The Reach and Persuasiveness of Viral Video Ads

Citation: Tucker, Catherine E. "The Reach and Persuasiveness of Viral Video Ads." Marketing Science 34, no. 2 (March 2015): 281–296.

As Published: http://dx.doi.org/10.1287/mksc.2014.0874

Publisher: Institute for Operations Research and the Management Sciences (INFORMS)

Persistent URL: http://hdl.handle.net/1721.1/99172

Version: Author's final manuscript: final author's manuscript post peer review, without publisher's formatting or copy editing

The Reach and Persuasiveness of Viral Video Ads

Catherine Tucker*

July 8, 2014

Abstract

Many video ads are designed to go 'viral,' so that the total number of views they receive depends on customers sharing the ads with their friends. This paper explores the relationship between number of views and how persuasive the ad is at convincing consumers to purchase or to adopt a favorable attitude towards the product. The analysis combines data on the total views of 400 video ads, and crowd-sourced measurement of advertising persuasiveness among 24,000 survey responses. Persuasiveness is measured by randomly exposing half of these consumers to a video ad and half to a similar placebo video ad, and then surveying their attitudes towards the focal product. Relative ad persuasiveness is on average 10% lower for every one million views that the video ad achieves. The exceptions to this pattern were ads that generated both views *and* large numbers of comments, and video ads that attracted comments that mentioned the product by name. There is suggestive evidence that such ads remained effective because they attracted views due to humor rather than because they were outrageous.

Keywords: Viral Advertising, Virality, Video Advertising, Internet
JEL Codes: L86, M37

*MIT Sloan School of Management, MIT, Cambridge, MA. and NBER. Thank-you to Ed Roberts for inspiring the project, Visible Measures for providing the data, and the NSF (CAREER Award. 6923256) and the Net Institute (http.www.NETinst.org) for providing financial support. Thank-you to Rene Algesheimer, Yakov Bart, Oliver Emrich, Gui Liberali and seminar participants at the University of Zurich for useful comments. All mistakes are mine alone.

1 Introduction

In the past few years, digital marketing strategy has shifted away from an emphasis on 'paid' media, where a brand pays to advertise, to 'earned' media, where the customers themselves become the channel of delivery (Corcoran, 2009). Reflecting this shift, social video advertising is among the fastest-growing segments in advertising today. In 2010, social video advertising views increased 230%, over nine times the growth in online search and display advertising (Olenski, 2010). These video ads are crucially different from rich-media banner ads. Rather than the advertiser paying for placement, these ads are designed to be transmitted by consumers themselves, either through consumers posting them on their social media feeds or sharing them directly with friends. This means that firms commission these video ads and post them on websites such as `YouTube.com`, in the hope and expectation that consumers themselves will encourage others to watch the video. This is evidently attractive for firms, as it implies a costless means of transmitting advertising. However, in common with other forms of 'earned' media, the return on investment from views obtained in this manner is not clear (Miller and Tucker, 2013).

This paper seeks to understand what the relationship is between the 'earning' of media and the persuasiveness of the media. The direction of the relationship is not clear. On the one hand, the very act of sharing a video ad suggests a degree of investment in the product and a liking of the ad that may speak well to its persuasiveness. On the other hand, advertisers may have to sacrifice elements of ad design in order to encourage people to share the ad, and that may damage the ad's persuasiveness.

The analysis uses historical data on the number of times that 400 different video ad campaigns posted on `YouTube.com` during 2010 were viewed. This data comes from a media metrics company that tracks major advertiser video ads and records the number of times these ads are viewed. The persuasiveness of these campaigns is then measured using tech-

niques pioneered by media metrics agencies such as Dynamic Logic and previously used in data analysis by Goldfarb and Tucker (2011a). I obtained 25,000 survey-responses through crowdsourcing and measured the effect of exposure to the video ad on purchase intent, using a randomized treatment and control methodology for each campaign. Respondents are either exposed to a focal product video or to a placebo video ad of similar length for another product in the data. They are then asked questions about their purchase intent and brand attitudes towards the focal product.

The randomization induced by the field-test procedure means that econometric analysis is straightforward. First, the analysis documents the direction of the relationship between the number of times an ad was viewed and traditional measures of advertising persuasiveness. Ads that achieved more views were less successful at increasing purchase intent. This is robust to different functional forms and to alternative definitions of the explanatory and dependent variable, such as brand favorability and consideration. It is robust to controls that allow the effect of exposure to vary by video ad length, campaign length, respondent demographics, product awareness and category. It is also robust to excluding respondents who had seen or heard of the ad before, meaning that the results do not reflect satiation.

Estimates of the magnitude of this negative relationship suggest that on average, ads that have received one million more views are 10% less persuasive. This drop in persuasiveness was compensated for by the increased number of views, so the paper also presents some rough projections to determine the point at which decreased persuasiveness outweighs the increased number of views in terms of the total persuasion exerted over the population. The estimates suggest that this point occurs between three and four million views, a viewership achieved by 6% of campaigns in the data.

The 'total views' reach measure, though, is only the static endpoint of a viral process. Therefore, the analysis also demonstrates that the results hold when looking at other more dynamic measures of virality, such as the pattern of the time trend of views and the views

that can be attributed to non-advertiser-seeded placement of video ads.

The crucial managerial question, though, is whether there are identifiable categories of ads for whom this negative relationship between organic reach and persuasiveness did not exist. Such cases are clearly very attractive to advertising managers, as they imply that organic reach does not have to be costly in terms of the persuasiveness of the ad design. Strikingly, the exceptions to this tradeoff are ads that either attract a lot of comments or that attract comments that mention the product by name. This has an important managerial implication. Marketing managers, as well as tracking total views for their ads, should also track the creation of user-generated content surrounding the ads. This should be used as an early indicator of the ads' likely ability to be persuasive as well as achieving high reach.

It is of course important to understand which underlying ad characteristics explain the results. The ads that did *not* exhibit this negative relationship between total views and persuasiveness were also less likely to be rated as being outrageous by participants. Instead, they were more likely to be rated as funny or, more weakly, as visually appealing. This is in line with an older advertising research literature that has emphasized that likeability (such as produced by humor) is an important determinant of ad appeal (Biel and Bridgwater, 1990; Weinberger and Gulas, 1992; Vakratsas and Ambler, 1999; Eisend, 2009), and that intentional outrageousness is less likely to be effective (Barnes and Dotson, 1990; Vzina and Paul, 1997). Therefore, an explanation for the results is that videos are achieving high reach because they are intentionally outrageous, and, as such, command attention (Tellis, 2004), but that an ad design of this kind ultimately harms the ad's persuasiveness.

This paper contributes to three existing academic literatures.

The first literature is on virality. Aral and Walker (2011) use data from a field experiment for an application on Facebook to show that forcing product users to broadcast a message is more effective than allowing users to post more personalized recommendations at their discretion. There have also been a few studies of campaigns that were explicitly designed

to go 'viral.' Toubia et al. (2011) presents evidence that a couponing campaign was more effective when transmitted using a 'viral' strategy on social media than when using more traditional offline methods. Chen et al. (2011) has shown that such social influence is most important at the beginning of a product's life.

Some recent papers have modeled the determinants of whether or not a video ad campaign goes 'viral.' This is increasingly important, given that 71% of online adults now use video-sharing sites (Moore, 2011). Porter and Golan (2006) emphasize the importance of outrageous content (specifically, sexuality, humor, violence, and nudity) as a determinant of virality; Brown et al. (2010) echo the importance of comedic violence and argue that the outrageous nature of these ads appears to be a key driver. Eckler and Bolls (2011) emphasize the importance of a positive emotional tone for virality. Outside of the video-ad sphere, Chiu et al. (2007) emphasized that hedonic messages are more likely to be shared by e-mail; Berger and Milkman (2012) emphasize that online news content is more likely to be shared if it evokes high or negative arousal as opposed to deactivating emotions such as sadness. Elberse et al. (2011) examined 12 months of data on popular trailers for movies and video games. They found evidence that the trailers' popularity was often driven by their daily advertising budget. Teixeira (2011) examines what drives people to share videos online and distinguishes between social utility and content utility in non-altruistic sharing behavior. Though these papers provide important empirical evidence about the drivers of virality, these papers did not actually measure how persuasive the video ads were and how this related to virality.

The second literature is on the persuasiveness of online advertising. Much of this literature has not considered advertising that is designed to be shared, instead focusing on non-interactive banner campaigns (Manchanda et al., 2006; Lambrecht and Tucker, 2013). Generally, this literature has only considered the persuasiveness of video-advertising tangentially or as part of a larger study. For example, Goldfarb and Tucker (2011a) presented a

result that video advertising is less persuasive when placed in a context which matched too closely the product being advertised. In the arena of video advertising, Teixeira et al. (2012) have shown that video ads that elicit joy or surprise are more likely to retain visual focus (as measured by eye-tracking) and are less likely to be fast-forwarded through. This is the first study, however, on the relationship between ad virality and ad persuasiveness, that is, how the ability of an ad to endogenously gain 'reach' is related to the ability of the ad to persuade.

