Super Returns to Super Bowl Ads?[*]

Seth Stephens-Davidowitz
seth.stephens@gmail.com

Hal Varian
hal@ischool.berkeley.edu

Michael D. Smith
mds@cmu.edu

December 10, 2016

Abstract

This paper uses a natural experiment—the Super Bowl—to study the causal effect of advertising on demand for movies. Identification of the causal effect rests on two points: 1) Super Bowl ads are purchased before advertisers know which teams will play; 2) home cities of the teams that are playing will have proportionally more viewers than viewers in other cities. We find that the movies in our sample experience on average incremental opening weekend ticket sales of about \$8.4 million from a \$3 million Super Bowl advertisement.

[*]This paper benefited greatly from discussions with Randall Lewis, David Reiley, Bo Cowgill, Lawrence Katz, and Lawrence Summers. The referees and editors were particularly helpful. We also thank participants at IO Fest at Berkeley and the NBER Summer Institute for helpful comments.

1 Introduction

The United States spends roughly 2 percent of its GDP on advertising (Galbi [2008]). Not surprisingly, whether, when, and why advertising increases product demand is of considerable interest to economists and marketers. However, empirically measuring the impact of advertising is notoriously difficult. Products that are heavily advertised tend to sell more, but this in itself does not prove causation (Sherman and Tollison [1971], Comanor and Wilson [1971]). A particular product often sees an increase in sales after increasing its ad expenditures, but here too the causation could run the other way (Heyse and Wei [1985], Ackerberg [2003]). For example, flower companies increase ad expenditures in the weeks leading up to Valentine's Day and see increased sales around Valentine's Day. But it is not easy to determine the causal impact of that ad expenditure since many of the same factors that affect consumer demand may also affect advertising purchase decisions (Schmalensee [1978], Lee et al. [1996]).

Testing for causal effects requires an exogenous shock to ad exposures. The gold standard, as usual, is a randomized experiment. For this reason, field experiments have become increasingly popular among economists and marketers studying advertising (Simester et al. [2009], Bertrand et al. [2010], Lewis and Rao [2012]). However, these experiments tend to be expensive and require access to proprietary data. Moreover, they tend to have low power, often do not produce statistically significant effects, and have not led to consensus on advertising effectiveness (Hu et al. [2007], Lewis and Reiley [2008], Lewis and Rao [2012]).

Further, field experiments tend to involve a particular subset of ads: those that a firm is uncertain enough about to agree to conduct an experiment. These ads may be quite different from ads that are *routinely* purchased by firms. By contrast the differential viewership associated with the the Super Bowl and other sports events yields natural experiments that can be used to estimate advertising effectiveness.

Two weeks prior to the Super Bowl, the NFC and AFC Championship games are played. Controlling for the point spread, the winners of these games are essentially random. On average, the Super Bowl will be watched by an additional eight percentage points, or roughly 20 percent, more households, in the home cities of the teams that play in the game compared to other cities. There is a similar increase in viewership for the host city of the Super Bowl. We refer to these boosts in viewership as the "home-city" and "host-city" effects respectively.

Super Bowl ads are typically sold out several weeks or months before these Championship games, so firms have to decide whether to purchase ads before knowing who will be featured in the Super Bowl. Hence the outcomes of the Championship Games are essentially random shocks to the number of viewers of Super Bowl ads in the home cities of the winning teams. The increased sales of advertised products in cities of qualifying teams, compared to sales in home cities of near-qualifying teams, can thus be attributed to advertisements.

There are three attractive features to studying movies advertised in the Super Bowl. First, movie advertisements are common for Super Bowls, with an average of about 7 per game in our sample. Second, different movies advertise each year. Third, Super Bowl ad expenditure represents a large fraction of a movie's expected revenue. For a Pepsi ad to be profitable, it only needs to move sales by a very small amount. As Lewis and Rao [2012] show, in their Super Bowl Impossibility Theorem, for products like Pepsi, it can be virtually impossible to detect even profitable effects. The cost of Super Bowl ads, on the other hand, can represent a meaningful fraction of a movie's revenue.

