NET Institute*

www.NETinst.org

Working Paper #11-06

October 2011

Virality, Network Effects and Advertising

Catherine Tucker
MIT Sloan School

* The Networks, Electronic Commerce, and Telecommunications ("NET") Institute, http://www.NETinst.org, is a non-profit institution devoted to research on network industries, electronic commerce, telecommunications, the Internet, "virtual networks" comprised of computers that share the same technical standard or operating system, and on network issues in general.

Virality, Network Effects and Advertising Effectiveness

Catherine Tucker*

September 30, 2011

Abstract

Many video ads are designed to go viral, so that their dissemination depends on customers sharing the ads with their friends. This paper explores whether there is a trade-off between achieving this virality and the effectiveness of the ad at persuading a consumer to purchase or adopt a favorable attitude towards a product. In other words, do ads, by being the kind of ads that achieve virality, sacrifice elements that would be better at persuading people to actually buy products? The analysis combines data on the real-life virality of 400 video ad campaigns, and crowd-sourced measurement of advertising effectiveness among 24,000 consumers. Effectiveness is measured by randomly exposing half of these consumers to a video ad and half to a placebo ad, and then surveying their attitudes towards the product. We find that ads that were more 'viral,' that is, ads that had achieved more views on websites such as `Youtube.com`, were indeed less effective at persuading consumers to purchase or adopt a favorable attitude to a brand. Relative ad-effectiveness dropped by roughly 10% for every million views. Taking into account the advantages of increased reach, this means that there was a decline in overall advertising effectiveness at 3-4 million views. Importantly, ads that generated both views *and* consumer engagement in the form of comments did not suffer from the same tradeoff. Such ads were also be less intentionally provocative or outrageous than ads and more likely to be viral due to humor or attractive visual-design.

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 Net Institute (http.www.NETinst.org) for providing financial support. All mistakes are mine alone.

1 Introduction

Social video advertising is among the fastest growing segments in advertising today. In 2010, social video advertising views increased 230%, over nine times more than search and display impression growth. These video ads are crucially different from rich-media banner ads. Rather than the advertiser paying for placement, such ads are designed to be shared and passed along consumers themselves. This means that these video ads are posted 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. Therefore, many video ad campaigns are intentionally designed to 'go viral' and achieve this costless form of reach.

This paper seeks to understand whether there is a trade-off between the creation of such endogenous network effects for advertising and advertising effectiveness. Can achieving 'virality' be costless for the firm, in terms of how well the advertising works in persuading people to purchase the product? This research represents a departure from the standard network effects literature, as network effects are no longer taken as exogenous; in order to increase the network effects associated with the ad, advertisers may have to sacrifice other aspects of ad design.

We use historical data on the number of times that 400 different video ad campaigns from the past year were shared. This data comes from a media metrics company that tracks major advertiser video ads and records the number of times these ads are shared and viewed. The effectiveness of these campaigns is then measured using techniques pioneered by media metrics agencies and previously used in data analysis by Goldfarb and Tucker (2011a). After recruiting 24,000 respondents through crowdsourcing, we measure the effect of exposure to the 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 of similar length. They are then asked questions about their purchase intents and brand attitudes towards the focal product.

The randomization induced by the field-test procedure means that econometric analysis is straightforward. First, we document whether or not there is a trade-off between creating endogenous network effects and advertising effectiveness. We find evidence that indeed ads that achieved more views were less successful at increasing purchase intent. We show that this is robust to different functional forms and alternative definitions of the explanatory and dependent variable. It is also robust to controls that allow the effect of exposure to vary by ad length and category. It is also robust to excluding respondents who had seen or heard of the ad before, meaning that the results do not simply represent satiation.

We present estimates of the magnitude of this trade-off and suggest that on average, every one million views is associated with a 10 percent drop in effectiveness. Of course, this drop of effectiveness was compensated for by the increased reach of the highly viral campaign, so we also present some rough projections concerning at what point the decreased effectiveness at persuasion outweighs the increased number of views in terms of the total persuasion exerted over the population. Our estimates suggest that this point occurs between 3-4 million views, a viewership that achieved by only 6% of campaigns in our data.