The final literature is a more recent one exploring the tradeoff between ad design and attention or reach, suggesting that the internet has reduced the tradeoff between richness and reach in information delivery in the internet era. Before the commercialization of the internet, firms had to choose between personal selling, which is an incredibly rich form of marketing communications but which has limited reach since there are no economies of scale, and media like television advertising, which achieves impressive reach but is not a rich form of marketing communications. Evans and Wurster (2000) argue that the easy replication and personalization facilitated by the internet reduced this tradeoff. This paper suggests, however, that advertisers who try to achieve scale on the internet through the actions of internet users rather than their own efforts may still face tradeoffs in terms of the persuasiveness of ads that users can be persuaded to view and subsequently share. This finding is more in line with the older literature on advertising content, which suggests that there is a substantial tradeoff between achieving persuasion and attention by the use of emotive ad characteristics (For a summary, see Tellis (2004), p. 151). For example, Steadman (1969) shows that the sexiness of advertising, though good at commanding attention, negatively affects brand recall. This echoes these results on the effects of outrageousness as a characteristic of viral ads. Though outrageousness is effective at increasing total views, ads that are outrageous are less effective at positively persuading consumers to buy the product.

2 Data

Visible Measures, a large video metrics company, provided the data for this paper. Data for movie campaigns provided by this company has also been used by Elberse et al. (2011) to study the effects of direct advertiser actions on video virality for movie previews. Visible Measures, founded in 2005, is an independent third-party media measurement firm for online video advertisers and publishers. It is the market leader in terms of tracking views and engagement for different types of social video ads. Visible Measures shared data for 2010 campaigns in the consumer goods category. They excluded from the data video ads for categories of products such as cars and other expensive items, for which most people were unlikely to be in the market. They also excluded video ads for entertainment products such as movies, video games, and DVDs, whose ads have a short shelf life.

29 percent of videos were for consumer packaged goods, 14 percent of videos were for electronics, 13 percent of videos were for apparel and eight percent were for fast food. The highest priced items was air travel (around three percent of campaigns). The lowest priced items were sodas and snacks in the consumer packaged goods category. Persuasiveness is allowed to vary by these different 'product' categories as controls in subsequent robustness checks.

The videos of 396 of these campaigns were still live, as measured by whether they were active on the video-sharing website `YouTube.com`, and consequently were included in this survey. These 396 videos covered 271 brands and 278 different products. All of these products had been advertised elsewhere, though in three percent of cases (all in the electronics category) the ad was for a new product release.Since Visible Measures is primarily employed as a media measurement company, it does not have data on the design costs or the creative process that lay behind the ads it tracks. Though Visible Measures did not share its proprietary system for collecting this data, descriptions on its website suggests that the data

is scraped on a daily basis from major video-sharing websites such as `Youtube.com` and `Vimeo.com`.

Table 1a reports the campaign-level summary statistics received from Visible Measures. 'Total views' captures the number of times these videos had been viewed by consumers. This encompasses both the views of the original video as placed by the ad agency, and views that were generated by copies of the ad and derivatives of the ad. It is clear from the standard deviation that there is a high variance in the number of total views across the ad campaigns, which is one of the reasons the for using log measures in the regressions. The results are also robust to a raw linear measure.

The analysis takes 'total views', the number of times in total the ad was viewed, is the initial static proxy measure of the outcome of the viral process.Such measures of reach are often loosely referred to as measuring 'virality of ads.'[1] This reflects a view that views of social video ads on pages such as `YouTube.com` are gained by an organic process where people find such videos on blogs or social media sites and then share the video ad further with their friends. However, since this process could be subject to manipulation[2] by advertisers, the paper presents alternative specifications that attempt to isolate only the views achieved from non-advertiser seeded placements and also measures of virality which reflect more closely the idea of a dynamic process. 'Total Comments' records the number of times that these videos had received a written comment from a consumer, typically posted below the ad on websites such as `Youtube.com`.

Of course, a simple regression that correlated firms' sales and the virality of their ad campaigns is unlikely to be useful, since the decision to launch a viral ad campaign is confounded with many other factors. Direct measurement of consumer response rates for online video ads is also difficult. Though it is possible to measure whether or not a `Youtube.com`

[1]See for example `https://www.facebook.com/help/285625061456389` where organic reach totals are referred to as 'viral reach'.

[2]See Wilbur and Zhu (2009) for a general discussion of manipulation of online ads.

user subscribes to a channel, not all users maintain accounts of the kind that allows them to subscribe. Typical 'direct response' methods of evaluating digital advertising, such as measuring click-throughs, are not appropriate. Many videos do not have embedded hyperlinks, and also many products that are advertised in the videos such as deodorant are not primarily sold online. As documented by Porter and Golan (2006) and Golan and Zaidner (2008), viral advertising very rarely has a clear 'call to action', such as visiting a website, that is measurable. Therefore, the analysis's advertising persuasiveness is based on industry standard techniques for measuring the persuasiveness of brand campaigns online. These techniques, developed by among others Dynamic Logic and Insight Express, combine a randomized control and exposure methodology with surveys on brand attitudes. Both major advertisers and major agencies use these same techniques for evaluating both banner campaigns and video campaigns.

Since such ad persuasiveness measures were not used as the campaigns were being rolled out, this data had to be collected retrospectively. Given the number of campaigns in the source data, this requires a large number of participants. To accomplish this, 25,000 survey responses were collected using the crowdsourcing platform Mechanical Turk. Similar crowdsourcing techniques have been used by Ghose et al. (2012) to design rankings for search results. Each of these participants visited a website that had been designed to resemble popular video sharing websites such as `Youtube.com`. The main difference between the study website and a traditional video-sharing website is that participants had no choice but to watch the video and that after watching the video, participants were asked to answer a series of questions concerning their brand attitudes. The other difference is that, as the video was embedded in the survey, survey takers were not exposed to the prior number of views or the number or nature of the comments that the video had received, as they would have been exposed to if they had been viewing the video on a regular video-sharing website.

For each campaign, on average 60 respondents were recruited. Half of the respondents

are allocated to a condition where they are exposed to the focal video ad for which there is virality data. The other half of respondents (the control group) see a placebo video ad for another unrelated (random) product that was also part of the data. The placebo ad that was shown among the control group was randomized to make sure that the choice of placebo ad did not influence the result.[3]

The randomization between whether someone saw the focal video ad or another one, means that in expectation all the respondents are identical. Therefore, the analysis can causally attribute any differences in their subsequent attitudes towards the product to whether they were exposed to the video ad or not.

The data recorded whether or not the respondent watches the video all the way through and the analysis excludes those who did not. Also excluded were participants who, despite the controls in place, managed to take the survey multiple times.[4]

There is obviously the potential for some multiple survey takers to have slipped through the cracks - for example if they masked their IP or used multiple IP addresses. However, at least the steps we did take should have weeded out the majority of multiple survey takers. This explains why there are 24,367 responses, which is fewer than the original 25,000 survey-responses recruited.[5] Table 1b summarizes the responses to the subsequent survey questions. These include questions about their purchase intent towards the focal product and likelihood of consideration of the focal product. There were also decoy questions about

[3]This could have occurred if the advertising was directly combative (Chen et al., 2009)

[4]A natural concern on Mechanical Turk is that not all survey takers are unique. We addressed this concern in three ways in our data processing:

1. We set up Mechanical Turk so that each Mechanical Turker was asked to complete a single survey.
2. We checked that no answers had identical Mechanical Turker IDs. We dropped subsequent responses from Mechanical Turkers who had responded multiple times to the survey.
3. Of course, one remaining problem is that Mechanical Turkers could potentially have multiple accounts. We tracked IP addresses to try to weed out duplicates. We dropped duplicates where identified.

[5]Also excluded were 161 survey-takers who incorrectly gave the length of the video as a veracity check for paying attention to the video.

another brand. All these questions are asked on a five-point scale in line with traditional advertising persuasiveness questioning (Morwitz et al., 2007). To allow comparison with the estimates of Goldfarb and Tucker (2011a) who use a similar methodology but in a non-forced exposure setting, this variable is converted from a five-point scale to a binary purchase intent measure that captures whether someone is very likely or likely to purchase the product for the main analysis. However, this is shown to be robust to the full scale in subsequent regressions. As seen in Table 1b, average purchase intent was relatively high, reflecting the mainstream nature of the products in the ads.

The use on purchase intent as the key dependent measure, means that the analysis is focused on the effect of advertising at the later stages of the purchase funnel or traditional purchase decision process (Vakratsas and Ambler, 1999). The paper's methodological approach, which necessitates forced exposure, makes it hard to analyze 'awareness' or other earlier stages of customer attitudes earlier in the purchase process.[6] The analysis does, however, control for heterogeneity in product awareness in subsequent regressions.

Survey responses are weaker measures of advertising persuasiveness than purchasing or profitability (as used by Reiley and Lewis (2009)), because although users may say they will purchase, they ultimately may not actually do so. However, as long as there is a positive correlation between whether someone intends to purchase a product and whether they actually do so, the directionality of the results should hold. Such a positive correlation between stated purchase intent and purchase outcomes has been broadly established (Bemmaor, 1995; Morwitz et al., 2007). However, a conservative view would be that the results reflect how total views is related to an established and widely-used measure of advertising persuasiveness that is used as an input when making advertising allocation decisions.