There are however, two notable disadvantages to studying movies. First, city-specific, movie sales data are costly to obtain. Nonetheless, we were able to acquire this data for a limited sample of movies and cities. However, we also have an additional proxy for movie demand—Google searches after the Super Bowl. Miao and Ma [2015] and Panaligan and Chen [2013] have illustrated that Google searches are predictive of opening week revenue, and Google searches have the advantage of being available for the full sample of cities.

The second disadvantage of studying movies is that movies do not have a standard measure of expected demand *prior* to the broadcast of the Super Bowl ads. Here too Google searches can be helpful in that they can serve as a proxy for pre-existing interest in the movie and help improve the prediction of the outcome (box office or searches) when the movie opens.

Wesley Hartmann and Daniel Kapper proposed the idea of using the Super Bowl as a natural experiment at a presentation at the June 7-9, 2012 Marketing Science conference. They subsequently circulated a June 2012 working paper examining the impact of the Super Bowl ads on beer and soft drink sales. The most recent version of their working paper is Hartmann and Klapper [2015].

We independently came up with a similar idea in February of 2013. We focused on Super Bowl movie ads and thought of "fans" as an instrumental variable for ad exposures. Our initial analysis used Google queries for movie titles as the response variable, but eventually we were able to acquire movie revenue data by DMA. Earlier versions of Hartmann and Klapper [2015] and this paper were presented at the same session at the 2014 summer NBER meeting in Cambridge.

Both papers find a substantial effect of advertising on purchases in quite different markets. Beer and soft drinks involve substantial repeat purchases and have familiar brands. Movies are typically purchased only once and each is unique. Given these quite different characteristics, it is comforting that both papers find an economically and statistically significant impact of advertising on sales.

In a related paper, Ho et al. [2009] build an econometric model of exhibitors' decisions to show a movie, and consumers' decisions to view a movie during its opening weekend. The first stage equation models the probability of placing a Super Bowl ad for a movie as a function of the movie's budget, genre, rating, and distributor, whether the movie is released on a holiday week, and the timing of the ad relative to the movies release. Using this estimate, the authors construct expected expenditure on the Super Bowl ad. In the second stage regressions, they use the *predicted* expenditure as an explanatory variable for exhibitor decisions to show the movie, and for consumers' decisions to view the movies during the opening weekend.

Our model differs from the approach in Ho et al. [2009] in that we do not model the studios' decisions to purchase ads. It is possible (though in our opinion not likely) that astute theater chains recognize that the home cities of the Super Bowl teams will be exposed to more ads and thus be more likely to want to see the advertised movies. If this is so, then our model is about the joint impact of advertising on both consumer and exhibitor decisions. However, in our view the primary response is likely consumer decisions since exhibitors typically have to construct their distribution schedules months in advance. We expand on this point in Section 5.

Our work is also related to Yelkur et al. [2004] who analyze the effectiveness of Super Bowl advertising by comparing box office revenue for movies that were advertised on the Super Bowl to a set of popular movies that did not have Super Bowl advertisements. The authors find that, controlling for budget size and release date, movies with Super Bowl advertisements had nearly 40 percent higher gross theatrical revenue than other non-promoted movies. Of course, the movies that were selected to be advertised were likely chosen for some reason, so there could potentially be bias in this estimate due to confounding variables.

Overall, with our method we find strong evidence of large effects of advertising on movie demand. Our results suggest that a 100 ratings point increase due to additional Super Bowl ad impressions increases opening weekend movie revenue by 50–70 percent. For the average movie in our sample, this translates into an incremental return of at least \$8.4 million in opening weekend ticket sales associated with a \$3 million Super Bowl advertisement.

We believe that researchers can use this methodology for other types of advertising. Sports events such as the World Series, basketball playoffs, college bowls, the Olympics, and the World Cup create many large, essentially random shocks to viewership of ads shown during these events that can serve as natural experiments to measure ad impact.