The crucial managerial question, though, is whether there are *categories* of ads for whom this trade-off between virality and effectiveness did not exist. Such cases would be a clear 'win-win' for advertising managers, where virality does not have to be costly in terms of the persuasiveness of the ad design. We found that viral ads that also induced consumers to comment on the ad, rather than just encouraging them to share it with others, did qualify as 'win-wins.' This has an important managerial implication. Marketing managers, as well as tracking total views for their ads, should also take into account other measures of viewer engagement such as the creation of user-generated content surrounding the ads. This should be used as an early indicator of successful engagement on the part of the ad, and its likely ability to be persuasive as well as viral.

We present evidence that such ads that were less likely to experience the trade-off were also less likely to be rated as being provocative or outrageous by participants. Instead, they were more likely to be rated as funny or visually appealing. Therefore, one explanation of our results is that videos are going viral because they are intentionally provocative or outrageous, but that such ad design does not necessarily make the ads more persuasive.

This paper contributes to two existing academic literatures.

The first literature is one of network effects and virality, which studies the potential for endogenous network effects. Aral and Walker (2011) studies this question in the context of product design. He found that, using evidence from a randomized field trial for an application on Facebook, forcing a product 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. (2009) 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.

Some recent papers have modeled the determinants of whether or not a video ad-campaign goes 'viral' This is increasingly important given that 71 percent of online adults now use video-sharing sites Moore (2011). Porter and Golan (2006) emphasize the importance of provocative content (specifically sexuality, humor, violence, and nudity) as a determinant of virality while Chiu et al. (2007) emphasized that hedonic messages are more likely to be shared by e-mail. Elberse et al. (2011) examined 12 months of data on popular trailers for movies and video games. She found evidence that their popularity was often driven by their daily advertising budget: Offline awareness stimulated network effects and virality. Teixeira (2011) examines what drives people to share videos online. He distinguishes between social utility and content utility and highlights the role of non-altruistic sharing behavior. Though they provided important empirical evidence about the drivers of virality, these papers did not actually measure how effective the video ads were.

The second literature is on the effectiveness of online advertising. Much of this literature has not considered the kind of advertising that is designed to be shared, instead focusing on non-interactive banner campaigns Manchanda et al. (2006); Lambrecht and Tucker (2011). Generally, this literature has only considered the effectiveness of video-advertising tangentially or as part of a larger study. Goldfarb and Tucker (2011a) presented results that video advertising is less effective when placed in a context which matched too closely the product being advertised. We believe that this is the first study of the trade-off between ad virality and ad effectiveness, that is, how the ability of an ad to endogenously gain 'reach' is related to the ability of the ad to persuade.

2 Data

2.1 Video Virality Data

We obtained data from a large video metrics company, Visible Measures. Different data for a smaller number of campaigns from this company has also been used Elberse et al. (2011) to study the effects of offline advertising budgets on video virality for movie previews. Visible Measures is an independent third-party media measurement firm for online video advertisers and publishers founded in 2005. It is the market leader in terms of tracking views and engagement for different types of social video ads. Visible Measures shared data with us for recent campaigns in the consumer goods category from 2010. We requested explicitly that they exclude from the data video ads for categories of products such as cars and other large ticket items, for which the majority of people were unlikely to be in the market. We also requested that they exclude video ads for entertainment products such as movies Video Games, and DVDs whose ads have a short-shelf life.

The videos of 396 of these campaigns were still live and online and consequently were included in this survey. Table 1a reports campaign-level summary statistics. 'Total views'

	Mean	Std Dev	Min	Max	Observations
Total Views (m)	777996.53	2705048.25	57	37761711	396
Total Comments	1058.54	4382.75	0	64704	396
Funny Rating	5.64	0.97	2	8	396
Provocative Rating	5.27	0.66	1	8	396
Outrageous Rating	5.13	0.74	1	8	396
Visual-Appeal Rating	6.74	0.66	1	9	396
Length Ad	56.22	33.35	0	120	396

(a) Campaign Level

	Mean	Std Dev	Min	Max	Observations
Exposed	0.50	0.50	0	1	24367
Purchase Intent	0.59	0.49	0	1	24367
Intent Scale	3.63	1.12	1	5	24367
Would Consider	3.67	1.10	1	5	24367
Age	29.57	9.44	18	65	24367
Male	0.70	0.46	0	1	24367
Income (USD)	35.53	24.22	20	100	24367
Weekly Internet Hours	26.23	10.93	1	35	24367

(b) Survey Level

Table 1: Summary Statistics

captures the number of times these videos had been viewed by consumers. This encompasses both 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 that we use a logged measure in our regressions. We also show the robustness of our results to the linear measure. '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`.