In addition to asking about purchase intent, the survey also asked participants about

[6]The data does not contain time-stamps for when consumers completed different parts of the survey, but in general the time they took to complete the task was only minutes more than it took to watch the video. This lack of interruption to survey taking prevented the collection of measures on ad memorability.

whether or not they recalled having seen the focal video ad before or had heard it discussed by their friends and media. This information is used in a robustness check to make sure that the fact that respondents are more likely to have seen viral videos before is not driving the results.

The survey also asked respondents about their gender, income, age, and the number of hours they spent on the internet. These descriptives are reported in Table 1b. They are used as controls in the regression, though since respondent allocation to exposed and control group was random, they mainly serve to improve efficiency. However, they do serve also as a check on how representative the survey-takers were. It is clear that respondents are more male than the general population, are younger, earn less, and also spend more time online. However, it is still possible they reflect the general population of viewers of video-sharing sites.

70% of survey-takers were male. This is similar to statistics reported by Moore (2011), that men are 28 percent more likely than women to have used a video-sharing site recently. The survey-takers were on average 30 years old. ComScore, a website audience tracking service, reports in its Searchplanner tool that in September 2010, `YouTube.com` users were on average 31.6 years old, which is reasonably similar to the Mechanical Turk population. In the comScore `YouTube.com` user data, 41 percent of users had an income under $40,000 a year, which is lower than this survey where 62.2 percent had a comparably low income. This suggests that the income level is different for the `Youtube.com` population and the Mechanical Turk population. However, work by Goldfarb and Prince (2008) suggests that this simple comparison of unique visitor income may overstate the difference, as it fails to adjust for time spent on the website. Their research instead suggests that poorer `YouTube.com` users are likely to account disproportionately for the time spent on `YouTube.com`. However, in general that since these participants were recruited via a crowdsourcing website, there also is the possibility that they may differ in unobserved ways from the population.

Another caveat with the representativeness of the responses is that the data collection takes place in a forced exposure setting which does not mimic the social process.[7] Therefore, 'persuasiveness' should be most properly thought of as the direct persuasiveness that can only be attributed to exposure to the video between the treated and exposed conditions and that does not reflect any incremental lift beyond that found in the video content that might come from a friend's recommendation.

The issue of how representative such respondents' answers are is faced by all research using survey-based evaluation techniques, as discussed in Goldfarb and Tucker (2011b). However, what is crucial is that there is no *a priori* reason to think that the kinds of ads that these participants would be favorably impressed by would differ from the more general video-sharing population, even if the magnitudes of their responses may differ. There is also evidence that the magnitudes of the measured effects match existing estimates of video-advertising efficacy that have been collected in less artificial settings (Goldfarb and Tucker, 2011a).

In addition, participants rated the videos on a ten-point sliding scale based on the extent to which they found it humorous, visually appealing, or outrageous.[8] Table 1c reports these ratings at the campaign level, based on the median response of the survey-takers.

[7]Supplementary analysis, also collected information on persuasiveness of videos for a randomly selected subset of 30 of the original 400 videos with slightly different instructions. In this slightly altered scenario, rather than been instructed to watch a video ad, survey takers were told to imagine they had just been sent the link to the `YouTube.com` video by a friend with the instruction to 'Check it out'. We compared the relative persuasiveness of this setting with our original setting, and found no statistically significant difference for purchase intent - (t= 0.47, p-value=0.63), favorable brand opinion (t= 0.78, p-value=0.43), or consideration (t= 0.34, p-value=0.72). Though this is somewhat encouraging, the findings should be treated as suggestive rather than conclusive, since this is a lab study which relies on subjects being capable of simulating their behavior in a social context in an artificial setting.

[8]There was supplementary data on provocativeness ratings, but this was highly correlated with outrageousness as a measure, so to avoid issues with collinearity and convergence, the analysis only uses the outrageousness measure.

Table 1: Summary Statistics

(a) Campaign level

	Mean	Std Dev	Min	Max
Total Views	777996.53	2705048.25	57	37761711
Total Comments	1058.54	4382.75	0	64704
Length Ad (sec)	56.24	33.31	10	120
Observations	396			

(b) Survey Participants' Responses

	Mean	Std Dev	Min	Max
Exposed	0.50	0.50	0	1
Purchase Intent	0.59	0.49	0	1
Intent Scale	3.63	1.12	1	5
Would Consider	0.60	0.49	0	1
Consideration Scale	3.67	1.10	1	5
Favorable Opinion Scale	3.75	0.99	1	5
Favorable Opinion	0.62	0.49	0	1
Aware of Product (Unexposed)	0.56	0.50	0	1
Age	29.57	9.44	18	65
Male	0.70	0.46	0	1
Income (000,USD)	35.53	24.22	20	100
Weekly Internet Hours	26.23	10.93	1	35
Lifetime tasks	6.18	33.68	0	251
Observations	24367			

(c) Average Median Campaign Ratings from Survey Participants

	Mean	Std Dev	Min	Max
Funny Rating	5.64	0.97	2	8
Outrageous Rating	5.13	0.74	1	8
Visual Appeal Rating	6.74	0.66	1	9

3 Empirical Analysis

The randomized procedure for collecting data makes the empirical analysis relatively straightforward.

For person i who was allocated to the testing cell for the video ad for product j, their purchase intent $Intent_{ij}$ is a binary variable which is a function of:

$$Intent_{ij} = I(\alpha Exposed_i + \beta Exposed_i \times LogViews^j + \theta X_i + \delta_j + \epsilon_j > 0) \tag{1}$$

Therefore, α captures the main effect of being exposed to a video ad on purchase intent. Purchase intent is a binary variable for whether the respondent said they were likely or very likely to purchase the product. β captures the core coefficient of interest for the paper, which is whether exposure is more or less effective if the ad has proven to have a high number of views; X_i is a vector of controls for gender, age, income, and time online and the vector θ is their associated coefficients; δ^j is a series of 396 campaign-level product fixed effects that control for heterogeneity in baseline purchase intent for the product in that campaign and includes the main effect of Ad Views ($LogViews_j$), which is why this lower-order interaction is not included in the specification.[9] Using a log measure of ad views avoids the results being driven by extreme values given the large variance in distribution of ad views. The initial specification assumes that the error term ϵ_j is normally distributed, implying a probit specification. Standard errors are clustered at the product level in accordance with the simulation results presented by Bertrand et al. (2004). This represents a conservative empirical approach, as in this setting there is randomization at the respondent level as well.

Table 2 builds up incrementally to the full specification in equation (1). Column (1)

[9]Since this does not give a baseline, the author also explored an alternative specification which used category fixed effects rather than product level fixed effects to allow separate identification of the baseline effect of $LogViews_j$. The coefficient was highly insignificant, suggesting a baseline of zero. Similarly, a coefficient which captured the average views for the placebo ad was also insignificant.

reports an initial specification measuring the main effect of *Exposed* on purchase intent. As expected, being exposed to the video ad has a positive and significant effect on the participant's purchase intent for that product.

The estimate in Column (1) suggests that exposure to a video ad increases purchase probability by 6.6 percentage points, which is similar to the average effect of exposure to 'in-stream' video ads reported by Goldfarb and Tucker (2011a). This is reassuring because that research used industry-collected data where survey-takers were people who had naturally come across the ad in the process of their web-browsing. This suggests that the recruitment method and forced exposure did not overly influence the measure.

Column (2) reruns this simple regression for the websites that had a below-median number of views. Column (3) reports results for the same regression for websites that have an above-median number of views. It is clear that on average the effect of exposure to the ad on purchase intent is greatest for video ads that have a below-median number of views. This is initial evidence of a negative relationship between the total views of the ad and its ability to persuade a viewer to purchase the product.

To test this more robustly, Column (4) provides an explicit test of the apparent difference in size in the coefficients for *Exposed* in Column (2) and (3) by reporting the results of a basic version of (1). The key variable of interest, $Exposed_i \times LogViews_j$, is negative and significant. This suggests that exposure to an ad which has received more views is less likely to be able to persuade an ad viewer to purchase the product.

Table 2: More viewed ads are less persuasive

	Probit (1)	Probit (2) $\geq$ Med Total View	Probit (3) $<$ Med Total View	Probit (4)	Probit (5)	Probit (6)	OLS (7)	OLS (8)
Exposed × LogViews				-0.0153**	-0.0153**	-0.0164**	-0.00658**	
				(0.00738)	(0.00747)	(0.00752)	(0.00268)	
Exposed × Total Views (m)								-0.0153**
								(0.00722)
Exposed	0.181***	0.150***	0.212***	0.246***	0.250***	0.259***	0.0951***	0.0738***
	(0.0177)	(0.0259)	(0.0239)	(0.0363)	(0.0368)	(0.0370)	(0.0134)	(0.00691)
Age					-0.00316***			
					(0.000965)			
Income (000,USD)					0.00116***			
					(0.000342)			
Weekly Internet Hours					-0.0000646			
					(0.000797)			
Male					0.310***	0.247***	0.0893***	0.0893***
					(0.0199)	(0.0208)	(0.00747)	(0.00689)
Product Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Demo Controls	No	No	No	No	No	Yes	Yes	Yes
Observations	24367	12221	12146	24367	24367	24367	24367	24367
Log-Likelihood	-15353.9	-7531.9	-7820.3	-15351.7	-15193.8	-14896.6	-15687.3	-15688.4
R-Squared							0.121	0.121

Dependent variable is binary indicator for whether or not participant states that they are likely or very likely to purchase the product. Probit estimates Columns (1)-(6) and OLS estimates (7)-(8). * p <0.10, ** p <0.05, *** p <0.01. Robust standard errors clustered at the product level.