2 Empirical specification

We use the following notation.

$$t = \text{date where outcome is measured (opening week)} \tag{1}$$

$$s = \text{date when ads are seen (Super Bowl)} \tag{2}$$

$$y_{mct} = \text{outcome for movie } m \text{ in city } c \text{ at time } t \tag{3}$$

$$x_{cms} = \text{adviews for movie } m \text{ in city } c \text{ at time } s \tag{4}$$

$$z_{cms} = \text{fans of team from city } c \text{ exposed to ad for movie } m \text{ at time } s \tag{5}$$

The variable `outcome` is the measure of ad performance, which in the initial specification is Google searches immediately prior to the opening weekend. Later we use opening weekend revenue for a subset of the movies advertised as our ad performance measure.

The `adviews` are the Nielsen ratings for the relevant Super Bowl. Nielsen ratings correspond to the percent of households watching the Super Bowl in an average half hour.

The `fans` variable in the initial specification consists of 3 dummy variables indicating whether the home team of the city in question is the AFC participant in the Super Bowl, whether the home team of the city in question is the NFC participant in the Super Bowl, and whether the city in question hosts the Super Bowl. Later on we investigate some refinements to this measure.

Our model specification is then a classic instrumental variable model.[1]

$$y_{cmt} = \alpha_0 + \alpha_1 x_{cms} + \epsilon_{cmt} \tag{6}$$

$$x_{cms} = \beta_0 + \beta_1 z_{cms} + \delta_{cms} \tag{7}$$

Equation (6) says that the outcome, y_{cmt}, depends on prior ad exposure, x_{cms}. We would not expect that estimating this single equation by ordinary least squares would produce a good estimate of the causal effect of advertising, since x_{cms} could be correlated with ϵ_{cmt}.

[1] We also include city and movie fixed effects along with an index of Google searches prior to the Super Bowl as control variables in our regressions.

There are a variety of ways that x_{cms} could be correlated with ϵ_{cmt}. For example, suppose that in some years, some cities are particularly interested in entertainment. These cities might watch the Super Bowl more than usual *and* attend movies more than usual. Or suppose different types of movies appealed to different geographic audiences. In this case, the teams that compete in the Super Bowl could affect the choice of movie advertised.

Another potential issue is measurement error. The city-level Nielsen ratings are based on a relatively small number of households. We would expect measurement error associated with the ratings numbers would attenuate the estimated effect of ad viewership on outcomes toward zero.

In order to estimate the causal impact of ad views on outcomes, we need an instrument—a variable that perturbs ad views exogenously.

Equation (7) contains such instruments, namely the home-city and the host-city effects we described earlier. We know from prior experience, and will verify in Section 4.1, that this instrument is a strong predictor of ad views. Furthermore, this instrument should be independent of ϵ_{cmt} since advertising expenditures typically are chosen well before it is known which teams will play in the Super Bowl. We present additional arguments for identification in Section 5.

3 Data

3.1 Ad views

We measured ad views using Nielsen ratings for the 2004-2014 Super Bowls, for 56 designated media markets (cities) from Street & Smith's *Sports Business Daily Global Journal*. Total local ad spend, which we use in Section 5.4, is taken from Kantar Media. This data is only available starting in 2009.

3.2 Movies

We looked at a sample of 70 movies that were advertised in the Super Bowl and were released within 6 months after the game date. The average gap between the Super Bowl and the movie release was about 66 days and the median was 54 days. The gap varied quite a bit, with a standard deviation of about 50 days. Roughly speaking, the median date of release was mid-March, but there is substantial variation in the release date.

We obtained the list of movies that advertised for the Super Bowl from the *USA Today's* AdMeter, which lists commercials and viewer ratings for all commercials after every Super Bowl. Release dates, distributor, budget, and national sales by week for every movie were found at `the-numbers.com`. Data on movie opening weekend sales is from Rentrak.