We wanted to gather data on advertising effectiveness. An issue with video advertising is that typical 'direct response' methods of evaluating digital advertising, such as measuring click-throughs, are not appropriate. As documented by Porter and Golan (2006); Golan and Zaidner (2008) viral advertising very rarely has a clear 'call to action', such as visiting a website, that is measurable. 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. Therefore, we test advertising effectiveness based on industry-standard techniques for measuring the effectiveness 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. As reported by Goldfarb and Tucker (2011a), both major advertisers and major agencies use these same techniques for evaluating both banner campaigns and video campaigns.

One issue is that because such ad effectiveness measures were not used as the campaigns were being rolled out, we have to collect this data retrospectively. Given the number of campaigns we want to evaluate, this requires a large number of participants. We obtain this large number using crowd-sourcing techniques. Specifically, we recruited 25,000 separate individuals using the crowd-sourcing platform Mechanical Turk. Similar crowd-sourcing techniques have been used by Ghose et al. (2011) 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 the 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.

For each campaign, we recruited on average 60 respondents. Half of the respondents are allocated to a condition where they are exposed to the focal video ad that we have virality data on. The other half of respondents (the control group) see a placebo ad for another unrelated (random) product that was also part of our data. This randomization means that in expectation all our respondents are identical. Therefore we can causally attribute any differences in their subsequent attitudes towards the product to ad exposure.

We record whether or not the respondent watches the video all the way through and exclude those who left the website from our data. We also exclude participants who, despite the controls in place, managed to take the survey multiple times. This explains why we have 24,367 respondents which is fewer than the original 25,000 respondents in our analysis. We then ask them a series of survey questions. Table 1b summarizes these responses. These include questions about their purchase intent towards the focal product and likelihood of consideration of the focal product. We also included decoy questions about another brand. All these questions are asked on a 5-point scale in line with traditional advertising effectiveness questioning (Morwitz et al., 2007). Following Goldfarb and Tucker (2011a), we converted this 5-point scale to a binary purchase intent measure that captures whether someone is very likely or likely to purchase the product for our main analysis. However, we show robustness to the use of the full scale in subsequent regressions.

One of the attractive features of this form of online ad measurement is that survey-based measures of purchase intent can be collected consistently across different products (Clark et al., 2009). Survey responses are weaker measures of advertising effectiveness than purchasing or profitability (as used by Reiley and Lewis (2009)), because though 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 our 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 our results should be interpreted as a reflecting an industry standard for a relative unitless measure of online effectiveness that is used as an input when making advertising allocation decisions.

In addition to asking about purchase intent, the survey also asked participants about whether or not they recall having seen the focal video ad before or had heard it discussed by their friends and media. We use this information in a robustness check to make sure that the fact that respondents are more likely to have seen viral videos before and there may be less of an effect the second time, is not driving our results. We also asked participants to rate the video on a 10-point sliding scale based on the extent to which they found it humorous, visually appealing, provocative or outrageous. We then use the median ratings for the campaign in our mediation analysis of what drives the effect. Table 1a reports these ratings at the campaign level, based on the median response of our survey-takers.

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. These 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 our survey-takers were. It is clear that they are more male than the general population, earn less, and also spend more time online. The fact that there were more males than females reflects video-sharing site usage. Based on a survey conducted by Moore (2011), men are 28 percent more likely than women to have used a video-sharing site recently. However, we accept that since these participants were recruited via a crowd-sourcing website, there is the possibility that they may differ in unobserved ways from the population. Another related issue is whether because respondents were in a forced exposure setting and a less natural setting than most people view video ads in, this may also have altered responses.

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.