This finding remains unchanged when adding log-linear controls for consumer characteristics in Column (5), which is as expected due to randomization. These log-linear controls suggest that richer, younger males who do more tasks are more likely in general to say they will purchase. Column (6) uses an alternative non-parametric set of controls for consumer characteristics which are indicators for six levels of income, age and internet usage. As can be seen in the log-likelihood, this non-parametric approach to controls is more efficient, which is why it forms our focal specification. In each case the use of such controls is indicated by a 'Yes' in the Demo Controls row at the bottom of the table.

An econometric concern is the interpretation of the main interaction terms. Research by Ai and Norton (2003) suggests that the interaction in a non-linear model may not capture the true cross-derivative. In order to ensure that the results are not a function of the nonlinearity of the estimation function, Column (7) demonstrates that a linear probability model gives qualitatively similar results, providing reassurance that the non-linear functional form does not drive the results. Column (8) shows that the result is also robust when using a linear version of the key explanatory variable 'Total Views' rather than $LogViews$. The r-squared in each of these columns is relatively low, but this is very much in line with previous studies in this area, such as Aral and Walker (2011).

On the basis of the probit model estimates for the linear measure of views and the appropriate Ai and Norton (2003) correction, Table 2 suggests roughly that for around every 1 million views an ad receives, the ad is on average 10% less persuasive. However, if a video ad is less persuasive for any individual viewer but has the potential to persuade more people because it has higher reach, that is not necessarily harmful to advertiser objectives. Figure 1 plots these rough estimates of a simulation which takes account of the total 'expected' persuasion from a video ad. This is defined as '*Reach*× *Persuasiveness*' and reflects how persuasive the ad was multiplied by how many consumers it was viewed by. The simulation suggests that there are eventually decreasing returns to achieving a large number of

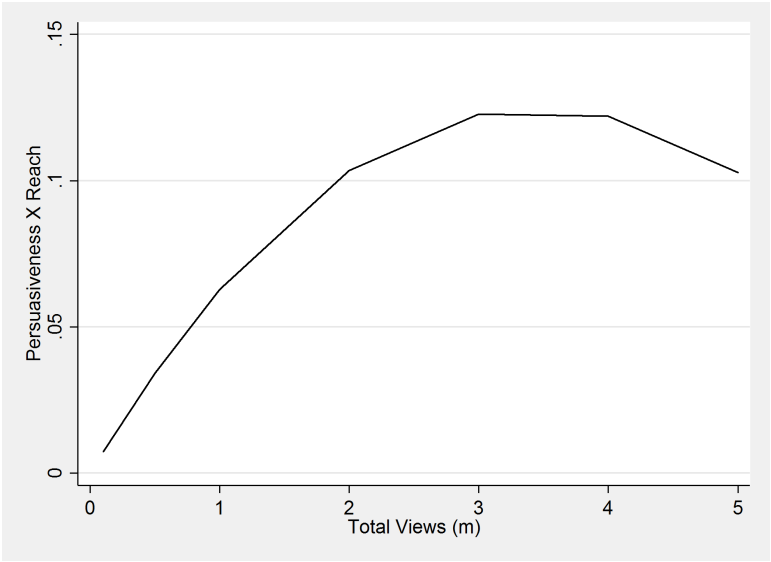

Figure 1: Predictions of Trade-off from Probit Model

total views overall, at three to four million total views. At this point the reduction in ad persuasiveness due to high total reach is large enough that incrementally more consumers viewing the ad achieves little. Only six percent of videos in the data achieved this level of organic reach, so the plot suggests that negative returns to high levels of total reach are limited. Figure 1 is a very rough calculation, but, the existence of inverse-U-shaped returns from achieving high total reach in viral forms of advertising is a new finding deserving of managerial attention.

3.1 Robustness

This section conducts a battery of robustness checks for the results in Table 2.

3.2 Alternative definitions of dependent variables

Table 3 checks the robustness of the results to alternative dependent variables. Column (1) shows robustness to using the entire purchase intent scale. In this OLS specification, the direction of the main effect of interest remains the same, which is to be expected given that the binary indicator for purchase intent was based on this scale. Column (2) repeats this robustness check, but this time uses an ordered probit specification to reflect the potential for non-linearities in interpretation of the scale.

Column (3) shows robustness to looking at an alternative measure of brand persuasiveness, which is whether or not the consumer would consider the brand. This is an important check, as most video advertising is explicitly brand advertising without a clear call to action. Therefore, it makes sense to see that the result applies to an earlier stage in the purchase process (Hauser, 1990). However, the results remain robust (both in significance and approximate magnitude) to a measure which attempts to capture inclusion in a consideration set. This suggests that the documented negative relationship holds across attempts to influence customer attitudes across different stages of the purchase cycle. In a similar spirit, Column (4) shows that the results are robust to using as a dependent variable whether or not the respondent had a 'favorable' or 'very favorable' opinion of the brand.

Table 3: Checking robustness to different dependent variables

	OLS	Oprobit	Probit	Probit
	(1)	(2)	(3)	(4)
	Intent Scale	Intent Scale	Would Consider	Favorable Opinion
Exposed × LogViews	-0.00829**	-0.0119*	-0.0145**	-0.0167**
	(0.00411)	(0.00692)	(0.00737)	(0.00744)
Exposed	0.115***	0.193***	0.274***	0.311***
	(0.0204)	(0.0333)	(0.0359)	(0.0361)
Product Controls	Yes	Yes	Yes	Yes
Demo Controls	Yes	Yes	Yes	Yes
Observations	24367	24367	24367	24367
Log-Likelihood	-25792.5	-21530.4	-14712.0	-14463.4
R-Squared	0.107			

OLS estimates in Column (1). Ordered Probit estimates in Column (2). Probit estimates Columns (3)-(4). Dependent variable is the full five-point purchase intent scale in Column (1)-(2). Dependent variable is whether or not the customer is likely or very likely to 'consider' purchasing the product in Column (3). Dependent variable is whether or not the customer is likely or very likely to have a 'favorable' opinion towards the product in Column (4). Robust standard errors clustered at the product level. * p <0.10, ** p <0.05, *** p <0.01.

3.3 Potential confounds

This section investigates the the robustness of the results in Table 2 for potential confounds that may provide alternative explanations.

One natural concern is that more viral video ads are less effective because the respondents have already been influenced by them, and repeated exposure is less effective (Tellis, 1988). To address this, Column (1) of Table 4 excludes the survey takers who stated they had seen or heard of the advertising campaign before. The results are robust to excluding such observations. This suggests that the explanation of the measured negative relationship is not wearout.

Another concern is that the results are driven by differences between the product categories that the ads were advertising. For example, more aspirational or hedonic categories of products may receive more views (Chiu et al., 2007; Berger and Milkman, 2012), but also find it less easy to persuade people to purchase via advertising. Column (2) of Table 4 addresses this concern, showing that the results are robust to allowing the persuasiveness of the ad to vary by the category of product (for example, whether it is food or a personal care item). The results remain robust to adding these interactions between category-specific indicators and the indicator for exposure, which would capture any differences in advertisers' potential ability to persuade respondents for that product category.

Column (3) addresses the concern that the results are driven by differences in ad length. For example, it could be that longer video ads are more persuasive but less likely to be viewed. To control for this, the specification in Column (3) includes an interaction between exposure and ad length. The results are robust to the inclusion of this control. They also suggest that ad length appears to have little relationship to the perceived persuasiveness of the ad.

Column (4) addresses the concern that the results are driven by differences in elapsed

time for the campaign. For example, it could be that older campaigns gained more views, but that products with older but still live campaigns (perhaps those that were more traditional and less fast-paced) found it more difficult to persuade people to purchase the product. To control for this, Column (4) includes an interaction between exposure and the number of days the campaign had run according to Visible Measures data. The results are robust to the inclusion of this control. They also suggest that on average longer campaigns are more persuasive, which makes sense as it is more likely that ineffective campaigns would be withdrawn.

Column (5) addresses the concern that the results could be an artifact of the fact that workers may have different levels of experience with Mechanical Turk, and that perhaps its overly sophisticated users were more likely to exhibit 'demand effects' by answering the questions in the way they thought that the questioner wanted, and that this might be driving the results if randomization failed. To control for this possibility, Column (5) allows the results to vary by the workers' number of previous tasks for other firms on Mechanical Turk. The results are again similar.

Column (6) addresses the concern that the result could be an artifact of the variation in ages of the survey-takers. For example, if video ads are targeted at young people, and young people are more likely to share ads that older people would disapprove or react poorly to, then this could explain the result. However, the addition of an interaction between the main effect with a variable for age does not change the focal estimates, suggesting that age is not a moderating factor.