3.3 Fans and Host City

As indicated above, the simplest proxy for fans of a team in a city is just a dummy variable that equals 1 if the team plays in the home city and 0 otherwise. We split the fans into AFC fans and NFC fans. We also add the host city in some specifications. Though the host city is known in advance, we argue in Section 5 that it represents such a small part of the total boost in viewership that it is unlikely to have a meaningful impact on advertiser choices. The advantage of including the host city is we get more power. However, the quantitative results are similar with and without host city, suggesting advertisers do not select ads considering which city is hosting the game.

To test the sensitivity of our results to alternate specifications, in Section 6.1 we refine the definition of fans using Google searches, and in Section 5.3 we adjusted the fans measure using Vegas odds in the playoffs so as to reflect the estimated fans at the time of the playoffs.

3.4 Searches

Movie titles frequently contain common words, making it difficult to use simple text matching to identify queries related to movies. For example, the word [wolverine] could refer to an animal, a university mascot, a brand of boots, or a Marvel comics character.

We address this problem by using the Google entity identifier associated with the movies in our sample. Google's entity identifier attempts to disambiguate different uses of a word by using contextual information associated with the search. So if a user searched for other animals in the session where a search for [wolverine] occurred, that user is likely looking for information about the animal. On the other hand, if a user included movie related terms along with a search for [wolverine] it is likely that they were using the word as short-hand for the movie *X-Men Origins: Wolverine*.

With the Google entity identifier, we generate a control variable in our regressions based on the Google Trends index prior to the Super Bowl for each city and movie in our data. The Google Trends index for the week preceding the opening weekend was used as an outcome variable in the initial specification. We interpret this index as a measure of "interest" in a movie. The Google Trends data has the advantage of being complete—available for all movies in the sample—and non-proprietary.[2] By contrast, the Rentrak data on opening weekend revenue is available only for a subset of movies and is proprietary and cannot be freely redistributed.

In addition to the Google Trends index of searches on the movie prior to the Super Bowl, we also use city and movie fixed effects.

We also confirm that a movie's opening weekend box office sales can be well-predicted by a few key features. In particular, we regress box office sales per capita on searches prior to the Super Bowl, the type of movie (comedy,

[2]The number of queries in a given city must be larger than an unspecified privacy threshold to show up in the index, so there are a few smaller cities that report zero searches on movie entities prior to the Super Bowl. We drop these cities from the analysis.

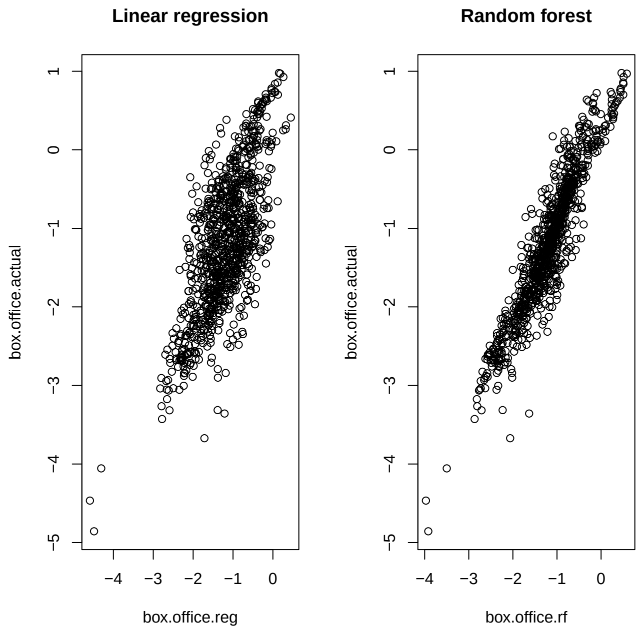

Figure 1: This shows how opening box office sales per capita compare to a prediction using a few features: Google searches prior to the Super Bowl, the type of movie (comedy, adventure, etc.), the distributor, the rating, and the DMA. We use both an ordinary linear regression and a random forest model.

adventure, etc.), the distributor, the rating, and the DMA. We use both an ordinary linear regression and a random forest model. The OLS prediction has an R^2 of 0.51 and the random forest has an R^2 of 0.87. Figure 3.4 shows how these two predictions compare to the actual box office sales.

