of experiments. In order to further explore the impact of the results of the first round of experiments on the TS method, I studied another example in which two arms were considered equally optimal in the first round when augmentation value equal to 0.7, the results are shown in Figure 10. It can be seen from the figure that the TS method ends the exploration of other arms after the second round, and it only conducts follow-up exploration on the two arms that are judged to be equally optimal in the first round. But the static method continues to explore in subsequent experiments. This is also because it allocates the same number of subjects to each arm in each round, but TS allocates more subjects to the arm that was judged to be optimal in the previous round.

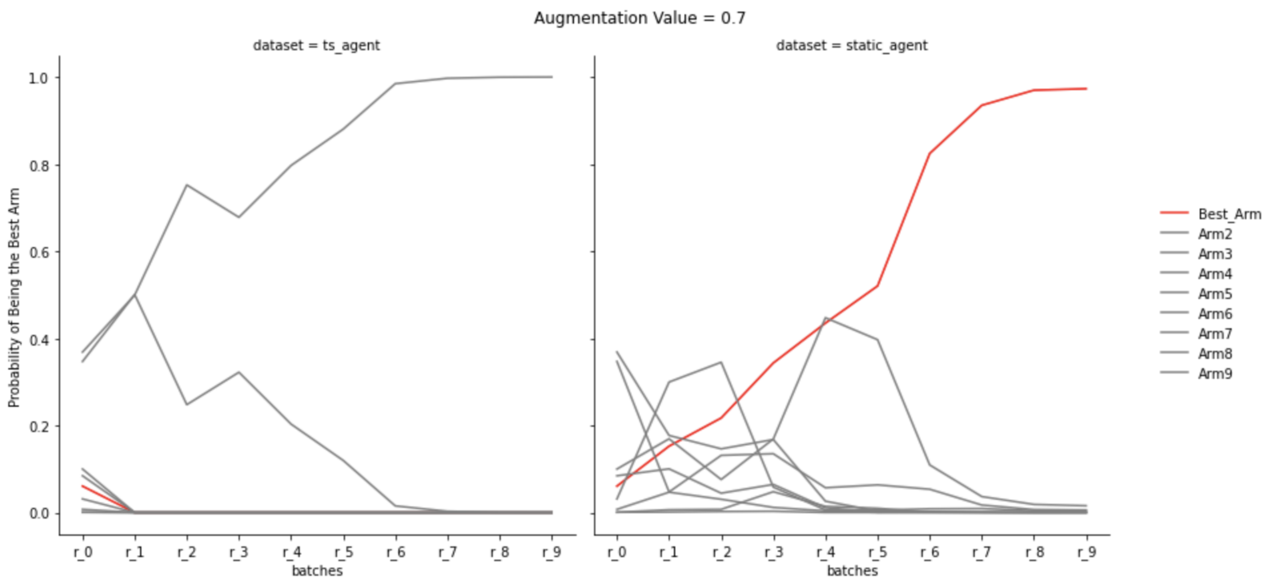

Figure 10. Case 3: Two arms are judged equally optimal in the first round.

From the above cases, I can conclude that when the augmentation value is large, the final conclusion of TS is greatly affected by the results of the first round of experiments. The arm that was judged to be the optimal solution in the first round is also more likely to be insisted on as the optimal solution in subsequent rounds of experiments. In other words, if one arm is not judged as one of the optimal solutions in the first round, then it will never be judged as the optimal solution in the subsequent rounds. What's more, I also find that when the augmentation value is small, the exploration phase of Thompson Sampling will be longer. With the augmentation value equal to 0.01, the TS method needs four rounds to converge, which means it is more exploring in the first 4 rounds and more capable of correcting mistakes. However, when the augmentation value is large, the TS method converges very fast, which means that its exploration phase is very short while the exploitation phase is very long. If it cannot draw correct conclusions in the shorter exploration phase, then the subsequent exploitation phase will continue the previous wrong conclusions.