Another concern is that potentially the ads could be mainly designed to promote awareness for new products. If the most viral ads were also for the newest products that in turn were harder to persuade consumers to purchase, this could explain the results. To test this, Column (7) adds an extra interaction with an indicator for whether the product had an above-average level of awareness as recorded among consumers who were not exposed to the

ad. The interaction $Exposed_i \times HighAwareness \times LogViews_j$ is insignificant, suggesting that awareness is not an important mediator of the effect.

Table 4: Exploring different explanations

	Seen Before (1)	Cat Int (2)	Ad Length (3)	Campaign Length (4)	Tasks (5)	Age (6)	Awareness (7)
Exposed × LogViews	-0.0181**	-0.0152**	-0.0158**	-0.0194**	-0.0159**	-0.0211***	-0.0158**
	(0.00775)	(0.00755)	(0.00775)	(0.00768)	(0.00743)	(0.00784)	(0.00806)
Exposed × Ad Length			-0.000140				
			(0.000504)				
Exposed × Campaign Length				0.000139*			
				(0.0000747)			
Exposed	0.227***	0.211***	0.264***	0.228***	0.258***	0.280***	0.274***
	(0.0379)	(0.0576)	(0.0420)	(0.0408)	(0.0364)	(0.0394)	(0.0379)
Exposed × Age × LogViews						0.0195	
						(0.0148)	
Exposed × Age						-0.0886	
						(0.0725)	
Age × LogViews						-0.0160	
						(0.0121)	
Exposed × High Aware × LogViews							0.0269
							(0.0240)
Lifetime Tasks					0.00322***		
					(0.000891)		
Exposed × Lifetime Tasks					0.0000373		
					(0.00118)		
LogViews × Lifetime Tasks					-0.000108		
					(0.000177)		
Exposed × LogViews × Lifetime Tasks					-0.0000381		
					(0.000249)		
Exposed × High Aware							-0.274**
							(0.138)
Category Interactions	No	Yes	No	No	No	No	No
Product Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Demo Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	22528	24617	24617	24617	24617	24617	24617
Log-Likelihood	-13807.3	-15056.9	-15060.0	-15058.4	-15033.3	-15059.1	-15055.9

Probit estimates. Dependent variable is binary indicator for whether or not participant states that they are likely or very likely to purchase the product. Robust standard errors clustered at the product level. * p <0.10, ** p <0.05, *** p <0.01.

3.4 Other measures of the viral process

One natural concern about the analysis so far is whether the use of 'reach' in terms of total views captures the essence of what is commonly thought or talked about as virality, since it does not measure directly the organic sharing of videos, but only the ultimate outcome. This section explores measures which use alternative approaches to approximate virality.

As discussed by Cruz and Fill (2008) and Elberse et al. (2011), there are many actions that marketers can take to increase 'reach' of a video ad which do not actually represent true sharing of a viral video. For example, external advertising expenditures can drive viewers towards the website where the advertising agency placed the video originally. An important finding in Elberse et al. (2011) is that the majority of views of a truly viral video stem from user-generated versions of the advertisement. Therefore, they suggest that one way of assessing the successfulness of a 'viral' campaign is to observe how many views are associated with copies of the video which are not placed by the agency but instead placed by fans of the video, since the latter is a more organic process.

Column (1) of Table 5 explores this by only looking at views that can be attributed to the non-advertiser seeded placements. Not all videos had copies, which explains why there are a smaller number of observations than in the main specification. There is still a similar negative relationship, where the persuasiveness of the ad appears to decrease in these non-advertiser-driven views.[10]

As described by Yoganarasimhan (2012) in her study of the effect of bloggers' social relationships on the propagation of `YouTube.com` videos, one way of conceptualizing the virality of videos is the extent to which they are shared across social networks. Unlike in that paper, this paper does not have data on the underlying social relationships between the

[10]Ealier versions of the paper also checked that the result was not an artifact of the fact that total views includes views of derivatives of the original ad. There is the possibility that if an ad were poorly executed, it could have invited scorn in the form of multiple parodies that could have artificially inflated the number of views of the original video. However, the robustness check shows that the results remain robust to excluding views that can be attributed to parodies.

viewers of videos, since rather than being organic and homemade content posted by bloggers who also make public their social networks, the videos were professionally produced content initially posted by professional advertising agencies. Instead, the analysis uses panel data on the growth of views to develop some measures to approximate a viral growth pattern and more accurately reflect the dynamic rather than static nature of virality.

Generally, 'virality' is used to define a process whereby an ad is shared by people successively. To capture this, Column (2) of Table 5 uses as a proxy measure of virality the inter-day correlation in views for that particular campaign. The idea is that ads whose views were the result of a successive sharing process are more likely to have daily views that are positively correlated with views from the previous day. This correlation is unlikely to be causal and highly likely to be biased upwards, as there is no exogenous shifter that allows identification of causal network effects (Tucker, 2008). With this caveat, the results are similar when using this alternative proxy measure.

The Oxford English Dictionary definition of virality is 'the tendency of an image, video, or piece of information to be circulated rapidly and widely from one Internet user to another; the quality or fact of being viral.' To capture this idea of diffusion speed, the paper uses another proxy measure for virality which is whether the time trend for the growth of views is linear or convex. A convex time trend is closer to the common idea of virality, reflecting an increasing spread of content across an ever-growing social network. Column (3) of Table 5 reports the result of this new interaction between persuasiveness and the convexity of the time trend as measured by the extent to which the daily views time trend follows a convex rather than a linear relationship with days elapsed. Though less precisely estimated, the estimate suggests that a convex pattern of growth of views is again associated with lower ad persuasiveness.

Column (4) of Table 5 investigates the relationship between ad persuasiveness and the number of comments that the video posting received. Total comments are 'user-generated

content.'[11] Figure A1 in the appendix displays how comments usually appear below the ad on a video-sharing website.

Of course, total comments are positively linked to the total number of views an ad receives, since without viewers there can be no comments, but it is conceptually distinct as well as requiring a different investment from the viewer. Reflecting these investments, when Visible Measures promote their data on total comments to advertisers, they label this viewer behavior as capturing viewer engagement.[12]

Column (4) of Table 5 explores what happens when $Exposed_i \times LogComments_i^j$ is added to the regression. The pattern for $Exposed_i \times LogViews_i^j$ is similar, if more precise than before. However, crucially, $Exposed_i \times LogComments_i^j$ is both positive and significant. This suggests that video ads that are successful at provoking users to comment on them and engage with them directly are also the ads that are more successful at persuading consumers to purchase the product. Since this is a striking result, Table A1 provides reassuring evidence on the robustness of this specification in the appendix.

However, the number of comments may be subject to spam and other forms of manipulation. To address this, additional data on text of the top five comments for each ad as rated by `YouTube.com` viewers was collected. In particular, this data allows identification of whether, as is the case in Figure A1 in the appendix, one of these comments mentioned the product by name.

Only 41 percent of campaigns had a top five comment which actually mentioned the product by name. Columns (5) and (6) report the results for a stratified analysis of the main sample using this distinction to help strengthen the linkage of the results to this measure of

[11] Such user-generated content is distinct from more general forms of online reputation systems (Dellarocas, 2003), and has been shown by Ghose and Han (2011); Ghose and Ipeirotis (2011) to correlate with product success. Moe and Schweidel (2012) have also shown that comment ratings themselves may be subject to cascades and herding.

[12] This is distinct from physical engagement as measured by Teixeira et al. (2012) using eye-tracker technology.

how well the ad engaged consumer attention around the product. Column (5) shows that ads that were successful in generating comments that actually mentioned the product by name do not experience the key tradeoff identified in the paper. By contrast, Column (6) reports results for campaigns where none of the top five comments mentioned the ad by name. For these campaigns, there is the familiar negative and significant relationship between ad reach and persuasiveness. These results appear to directly underpin the suggested theoretical mechanism for the results, which is that many viral video ads fail to engage consumers around the product as opposed to the non-product-related contents of the video ad.[13]

The next section seeks to enrich these findings by determining how total views, total comments and ad persuasiveness are jointly determined by underlying ad characteristics.

[13]In the spirit of Ghose et al. (2012), the analysis also looked at average comment length and spelling mistakes and mentions of television as potential moderators of the effect, but did not find any statistically significant relationship.

Table 5: Exploring different measures of virality

	Non-Seeded Placements (1)	Correlation (2)	Non-Linear Time Trend (3)	# Comments (4)	Product Mentioned (5)	Product Not Mentioned (6)
Exposed × Copy Views	-0.0189*					
	(0.0102)					
Exposed × Daily Views Correlation		-0.125**				
		(0.0612)				
Exposed × Non-Linear Time Trend			-1.731*			
			(1.012)			
Exposed × Log Views				-0.0379***	-0.00689	-0.0220**
				(0.0143)	(0.0119)	(0.00938)
Exposed × Log Comments				0.0281**		
				(0.0141)		
Exposed	0.340***	0.250***	0.220***	0.420***	0.206***	0.291***
	(0.0845)	(0.0349)	(0.0248)	(0.0931)	(0.0554)	(0.0477)
Product Controls	Yes	Yes	Yes	Yes	Yes	Yes
Demo Controls	Yes	Yes	Yes	Yes	Yes	Yes
Observations	17138	24617	24617	24617	10189	14428
Log-Likelihood	-10589.9	-15159.9	-15061.0	-15057.8	-6264.2	-8789.8

Probit estimates. Dependent variable is binary indicator for whether or not participant states that they are likely or very likely to purchase the product. Robust standard errors clustered at the product level. * p <0.10, ** p <0.05, *** p <0.01.