3 Empirical Analysis

This randomized procedure for collecting data makes our empirical analysis relatively straightforward.

For person i who was allocated to the testing cell for video ad for product j, their purchase intent reflects:

$$Intent_{ij} = I(\alpha Exposed_{ij} + \beta Exposed_{ij} \times LoggedAdViews_i^j + \theta X_{ij} + \delta_j + \epsilon_{ij} > 0) \quad (1)$$

Therefore, α captures the main effect of being exposed to a video ad on purchase intent; β captures the core coefficient of interest for the paper - whether exposure is more or less effective if the ad has proven to be viral; X_{ij} is a vector of controls for gender, age, income, and time online; δ^j is a series of 397 consumer good product-level fixed effects that control for heterogeneity in baseline purchase intent for that product and includes the main effect of Ad Views ($LoggedAdViews_i^j$), which is why this lower-order interaction is not included in our specification. We used a logged measure of ad views, because we do not want our results to be biased by extreme values given the large variance in distribution of ad views. However, we show robustness to other measures subsequently. In our initial regressions, we assume that the ϵ_{ij} is normally distributed, implying a probit specification, though we subsequently show robustness to other functional forms. 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 our setting we have randomization at the respondent level as well.

Table 2 shows our initial results that investigate the relationship between ad-effectiveness and virality as measured by total views of the video. Column (1) reports an initial specification where we simply measure 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 magnitude suggests that exposure to a video increases purchase probability by 6.6 percentage points, which is a similar order of magnitude to the average effect of exposure to video ads reported by Goldfarb and Tucker (2011a). This is reassuring because Goldfarb and Tucker (2011a) used industry-sponsored data where survey-takers were people who had naturally come across the ad in the process of their web-browsing. This therefore suggests that the recruitment method and forced exposure did not noticeably influence our 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 our first evidence that there may be a trade-off between the virality of the ad and its effectiveness at persuading a viewer to purchase the product.

Table 2: More viewed ads are less effective

	Probit	Probit	Probit	Probit	Probit	Probit	OLS	OLS
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
		$\geq$ Med Total View	$<$ Med Total View					
Exposed × Logged Total Views				-0.0153**	-0.0153**	-0.0169**	-0.00678**	
				(0.00738)	(0.00747)	(0.00746)	(0.00267)	
Exposed × Total Views (m)								-0.0154**
								(0.00784)
Exposed	0.181***	0.150***	0.212***	0.246***	0.250***	0.261***	0.0961***	0.0740***
	(0.0177)	(0.0259)	(0.0239)	(0.0363)	(0.0368)	(0.0368)	(0.0135)	(0.00753)
Age					-0.00316***			
					(0.000965)			
Income (USD)					0.00116***			
					(0.000342)			
Weekly Internet Hours					-0.0000646			
					(0.000797)			
Male					0.310***	0.254***	0.0913***	0.0912***
					(0.0199)	(0.0206)	(0.00742)	(0.00742)
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	-14958.3	-15747.5	-15748.7
R-Squared							0.117	0.117

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Robust standard errors clustered at the product-level. 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).

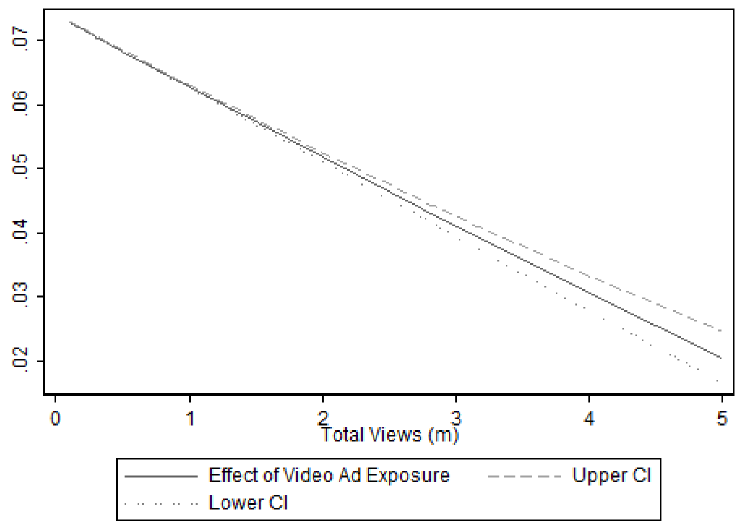
To test this more robustly, Column (4) provides an explicit test of the statistical differences on the coefficients for *Exposed* in Column (2) and (3) by reporting the results of a basic version of (1). The key variable of interest $Exposed_{ij} \times LoggedAdViews_i^j$ is negative and significant. This suggests that exposure to an ad which has received more views is less likely to be effective at persuading an ad viewer to purchase the product.