4 Results for Google Searches

4.1 First stage

This section examines the effects of advertising on Google searches in the week prior to opening weekend. Table 1 shows that the Nielsen ratings in a given city are strongly related to whether that city is a home city for one of the teams playing or the host city for the game.

Table 1: First Stage: Super Bowl Ratings and Fans of Teams

	Nielsen Ratings	
	(1)	(2)
City of AFC Championship Game Winner		0.077***
		(0.009)
City of NFC Championship Game Winner		0.076***
		(0.008)
Super Bowl Host City		0.063***
		(0.008)
Constant	0.455***	0.451***
	(0.004)	(0.003)
Adjusted R-squared	0.66	0.75
Observations	616	616

* $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

Notes: Robust standard errors clustered at the city-year level are shown in parentheses. City and year fixed effects are included in all specifications. Nielsen ratings correspond to the percent of households watching the Super Bowl in an average half hour. Home city is a dummy variable that takes the value 1 if a team plays in a city; 0 otherwise. The Green Bay Packers' Home city is Milwaukee, since we do not have ratings data on Green Bay. Data sources are discussed in more detail in Section 3.

Column (1) of Table 1 shows the R^2 for the regression that only uses city and movie fixed effects. Column (2) shows what happens to R^2 when

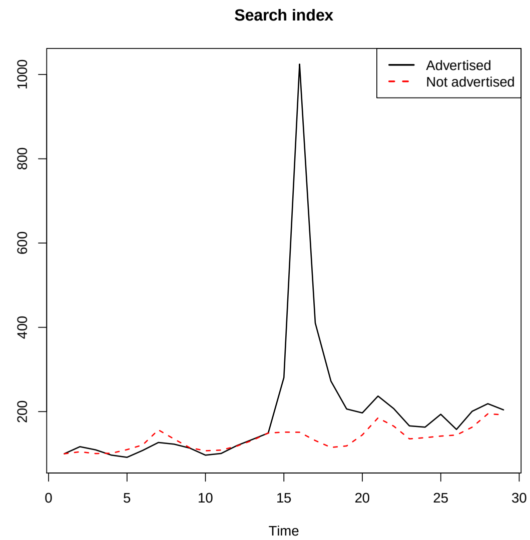

Figure 2: Nationwide searches for movies advertised during the Super Bowl and similar movies that were not advertised during the Super Bowl.

we include dummy variables for the teams that are playing and the host city. The R^2 moves from 66 percent to 75 percent, indicating that these instruments significantly improve the prediction of Nielsen ratings.

As the regression shows, about 8 percentage points more households will watch the Super Bowl in the home city of qualifying teams. This is about a 20 percent increase in ratings compared to the sample average.

4.2 Second stage

Figure 2 shows the nationwide queries on movie titles advertised in the Super Bowl.

It is clear that movies advertised in the Super Bowl see a significant bump in searches. We also contrast these searches with national search volume for a set of placebo movies that had similar qualities to the advertising movies but did not advertise in the Super Bowl. We discuss how we select these movies in Section 6.3.

While it is clear there is an increase in interest in advertising movies immediately after the ads are shown, it is not apparent how much of that initial interest translates into box office revenue. That question is what our model is designed to answer.

The regression results in Table 2 use an ordinary least squares regression in Column (1) to show that, for movies that advertised in the Super Bowl, Google searches on release week are notably higher in cities with higher Super Bowl ratings than in other cities. Note that Google Trends numbers for the search volume in a particular geo are measured relative to the total number of searches in that geo. Hence the Trends numbers are already normalized for population size.

Column (2) uses both home and host cities as instruments and finds about twice as large an effect as the OLS estimate. Our baseline model uses host cities as an instrument but Table 5 shows the estimated effect is similar if we use only home cities.