4 When is there no negative relationship?

4.1 Ad Characteristics

So far, this paper has documented there is a negative relationship between the total views that ads achieve and their persuasiveness. It has also documented that the trade-off is weaker if there are multiple user comments, or if the comments tend to mention the product by name. However, of crucial interest to managers is what actions they can take when designing ads to mitigate this negative relationship.

To explore this, Table 6 provides evidence about how different advertising characteristics moderate this negative relationship. It repeats the estimation from Table 2 but stratifies by whether survey-takers rated the ad as being below or above median in terms of how humorous, visually appealing or outrageous it was. It shows that the tradeoff is weak for ads that are above the median in terms of being funny or visually appealing, but that the tradeoff is larger and more significant for ads that are rated as highly outrageous or not funny or not visually appealing.

Table 6: Effects of underlying ad characteristics

	Not Funny	Funny	Not Visual	Visual	Not Outrageous	Outrageous
	(1)	(2)	(3)	(4)	(5)	(6)
	Purchase Intent	Purchase Intent	Purchase Intent	Purchase Intent	Purchase Intent	Purchase Intent
Exposed × LogViews	-0.0211**	-0.0133	-0.0254**	-0.0109	-0.0152	-0.0225**
	(0.00970)	(0.0119)	(0.0116)	(0.00986)	(0.0108)	(0.0111)
Exposed	0.255***	0.270***	0.334***	0.203***	0.294***	0.245***
	(0.0460)	(0.0612)	(0.0580)	(0.0474)	(0.0579)	(0.0487)
Product Controls	Yes	Yes	Yes	Yes	Yes	Yes
Demo Controls	Yes	Yes	Yes	Yes	Yes	Yes
Observations	12341	12026	12381	11986	12270	12097
Log-Likelihood	-7555.7	-7326.3	-7321.4	-7567.6	-7443.0	-7436.2

Probit estimates. Dependent variable is binary indicator for whether or not participant states that they are likely or very likely to purchase the product. Robust standard errors clustered at the product level. * p <0.10, ** p <0.05, *** p <0.01.

4.2 Combined System of Equations

The results in Table 6 are suggestive as to the potential mechanism which underlines the results. As formalized in (Tellis, 2004), p. 151, ads can achieve high levels of attention but simultaneously experience decreased persuasiveness if they use emotional responses from their viewer to evoke attention. It seems likely that some video ads are purposely being designed to be outrageous in order to command attention and incite consumers to share the video with their friends (Porter and Golan, 2006; Brown et al., 2010; Moore, 2011). However, on average, outrageous ads are not succeeding in persuading consumers to buy products. This is in line with existing research (Vzina and Paul, 1997) on how outrageousness affects ad response. By contrast, ad characteristics such as humor appear to be successful at both promoting user response to the ad and encouraging high levels of organic reach. This is underpinned by behavioral research into humor in ads, which suggests that, unlike with strong emotional stimuli, on average humor does not harm the advertising message and can sometimes enhance it by increasing engagement (Weinberger and Gulas, 1992).

Speculatively, the difference in effect of humor and outrageousness may be because, as discussed by Percy and Rossiter (1992), the tradeoff between attention and persuasion at higher levels of that emotion can be avoided if the stimulus is closely linked to the message of the ad. It is possible that humor and perhaps visual appeal are emotional characteristics of ads that are easier to link to the ads' message. Perhaps echoing this, the majority of the literature such as Duncan and Nelson (1985) has found positive effects on both attention and persuasiveness from incorporating humor into ad messages.

To reflect this, it is possible to expand the analysis to reflect a joint system of equations[14] for both survey-taker i's stated purchase intent and campaign j's total views and comments.

[14]The author owes an anonymous reviewer a great deal of gratitude for helping lay out this system of equations.

$$Intent_{ij} = I(\alpha_{j1} + \alpha_{j2}Exposed_i + \theta DemoVariables_{ij} + \epsilon_{ij})$$

$$\alpha_{j1} = \mu_0 + \mu_1 Views_j + \mu_1 Comments_j + \mu_{31} Funny_j + \mu_{32} Visual_j + \mu_{32} Outrageous_j + \lambda_{j1}$$

$$\alpha_{j2} = \mu_4 + \mu_5 Views_j + \mu_6 Comments_j + \mu_{71} Funny_j + \mu_{72} Visual_j + \mu_{73} Outrageous_j + \lambda_{j2}$$

$$Views_j = \gamma_1 + \gamma_{21} Funny_j + \gamma_{22} Visual_j + \gamma_{23} Outrageous_j + \zeta_{j1}$$

$$Comments_j = \gamma_3 + \gamma_{41} Funny_j + \gamma_{42} Visual_j + \gamma_{43} Outrageous_j + \zeta_{j2} \quad (2)$$

The random effects λ_{j1}, λ_{j2}, ζ_{j1} and ζ_{j2} are jointly estimated using a multivariate normal as a generalized structural equation model. The key vectors of coefficients for the purposes of understanding the effect of ad characteristics on views and comments are $\gamma_{2...}$ and $\gamma_{4...}$. The key vectors of coefficients for understanding the effect of ad characteristics on ad persuasiveness are captured by the vector $\mu_{7...}$.

Column (1) of Table 7 reports the initial results of this approach. The results for $\gamma_{2...}$ suggest that total views increase significantly in the rating for outrageousness and the rating for humor, but do not significantly increase in the rating for visual appeal. By contrast, the estimates for $\gamma 4...$ suggest that total comments increase predominantly in humor, but not in outrageousness or visual appeal. The positive estimate for μ_2 and the negative estimates for μ_5 suggest that in general, product categories with more views have a higher underlying purchase intent for those who see the placebo ad, but that after watching the ad this is less pronounced and instead purchase intent is driven by ad characteristics.

The estimates for $\mu_{7...}$ suggest that ad persuasiveness is a positive function of humor and visual appeal but is negatively affected by ad outrageousness. These estimates for $\mu_{7...}$ are in line with the estimates observed in Table 6, which suggested the persuasiveness and reach tradeoff was less severe for ads that are popular due to their humor or visual appeal. In particular, ads that are rated as humorous can achieve both high persuasiveness and reach.

It also sheds light on the result that ad outrageousness appears to augment the tradeoff. Outrageousness increases ad views, but decreases persuasiveness.

There is also the possibility that there are non-linearities in the effects of ad characteristics on both total views and the effects of exposure. As discussed in (Tellis, 2004), p. 151, while the relationship between attention and the strength of the emotional stimulus may be increasing and linear, the persuasiveness of the ad may be concave, in that persuasiveness originally increases in emotion but at some point decreases. To explore this possibility, equation (3) adds polynomials for the ratings for ad characteristics: Humor, outrageousness and visual appeal.

$$
\begin{aligned}
Intent_{ij} &= I(\alpha_{j1} + \alpha_{j2} Exposed_i + \theta DemoVariables_{ij} + \epsilon_{ij}) \\
\alpha_{j1} &= \mu_0 + \mu_1 Views_j + \mu_1 Comments_j + \mu_{31} Funny_j + \mu_{32} Visual_j + \mu_{33} Outrageous_j + \\
&\quad \mu_{34} Funny_j^2 + \mu_{35} Visual_j^2 + \mu_{36} Outrageous_j^2 + \lambda_{j1} \\
\alpha_{j2} &= \mu_4 + \mu_5 Views_j + \mu_6 Comments_j + \mu_{71} Funny_j + \mu_{72} Visual_j + \mu_{73} Outrageous_j \\
&\quad + \mu_{74} Funny_j^2 + \mu_{75} Visual_j^2 + \mu_{76} Outrageous_j^2 + \lambda_{j2} \\
Views_j &= \gamma_1 + \gamma_{21} Funny_j + \gamma_{22} Visual_j \\
&\quad + \gamma_{23} Outrageous_j + \gamma_{24} Funny_j + \gamma_{25} Visual_j + \gamma_{26} Outrageous_j + \zeta_{j1} \\
Comments_j &= \gamma_3 + \gamma_{41} Funny_j + \gamma_{42} Visual_j \\
&\quad + \gamma_{43} Outrageous_j + \gamma_{44} Funny_j + \gamma_{45} Visual_j + \gamma_{46} Outrageous_j + \zeta_{j2} \quad (3)
\end{aligned}
$$

Column (2) of Table 7 reports the results for this extended model. From the estimates of γ, it seems clear that the major non-linearity in the effect of ad characteristics on total views and total comments is in the effect of visual appeal. It appears that both comments and views exhibit a convex relationship with visual appeal, which suggests that visual appeal

only matters at very high ratings when it comes to provoking either views or comments. When it comes to the effect of ad characteristics on persuasiveness as captured by $\mu_{7...}$, it appears that the major non-linearities are for visual appeal and humor, which exhibit convexity. This suggests that very funny or very visually appealing ads are disproportionately appealing relative to quite funny or quite visually appealing ads. The point estimate for outrageousness, while suggesting some degree of concavity, is not significant at conventional levels. This finding of non-increasing returns echoes Vzina and Paul (1997), who finds a lack of resonant emotional appeal in outrageous ads. This contrasts with emotions such as sadness or anger, which as shown by Kamp and MacInnis (1995) tend to be more strongly associated with concavity.