This finding remains unchanged when we add linear controls for consumer characteristics in Column (5) which is as expected due to randomization. These linear controls are such that richer, younger males 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 simply 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 we use it for the rest of the specifications. In each case the use of such controls is indicated by a 'Yes' in the Demo Controls row at the bottom of the table.

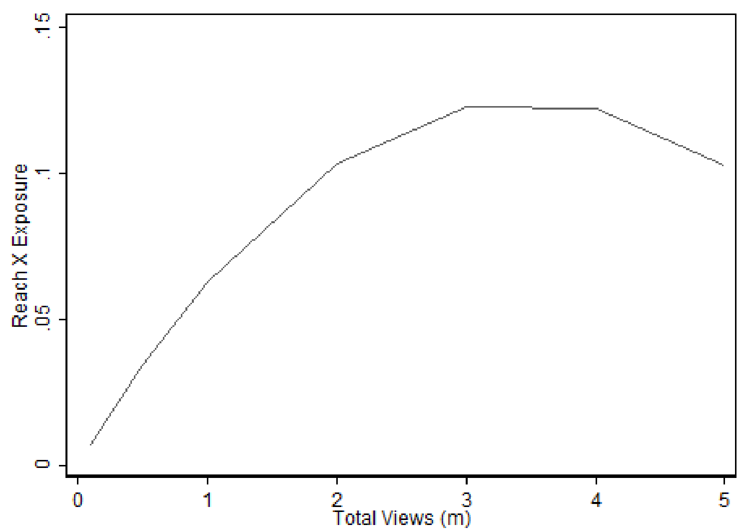
An econometric concern is the interpretation of the main interaction terms. Research by Ai and Norton (2003) suggests that the interaction in a non-linear model may not capture the true cross-derivative. In order to ensure that our results are not a function of the nonlinearity of the estimation function, we also show in Column (7) that a linear probability model gives qualitatively similar results, providing reassurance that the non-linear functional form does not drive our results. In Column (8), we show that the result is also robust if we use a linearized form of our key explanatory variable 'Total Views' rather than the logged form. 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); Goldfarb and Tucker (2011a).

To give an idea of the magnitude of these estimates, we used a probit model and the appropriate correcting of Ai and Norton (2003) to calculated different predicted values at different numbers of total (non-logged) views. Figure 1a presents the results. There is a sizeable loss of effectiveness for ads that received a larger number of views, and it suggests that roughly for around every 1 million views there is a 10 percent decline in effectiveness.

However, this is not the whole story, as of course by definition the most viral videos had improved reach, meaning that they while they were less effective for any individual viewer, they also potentially were able to persuade more people. To take account of this, we did a rough simulation where we took account of the total expected persuasion from a video ad, which of course reflects how persuasive the ad was and by how many consumers it was viewed by. Figure 1b plots these rough estimates. Our simulation suggests that there are eventually decreasing returns to achieving virality overall, at the 3-4 million total views mark. At this point ad-effectiveness is low enough that incrementally more consumers viewing the ad achieves little. Only 6 percent of videos in our data achieved this level of virality, so our plot suggests that negative returns to virality are generally limited. Figure 1b is also a very rough back-of-the envelope calculation. However, the existence of inverse-U-shaped returns from achieving virality in advertising is a new finding and one that deserves managerial attention.