Increasing First Batch Size

According to the above findings, the TS method over-relies on early stage results. In this section, I further explore whether increasing the first batch size will fix this problem. I still use 4000 samples and 9 arms, fix the number of batches to 10, set the augmentation value to 0.7, and vary the size of the first batch. The size of the first batch is set to 400, 500, 1000, 1500, 1600, 1700, 1800, 1900 and 2000. I only use the clear winner case for this part and run 5000 replications for each size to get the accuracy. The results are shown in Figure 11. We can find that as the first batch size increases, the performance of TS is indeed getting better. Compared with the benchmark example where the first batch size is equal to 400 and the TS accuracy rate is only 0.8414, after increasing the first batch size to 1000, the performance of TS is indeed improved, and there is an accuracy rate of 0.9494, but the performance of TS is still worse than that of static design. What's more, as shown in the results in Study 2, when the first batch size equals 1000, the number of batches is 10, and augmentation equals 0.2, the accuracy of TS is 0.9904. Through these comparative studies, we can conclude that choosing a relatively larger first batch size can partially offset the impact of augmentation value, but cannot completely eliminate it. If we keep increasing the first batch size, the accuracy of TS will eventually exceed 0.99. When the first batch size reaches at least 1700, the accuracy rate of TS will exceed 0.99, but static method can always maintain an accuracy rate of 0.99. What's more, the charm of Thompson Sampling lies in the balance between exploration and exploitation, but if we need to invest a large number of subjects in the first stage, that is, the exploration stage, to pursue the accuracy of TS, then the benefits of using Thompson Sampling will be greatly reduced.

Augmentation Value = 0.7 Assignment Algorithm	First Batch Size	Probability Best arm selected
TS	400	0.8414
	500	0.858
	1000	0.9494
	1500	0.9782
	1600	0.985
	1700	0.9902
	1800	0.9918
	1900	0.993
	2000	0.996
Static	400	0.9976
	500	0.9968
	1000	0.9932
	1500	0.992
	1600	0.9908
	1700	0.991
	1800	0.9926
	1900	0.9914
	2000	0.9906

Figure 11. Simulation Statistics, Increasing First Batch Size

Conclusion

In online advertising campaigns, before mass distributing, advertisers display various versions of advertisements on social medias to do experiments and select the optimal one. The thought of exploiting while exploring in adaptive design method attracts many practitioners. If units are totally independent, Thompson Sampling usually outperforms static design, however, under the social network spillover effects, Thompson Sampling algorithm becomes unstable.

In this article, I try to explore the performance of Thompson Sampling and static design methods under the appearance of social network spillover effect and get some absorbing findings. From simulation results, I conclude that social network spillover effect has evident effect on both Thompson Sampling and the static design. Their estimation of ATE is inflated under this effect, and the more batches the more evident the inflation. In terms of accuracy, social network spillover effect has a much more remarkable impact on Thompson Sampling than static design. With social network spillover effect, Thompson Sampling's accuracy increases with larger first batch size. The higher the augmentation value, the less accurate Thompson Sampling is. Thompson Sampling is sensitive to varying best value, as a small increase in the best value could greatly boost the performance of Thompson Sampling.

In the clear winner case, Thompson Sampling is more negatively impacted than the static design, while its accuracy holds higher than the static design in cases where there is no clear winner or another competing second best arm.

Limitation and Future Work

In this paper, through the simulations which compare multiple groups of experiments, it is found that both static design and TS method yield biased estimation. Under the appearance of social network spillover effects, the accuracy of Thompson Sampling method selecting the true best arm is less than that of static design and the performance of Thompson Sampling is also worse. There are two main limitations in this article and I can take further steps in the future. The first is that I lack the real experiments data. I can only do simulations in current stage. In the future, I can design field experiments and collect real world data to test my findings. The second one is that I can further figure out an unbiased estimator for Thompson Sampling and static design under the existence of social network spillover effects.