Table 7: The joint effects of ad characteristics on persuasiveness, views and comments

	Simple (1)	Polynomials (2)
TotalViews		
γ_1	4.146***	4.434***
	(0.0485)	(0.0922)
γ_{21}	0.0133*	0.0372
	(0.00766)	(0.0296)
γ_{22}	0.00265	-0.130***
	(0.00784)	(0.0336)
γ_{23}	0.0159**	0.0103
	(0.00692)	(0.0275)
γ_{24}		-0.00227
		(0.00268)
γ_{25}		0.0112***
		(0.00279)
γ_{26}		0.000612
		(0.00268)
TotalComments		
γ_3	-2.503***	-2.395***
	(0.0459)	(0.0872)
γ_{41}	0.0128*	0.0466*
	(0.00724)	(0.0280)
γ_{42}	-0.00760	-0.0725**
	(0.00741)	(0.0318)
γ_{43}	0.00952	0.000104
	(0.00654)	(0.0260)
γ_{44}		-0.00318
		(0.00254)
γ_{45}		0.00551**
		(0.00264)
γ_{46}		0.000940
		(0.00253)
Main		
μ_0	-1.142***	-0.608***
	(0.0912)	(0.111)
μ_1	0.0754***	0.0746***
	(0.00901)	(0.00903)
μ_2	-0.0335***	-0.0329***
	(0.00942)	(0.00944)
μ_{31}	0.0772***	0.0335
	(0.00596)	(0.0233)
μ_{32}	0.0633***	-0.0324
	(0.00610)	(0.0264)
μ_{33}	0.0412***	-0.0397*
	(0.00544)	(0.0218)
μ_{34}		0.00394*
		(0.00213)
μ_{35}		0.00757***
		(0.00220)
μ_{36}		0.00847***
		(0.00215)
μ_4	0.0434	0.428***
	(0.0988)	(0.134)
μ_5	-0.0407***	-0.0442***
	(0.0125)	(0.0125)
μ_6	0.0206	0.0228*
	(0.0132)	(0.0133)
μ_{71}	0.0147*	0.0438
	(0.00861)	(0.0338)
μ_{72}	0.0746***	-0.0303
	(0.00874)	(0.0379)
μ_{73}	-0.0355***	-0.133***
	(0.00791)	(0.0317)
μ_{74}		0.00886***
		(0.00320)
μ_{75}		0.0105***
		(0.00316)
μ_{76}		-0.00278
		(0.00311)
Demo Controls	Yes	Yes
Observations	23673	23673
Log-Likelihood	-120521.5	-120369.2

Joint estimates for purchase intent, total views and total comments using multivariate normal distribution. Column (1) reports results for estimation based on equation (3). Column (2) reports results for estimation based on equation (3). Robust standard errors clustered at the product level. * p <0.10, ** p <0.05, *** p <0.01.

5 Implications

Firms online are increasingly switching their emphasis from 'paid media' such as online display advertising, to 'earned media' where consumers themselves transmit the message. This has been reflected in the growth of social video advertising, where video ads are now designed to go viral and achieve costless reach. This is a very different distribution system for advertising, compared to a typical placement process where an advertising manager simply decides on how many exposures they want and on what medium to purchase them for. Instead, with viral advertising the advertising manager is responsible for designing ads that will generate their own exposures.

The aim of this paper is to quantify the empirical relationship in social advertising between ads that earn multiple views and ads that are persuasive. Combining historical data and a randomized treatment and control methodology among a large crowdsourced population of survey-takers, the analysis measures this relationship empirically. There is a significant negative relationship between total ad views and ad persuasiveness. The ads that receive the most views are also the ones that are relatively less able to persuade consumers to purchase the product. Accounting for viral ad's larger reach, the negative relationship between views and persuasiveness leads to negative consequences after an ad reaches 3-4 million views. This result is robust to a variety of robustness checks.

Videos receive more comments alongside their views and comments that mention the product were less likely to experience this tradeoff. In other words, ads that are successful at not just at provoking consumers to share the ad with others but also to take time to respond to the ad itself, appear more successful. This suggests that managers should not simply track views but also the nature of user-generated content surrounding their campaigns when evaluating campaign success.

Underlying ad characteristics appear to explain this phenomenon. A joint specification

suggest that ads that achieve high views because they are outrageous, are also less persuasive as a result of this same outrageousness. Though outrageousness is sufficient to induce participants to share an ad, it has a negative effect on the persuasiveness of the ad. By contrast, ads that are humorous can achieve high views and simultaneously be persuasive.

There are of course limitations to this study. First, despite the extensive data collection, these results hold for 400 ad campaigns from the consumer goods category from 2010. It is not clear whether the results would hold for other products or across time. Second, the recruited participants may not be representative of the population, though may be closer to the `YouTube.com` population. This is likely to mean that the estimates are not representative. However, unless this group responds very differently to different ads from the rest of the population, then the general conclusions should hold. Third, all ad design and consequently organic reach or virality is exogenous to the study and was not explicitly manipulated. Fourth, advertisers on video sharing websites may have other objectives such as gathering subscriptions to their online video channel or another form of direct response which is a separate objective from simply shifting purchase intent. The analysis does not have data on these other forms of consumer response. Last, since the data is video ads for well-known consumer goods, it does not allow the study of viral video ads on product awareness. Notwithstanding these limitations, this study does document the potential for an empirical negative relationship between earned reach and ad persuasiveness for ad managers who are trying to exploit the new medium of video advertising.

References

Ai, C. and E. C. Norton (2003, July). Interaction terms in logit and probit models. *Economics Letters 80*(1), 123–129.

Aral, S. and D. Walker (September 2011). Creating social contagion through viral product design: A randomized trial of peer influence in networks. *Management Science 57*(9), 1623–1639.

Barnes, J. H. and M. J. Dotson (1990). An exploratory investigation into the nature of offensive television advertising. *Journal of Advertising 19*(3), 61–69.

Bemmaor, A. C. (1995). Predicting behavior from intention-to-buy measures: The parametric case. *Journal of Marketing Research 32*(2), 176–191.

Berger, J. and K. L. Milkman (2012). What makes online content viral? *Journal of Marketing Research 49*(2), 192–205.

Bertrand, M., E. Duflo, and S. Mullainathan (2004). How much should we trust differences-in-differences estimates? *The Quarterly Journal of Economics 119*(1), 249–275.

Biel, A. L. and C. A. Bridgwater (1990). Attributes of likable television commercials. *Journal of Advertising Research 30*(3), 38 – 44.

Brown, M. R., R. K. Bhadury, and N. K. Pope (2010). The impact of comedic violence on viral advertising effectivensss. *Journal of Advertising 39*(1), 49–65.

Chen, Y., Y. V. Joshi, J. S. Raju, and Z. J. Zhang (January/February 2009). A theory of combative advertising. *Marketing Science 28*(1), 1–19.

Chen, Y., Q. Wang, and J. Xie (2011). Online social interactions: A natural experiment

on word of mouth versus observational learning. *Journal of Marketing Research 48*(2), 238–254.

Chiu, H.-C., Y.-C. Hsieh, Y.-H. Kao, and M. Lee (2007). The determinants of email receivers' disseminating behaviors on the internet. *Journal of Advertising Research 47*(4), 524.

Corcoran, S. (2009, December 16). Defining earned, owned and paid media. Forrester Research.

Cruz, D. and C. Fill (2008). Evaluating viral marketing: isolating the key criteria. *Marketing Intelligence & Planning 26*(7), 743 – 758.

Dellarocas, C. (2003). The digitization of word of mouth: Promise and challenges of online feedback mechanisms. *Management Science 49*(10), 1407–1424.

Duncan, C. P. and J. E. Nelson (1985). Effects of humor in a radio advertising experiment. *Journal of Advertising 14*(2), 33–64.

Eckler, P. and P. Bolls (2011). Spreading the virus: Emotional tone of viral advertising and its effect on forwarding intentions and attitudes. *Journal of Interactive Advertising 11*(2), 1–11.

Eisend, M. (2009). A meta-analysis of humor in advertising. *Journal of the Academy of Marketing Science 37*, 191–203. 10.1007/s11747-008-0096-y.

Elberse, A., C. Lee, and L. Zhang (2011). Viral videos: The dynamics of online video advertising campaigns. *Mimeo, HBS*.

Evans, P. and T. S. Wurster (2000). *Blown to bits: How the new economics of information transforms strategy*. Boston: Harvard Business School Press.