(a) Linear Prediction of Tradeoff



(b) Aggregate Effects of Tradeoff

Figure 1: Predictions from Probit Model

3.1 Robustness

One natural concern given our use of historical data is that our results may be biased because a general prior awareness of a campaign or its success may influence respondents' answers to questions about advertising effectiveness. This would provide an alternative explanation of our findings, that the reason that more viral video ads are less effective is because the respondents have already been influenced by them, and repeated exposure is less effective. We address this in Column (1) of Table 3 as we exclude our crowd-sourced field testers who stated they had seen or heard of the advertising campaign before. Our results are robust to excluding such observations. This suggests that the explanation to the trade-off lies in the ads' own characteristics rather than in wearout among the general population.

In Column (2) we address another natural concern, which is that the number of placements (that is the number of websites) that the video was posted on drove the result. As discussed by Cruz and Fill (2008), the process whereby an ad agency determines the number of placements is highly strategic. Therefore an alternative interpretation of the measured trade-off would simply be that videos with multiple placements got more views but the multiple placements themselves were in response to acknowledged ad ineffectiveness. However, even when we control for placements by using a measure of the average number of views per placement, the result holds, and if anything is more precisely estimated.

Column (3) addresses the concern that our result could be 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 parodic derivatives that could have artificially inflated the number of views. However, the robustness check shows that our results remain robust to excluding views that can be attributed to derivatives.

Another concern is that our results may simply be being driven by category difference between the ads. For example, more aspirational or hedonic categories of products may receive more views but also be less easy to persuade people to purchase via advertising. Column (4) addresses the concern and shows that the results are robust to our allowing the effectiveness 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 the addition of these interactions between category-specific indicators and the indicator for exposure.

Column (5) addresses the purely mechanical concern that the results are driven by differences in ad length. For example, it could be more likely that longer video ads are more persuasive but less likely to be shared. To control for this, we included an interaction between exposure and ad length. Our results are robust to the inclusion of this control. They also suggest, interestingly, that ad length appears to have little relationship with the perceived persuasiveness of the ad.

In Columns (6) and (7) of Table 3, we check the robustness of our results to alternative dependent variables. Columns (6) show 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 (7) 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 our result applies to an earlier stage in the purchase process (Hauser and Wernerfelt, 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 trade-off holds across attempts at influence customer attitudes across different stages of the purchase cycle.

Table 3: Checking Robustness

	Not Seen (1) Purchase Intent	Adj Placements (2) Purchase Intent	Adj Deriv. (3) Purchase Intent	Cat Int (4) Purchase Intent	Ad Length (5) Purchase Intent	Ord Probit (6) Intent Scale	OLS (7) Would Consider
Exposed × Logged Total Views	-0.0186**			-0.0158**	-0.0165**	-0.00860**	-0.0152**
	(0.00773)			(0.00780)	(0.00773)	(0.00410)	(0.00736)
Exposed × Placement Adjusted Views		-0.0235**					
		(0.0115)					
Exposed × Logged Non-Derivative Views			-0.0190**				
			(0.00752)				
Exposed × Ad Length					-0.000120		
					(0.000507)		
Exposed	0.226***	0.248***	0.270***	0.214***	0.266***	0.117***	0.277***
	(0.0381)	(0.0337)	(0.0365)	(0.0496)	(0.0422)	(0.0205)	(0.0358)
Category Interactions	No	No	No	Yes	No	No	No
Product Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Demo Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	22298	24367	24301	24367	24367	24367	24363
Log-Likelihood	-13718.9	-14958.5	-14921.1	-14955.1	-14958.3	-25855.5	-14773.9
R-Squared						0.102	

* $p < 0.10$, ** $p < 0.05$,*** $p < 0.01$. Robust standard errors clustered at the product-level. Probit estimates in Columns (1)-(5), (7). OLS estimates in Column (6). Dependent variable is binary indicator for whether or not participant states that they are likely or very likely to purchase the product in Columns (1)-(5). Dependent variable is the full five-point purchase intent scale in Column (6). Dependent variable is whether or not the customer is likely or very likely to 'consider' purchasing the product in Column (7).

4 Engagement and Virality

We want to explore whether there are some circumstances when virality and advertising effectiveness are complements and others when they are substitutes. The intent of this analysis is to identify when there is a trade-off between ad effectiveness and virality, and more importantly for managers when there is not. This means that we can offer some practical guidance as to occasions when ads can both be inherently viral and effective at inducing purchase intent.