References

1. L Busca, L. B. A framework for digital marketing research: investigating the four cultural eras of digital marketing. *J. Interact. Mark.* **49**, 1–19 (2020).
2. Ko, E. Bridging asia and the world: global platform for the interface between marketing and management. *J. Bus. Res* **99**, 350–353 (2019).
3. Mandal, P. Understanding digital marketing strategy. *Int. J. Sci. Res. Manag.* (2017).
4. J R Hauser, G. L. U., G Liberali. Website morphing 2.0: Technical and implementation advances and a field experiment. *Manag. Sci* **60**, 1594–1616 (2014).
5. M Faruk, S. H., M Rahman. How digital marketing evolved over time: A bibliometric analysis on scopus database. *Heliyon* **7** (2021).
6. Eric M. Schwartz, P. S. F., Eric T. Bradlow. Customer acquisition via display advertising using multi-armed bandit experiments. *Mark. Sci.* **36**, 500–522, DOI: 10.1287/mksc.2016.1023 (2017).
7. A Lambrecht, C. T. When does retargeting work? information specificity in online advertising. *J. Mark. Res* **50**, 561–576 (2013).
8. A Goldfarb, C. T. Online display advertising: Targeting and obtrusiveness. *Mark. Sci* **30**, 389–404 (2011).
9. Robbins, H. Some aspects of the sequential design of experiments. *Bull. Amer. Math. Soc* **58**, 527–535 (1952).
10. G Keller, A. O. Branching bandits: A sequential search process with correlated pay-offs. *J. Econ. Theory* **113**, 302–315 (2003).
11. B C May, A. L. D. S. L., N Korda. *Optimist. Bayesian sampling contextual bandit problems* (2011).
12. Scott, S. A modern bayesian look at the multi-armed bandit. *Appl. Stoch. Model. Bus. Ind.* **26**, 639–658, DOI: 10.1002/asmb.874 (2010).
13. P M Aronow, C. S. Estimating average causal effects under general interference, with application to a social network experiment. *The Annals Appl. Stat.* **11**, 1912–1947 (2017).
14. I Chae, Y. B. D. Y., A T Stephen. Spillover effects in seeded word-of-mouth marketing campaigns. *Mark. Sci.* **36**, 89–104 (2017).
15. Thompson, W. R. On the likelihood that one unknown probability exceeds another in view of the evidence of two samples. *Biometrika* **25**, 285–294 (1933).
16. M Offer-Westort, D. P. G., A Coppock. Adaptive experimental design: Prospects and applications in political science. *Am. J. Polit. Sci.* **65**, 826–844 (2021).
17. Leskovec, J. & Mcauley, J. Learning to discover social circles in ego networks. In Pereira, F., Burges, C., Bottou, L. & Weinberger, K. (eds.) *Advances in Neural Information Processing Systems*, vol. 25 (Curran Associates, Inc., 2012).
18. A Sallaberry, G. M., F Zaidi. Model for generating artificial social networks having community structures with small-world and scale-free properties. *Soc. Netw. Analysis Min.* **3**, 597–609 (2013).
19. L Hamill, G. G. Social circles: A simple structure for agent-based social network models. *J. Artif. Soc. Soc. Simul.* **12** (2009).
20. D Hales, S. A. *Friends for free: Self-organizing artificial social networks for trust cooperation* (2005).
21. S Elkosantini, D. G. Human behavior and social network simulation: fuzzy sets/logic and agents-based approach. *SpringSim* 102–109 (2007).
22. S S Villar, J. B., J Wason. Response-adaptive randomization for multi-arm clinical trials using the forward looking gittins index rule. *Biometrics* **71**, 969–978 (2015).

23. R A Lewis, D. N. Display advertising's competitivespillovers to consumer search. *Quant. Mark. Econ.* **13**, 93–115 (2015).

24. A Borah, G. J. T. *Halo (spillover) effects social media: Doproduct recalls one brand hurt or help rival brands?J. Mar-keting Res* **53**, 143–160 (2016).

25. Y Trope, N. L. Construal-level theory of psychologicaldistance. *Psych. Rev* **117**, 440–463 (2010).

26. A T Stephen, D. L. How word-of-mouth transmis-sion encouragement affects consumers' transmission decisions,receiver selection, and diffusion speed. *Internat. J. Res. Mark.* (2016).

27. Anderson, E. W. The formation of market-level expectationsand its covariates. *J. Consumer Res* **30**, 115–124 (2003).