Ghose, A. and S. P. Han (September 2011). An empirical analysis of user content generation and usage behavior on the mobile internet. *Management Science 57*(9), 1671–1691.

Ghose, A. and P. G. Ipeirotis (2011). Estimating the helpfulness and economic impact of product reviews: Mining text and reviewer characteristics. *IEEE Trans. Knowl. Data Eng. 23*(10), 1498–1512.

Ghose, A., P. G. Ipeirotis, and B. Li (May/June 2012). Designing ranking systems for hotels on travel search engines by mining user-generated and crowdsourced content. *Marketing Science 31*(3), 493–520.

Golan, G. J. and L. Zaidner (2008). Creative strategies in viral advertising: An application of Taylor's six-segment message strategy wheel. *Journal of Computer-Mediated Communication 13*(4), 959–972.

Goldfarb, A. and J. Prince (2008). Internet adoption and usage patterns are different: Implications for the digital divide. *Information Economics and Policy 20*(1), 2–15.

Goldfarb, A. and C. Tucker (2011a, May). Online display advertising: Targeting and obtrusiveness. *Marketing Science 30*, 389–404.

Goldfarb, A. and C. Tucker (2011b, January). Privacy regulation and online advertising. *Management Science 57*(1), 57–71.

Hauser, J. R. B. W. (1990, March). An evaluation cost model of consideration sets. *Journal of Consumer Research 16*, 393–408.

Kamp, E. and D. J. MacInnis (1995). Characteristics of portrayed emotions in commercials: When does what is shown in ads affect viewers? *Journal of Advertising Research 35*, 19–28.

Lambrecht, A. and C. Tucker (2013). When does retargeting work? Information specificity in online advertising. *Journal of Marketing Research 50*(5), 561–576.

Manchanda, P., J.-P. Dube, K. Y. Goh, and P. K. Chintagunta (2006). The effect of banner advertising on internet purchasing. *Journal of Marketing Research 43*(1), 98 – 108.

Miller, A. R. and C. Tucker (2013, March). Active social media management: The case of health care. *Information Systems Research 24*(1), 52–70.

Moe, W. W. and D. A. Schweidel (2012). Online product opinions: Incidence, evaluation, and evolution. *Marketing Science 31*(3), 372–386.

Moore, K. (2011, July). 71 percent of online adults now use video-sharing sites. *Pew Internet Project*.

Morwitz, V. G., J. H. Steckel, and A. Gupta (2007). When do purchase intentions predict sales? *International Journal of Forecasting 23*(3), 347–364.

Olenski, S. (2010, December 16). Social video advertising: The explosion of 2010. *Social Media Today*.

Percy, L. and J. R. Rossiter (1992). Advertising stimulus effects: A review. *Journal of Current Issues & Research in Advertising 14*(1), 75–90.

Porter, L. and G. J. Golan (2006). From subservient chickens to brawny men: A comparison of viral advertising to television advertising. *Journal of Interactive Advertising 6*(2).

Reiley, D. and R. Lewis (2009). Retail advertising works! Measuring the effects of advertising on sales via a controlled experiment on Yahoo!". Working Paper, Yahoo! Research.

Steadman, M. (1969). How sexy illustrations affect brand recall. *Journal of Advertising Research 9*(1), 15–19.

Teixeira, T., M. Wedel, and R. Peters (2012). Emotion-induced engagement in internet video ads. *Journal of Marketing Research 49*(2), 144–159.

Teixeira, T. S. (2011). Video ads virality. *Mimeo, HBS*.

Tellis, G. (1988, May). Advertising exposure, loyalty, and brand purchase: A two-stage model of choice. *Journal of Marketing Research 25*(2), 134–144.

Tellis, G. J. (2004). *Effective advertising: Understanding when, how, and why advertising works*. Sage.

Toubia, O., A. Freud, and A. T. Stephen (2011). Viral marketing: A large-scale field experiment. *Economics, Management, and Financial Markets* (3), 43–65.

Tucker, C. (2008). Identifying formal and informal influence in technology adoption with network externalities. *Management Science 54*(12), 2024–2038.

Vakratsas, D. and T. Ambler (1999). How Advertising Works: What Do We Really Know? *The Journal of Marketing 63*(1).

Vzina, R. and O. Paul (1997). Provocation in advertising: A conceptualization and an empirical assessment. *International Journal of Research in Marketing 14*(2), 177 – 192.

Weinberger, M. G. and C. S. Gulas (1992). The impact of humor in advertising: A review. *Journal of Advertising 21*(4), pp. 35–59.

Wilbur, K. C. and Y. Zhu (March/April 2009). Click fraud. *Marketing Science 28*(2), 293–308.

Yoganarasimhan, H. (2012). Impact of social network structure on content propagation: A study using YouTube data. *Quantitative Marketing and Economics 10*(1), 111–150.

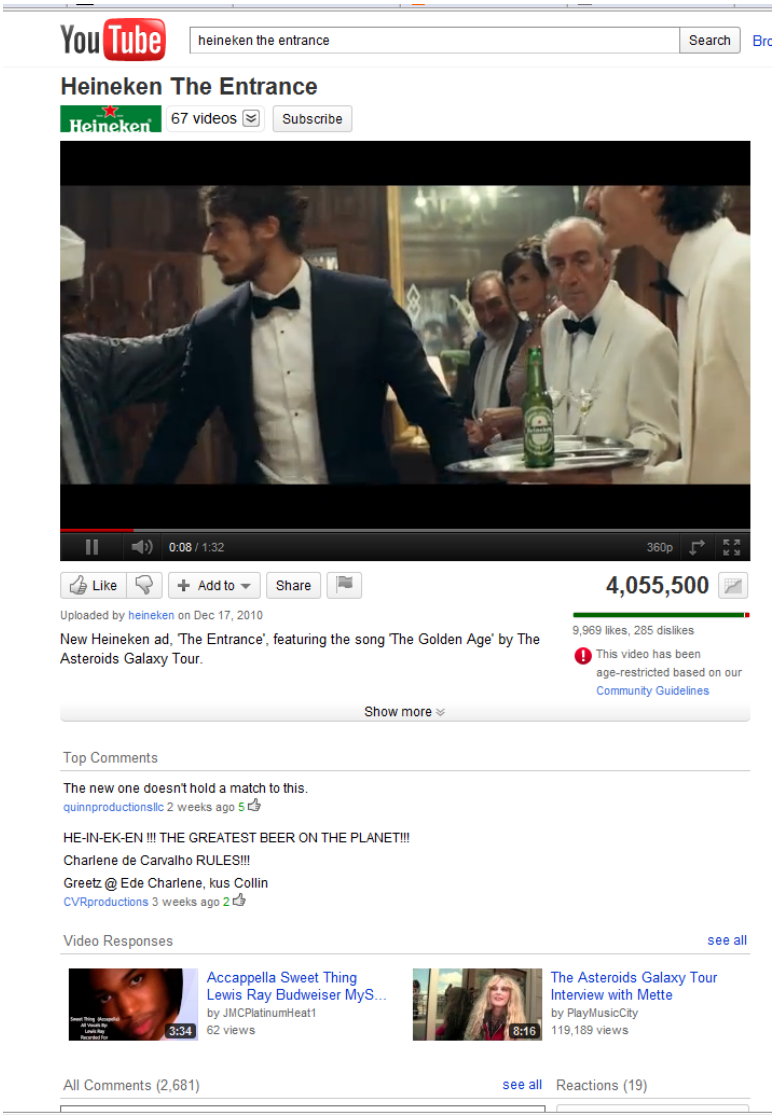

Figure A1: Screen Shot from Typical Video Ad Campaign Showing Comments and Total Views

Table A1: Checks for robustness of Total Comments interaction in Column (4) of Table 5

	Probit	Probit:Not Seen	OLS	OlS	Probit	Probit
	(1)	(2)	(3)	(4)	(5)	(6)
	Purchase Intent	Purchase Intent	Purchase Intent	Intent Scale	Would Consider	Favorable Opinion
Exposed × Logged Views	-0.0379***	-0.0397***	-0.0145***	-0.0202***	-0.0327***	-0.0416***
	(0.0143)	(0.0132)	(0.00510)	(0.00774)	(0.0126)	(0.0127)
Exposed × Logged Comments	0.0281**	0.0282**	0.0103**	0.0156**	0.0238*	0.0326**
	(0.0142)	(0.0139)	(0.00504)	(0.00765)	(0.0133)	(0.0133)
Product Controls	Yes	Yes	Yes	Yes	Yes	Yes
Demo Controls	Yes	Yes	Yes	Yes	Yes	Yes
Observations	24367	22298	24367	24367	24367	24367
Log-Likelihood	-14894.3	-13653.6	-15684.8	-25790.0	-14710.4	-14460.4
R-Squared			0.121	0.107		

In Column (2) all respondents who had seen or heard of the ad before are excluded. Dependent variable is whether someone is likely or very likely to purchase the product in Columns (1)-(3). Dependent variable is the 5-point purchase intent scale in Column (4). Dependent variable is whether someone is likely or very likely to consider the product in Column (5). Dependent variable is whether someone is likely or very likely to have a favorable opinion of the product in Column (6). Robust standard errors clustered at the product level.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.