We do this by introducing an explicit measure of viewer engagement (rather than simply virality) to our regressions. This is the 'total comments' that an ad receives. It is a measure of the success of the ad at promoting engagement online. Of course it is linked to the ultimate virality of the ad, since without viewers there can be no comments, but it is conceptually distinct.

Total comments are best understood as a form of related 'user-generated content', distinct from online product ratings, such has been shown by Ghose and Han (2011); Ghose and Ipeirotis (2011) to correlate with product success. Moe and Schweidel (2011) has also show that comment ratings themselves may be subject to cascades and herding.

Table 4 explores what occurs when we include this measure of engagement into our regressions. In Column (1) we show what happens when we add $Exposed_{ij} \times LoggedComments_i^j$ to our regression. The pattern for $Exposed_{ij} \times LoggedViews_i^j$ is similar if more precise than before. However, crucially $Exposed_{ij} \times LoggedComments_i^j$ 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. Table A1 in the appendix shows that this result holds across the different functional forms and dependent variables that we did robustness checks for in Table 3.

Column (2) shows our results hold when we explicitly measure the effect of a ratio of $\frac{TotalComments}{TotalViews}$. The larger the ratio, the more likely the ad was to have been at promoting purchase intent. Column (3) and (4) show what this result implies when we stratify our sample by ratio. Column (3) shows that there was a strong trade-off for ads that were in the bottom 50 percent in terms of their ratio of comments to views. Column (4) shows that ads that were in the top 50 percent of their comments:views ratio experienced very little trade-off between ad-effectiveness and virality.

Table 4: What Mediates the Trade-off?

	Probit	Probit	Probit	Probit
	(1)	(2)	(3)	(4)
	Both	Ratio	Bottom 50% Ratio	Top 50% Ratio
Purchase Intent				
Exposed × Logged Total Views	-0.0387***		-0.0303***	0.00311
	(0.0143)		(0.00990)	(0.0110)
Exposed × Ratio Views: Comments		0.0634***		
		(0.0189)		
Exposed	0.424***	0.630***	0.347***	0.136***
	(0.0932)	(0.132)	(0.0501)	(0.0508)
Exposed × Logged Total Comments	0.0285**			
	(0.0142)			
Product Controls	Yes	Yes	Yes	Yes
Demo Controls	Yes	Yes	Yes	Yes
Observations	24367	24367	12207	12160
Log-Likelihood	-14955.9	-14954.0	-8002.8	-7957.2

* p <0.10, ** p <0.05,*** p <0.01. Robust standard errors clustered at the product level.

We then go on to explore what underlying ad characteristics drive this distinction between the effect of overall virality and engagement. Table 5 indicates the ad characteristics that are linked both with high views and with this desirable high ratio between comments and views. It is clear that the ads that are both more likely to attract a large number of total views but less likely to attract a high ratio of comments to views are the ones that are intentionally provocative or outrageous in their ad design. On the other hand, the ads which are visually appealing and funny appear successful at eliciting comments and,

Table 5: Correlation of Ad Characteristics with Total Views and Comments Ratio

	Total Views (m)	Total Comments:Total Views Ratio
Outrageous Rating	0.169***	-0.0191**
Provocative Rating	0.155***	-0.0381***
Funny Rating	0.144***	0.0131*
Visual-Appeal Rating	0.115***	0.0203**

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

though successful at attracting more views, are less viral than those that are provocative or outrageous.

This provides evidence about why the measured trade-off exists between advertising virality and advertising effectiveness. Some video ads are purposely being designed to be outrageous or provocative with the aspiration of inciting consumers to share the video with their friends. However, on average, they are neither provoking responses among viewers to the actual ad itself nor by implication succeeding in persuading users to purchase the product. In other words, being outrageous is a reliable strategy for encouraging virality, but it reduces the persuasiveness of ads. On the other hand, ad characteristics such as humor appear to be successful at both promoting user response to the ad as well as virality.

5 Implications

Many 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 and purchases them. Instead, the advertising manager is responsible for designing ads that will generate their own exposures and therefore (goes the thinking) be effective.

The aim of this paper is to quantify whether there is a trade-off in ad-design between making sure that ads are effective in terms of actually leading consumers to be more likely to purchase the product and ads that are viral. Combining historical data and a randomized treatment and control methodology among a large crowd-sourced population of survey-takers, we measure this trade-off empirically. We find evidence that there is a significant trade-off between virality and ad effectiveness. The ads that receive the most views are also the ones that are least effective at persuading consumers to purchase the product. We present evidence that after adjusting for the improved reach (that is the larger number of people who view the ads) of viral videos, this trade-off only leads to negative consequences after an ad reaches 3 million views, therefore affecting the top six percent of ads by views in our sample. We check the robustness of our results in a variety of ways.

We then provide some evidence about why this occurs. Videos that receive many comments relative to views do not suffer this trade-off. Instead they appear to continue to be effective. In other words, ads that are successful 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. We present some suggestive evidence about why this occurs. The ads that do worst by these metrics are ads that are viral by virtue of being of their being rated as outrageous or provocative. Though provocative ad design is sufficient to induce participants to share an ad, it is not sufficient to induce them to respond or be influenced by the ad. On the other hand, ads that are viral by virtue of their humor or their visual design appear to be less likely to suffer this trade-off.

There are of course limitations to this study. First, despite the extensive data collection, these results hold for 400 ad campaigns for the consumer goods category from 2010. It is not clear that the results would hold for other products or across time. Second, the participants that we recruited to measure may not be representative of the population. Though this is likely to mean there may be measurement error when it comes to our precise calibration of the trade-off, unless this group responds very differently to different ads from the rest of the population then our general conclusions should hold. Third, all ad design and consequently virality is exogenous to the study and was not explicitly manipulated. Notwithstanding these limitations, this study does document a new trade-off between the creation of endogenous network effects and ad virality 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. *Economic 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.

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

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.

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.

Clark, C. R., U. Doraszelski, and M. Draganska (2009, June). Information or Persuasion? An Empirical Investigation of the Effect of Advertising on Brand Awareness and Perceived Quality using Panel Data. *Quantitive Marketing and Economics 7*(2), 207–236.

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

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

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 (2011). Designing ranking systems for hotels on travel search engines by mining user-generated and crowd-sourced content. *Mimeo, NYU*.

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

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

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

Hauser, J. R. and B. Wernerfelt (1990, March). An evaluation cost model of consideration sets. *Journal of Consumer Research 16*(4), 393–408.

Lambrecht, A. and C. Tucker (2011). Online consumer behavior: Retargeting and information specificity. *Mimeo, LBS*.

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.

Moe, W. W. and D. A. Schweidel (2011). Online product opinions: Incidence, evaluation, and evolution. *Marketing Science*.

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.

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.

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

Toubia, O., A. T. Stephen, and A. Freud (2009). Viral Marketing: A Large-Scale Field Experiment. *Mimeo, Insead*.

Table A1: What Mediates the Trade-off?

	Probit:Not Seen	Probit	Probit	OLS	OlS
	(1)	(2)	(3)	(4)	(5)
	Purchase Intent	Purchase Intent	Purchase Intent	Intent Scale	Would Consider
Exposed × Logged Total Views	-0.0402***		-0.0147***	-0.0206***	-0.0336***
	(0.0150)		(0.00511)	(0.00775)	(0.0126)
Exposed × Logged Total Comments	0.0283*		0.0103**	0.0157**	0.0242*
	(0.0150)		(0.00506)	(0.00767)	(0.0132)
Exposed × Placement Adjusted Views		-0.0763***			
		(0.0226)			
Exposed × Placement Adjusted Comments		0.0574***			
		(0.0186)			
Exposed	0.388***	0.636***	0.156***	0.207***	0.415***
	(0.0980)	(0.134)	(0.0336)	(0.0505)	(0.0841)
Product Controls	Yes	Yes	Yes	Yes	Yes
Demo Controls	Yes	Yes	Yes	Yes	Yes
Observations	22298	22968	24367	24367	24363
Log-Likelihood	-13716.7	-14055.6	-15745.0	-25852.9	-14772.2
R-Squared			0.117	0.102	

* p <0.10, ** p <0.05,*** p <0.01. Robust standard errors clustered at the product level.