Table 2: Effects of Advertising

	log(Google Searches on Release Week)		log(Box Office PC)	
	(1)	(2)	(3)	(4)
Nielsen Ratings	0.314	0.762**	0.484**	0.771**
	(0.243)	(0.318)	(0.225)	(0.362)
log(pre Search)	0.068***	0.069***	0.035***	0.035***
	(0.018)	(0.017)	(0.013)	(0.012)
Adj. R-squared	0.89	0.89	0.96	0.96
Observations	3,080	3,080	1,088	1,088
Specification	OLS	2SLS	OLS	2SLS

* $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

Notes: Robust standard errors clustered at the city-year level are shown in parentheses. City and year fixed effects are included in all specifications. Super Bowl ratings are Nielsen ratings, corresponding to percent of households watching the Super Bowl in an average half hour. Instruments include dummy variables for the home and host cities. Data sources are discussed in more detail in Section 3.

5 Identification

In this section we consider arguments questioning the validity of the `fans` instrument and present rebuttals to these arguments.

We note that there could be a potential problem in our estimates above if East Coast and Midwest football fans liked different kinds of movies. In such a case, the movie that a studio chooses to advertise could, in principle, depend on which teams play in the Super Bowl. In our view, this is conceivable, but not likely.

The reason this shouldn't impact our estimates is that the decisions about which movies to promote and how much to spend on promotion are made at a *national* level. This means that variations in attendance will be determined primarily by local tastes. The only role that advertiser decisions might make is in determining which movies to advertise nationwide. This will typically not depend on which teams end up playing since the choice of which movies to advertise 1) is made well in advance and 2) has a tiny impact on the total size of the audience, as we establish below.

5.1 Ad decisions are made in advance

The decision to show an ad in the Super Bowl is typically made far in advance of the actual game, when advertisers would have little idea which teams would play. (They would know the host city, which we deal with shortly.) Table 3 presents a list of press reports about the status of Super Bowl ad sales. (We report short URLs for reasons of space; complete URLs are provided in a spreadsheet in the Appendix.) Of course most advertisers do not wait until the last minute to purchase ads. According to our discussions with film industry executives, the decision about which movies to advertise in the Super Bowl are decided well in advance of the game. Generally studios only have a few choices of movies that will be released in an appropriate time frame, and a great deal of care goes into planning and executing marketing for the hoped-for blockbusters.

Table 3: Ad sales for Super Bowl.

Year	Snippet	Date	Source
2003	fewer than 10 spots available	Jan 06 2003	superbowl.ads.com
2004	–NA–	–NA–	–NA–
2005	said Thursday all 59 slots had been sold	Feb 02 2005	money.cnn.com
2006	80% sold	Dec 18 2005	www.mediapost.com
2007	first half sold out	Jan 03 2007	money.cnn.com
2008	90% sold out by first week in Nov	Nov 07 2007	money.cnn.com
2009	much was sold out by September	Jan 09 2008	money.cnn.com
2010	had finished selling commercial time	Feb 01 2010	articles.latimes.com
2011	3 months before	Oct 29 2010	adage.com
2012	has sold out	Jan 02 2012	www.bloomberg.com
2013	advertisers need to announce 5 months out	Sep 03 2013	www.usatoday.com

Notes: The columns show the Super Bowl year and extracts from news articles that appeared on the indicated date from the indicated source. The full URL for these snippets is available in the online Appendix.

5.2 Home-city and host-city effects are small

It may well be that studios would advertise different movies in different geographies if they were able to do so, but in this case there is a single nationwide audience and advertisers must choose one movie for the entire audience. This restriction makes it implausible that the host cities and home cities of the teams playing in the Super Bowl would have any impact on advertising decisions since the *aggregate* audience for the ad is not very sensitive to which teams actually play and where they play.

To see this, we constructed an estimate of what *would have happened* to viewership if the teams that lost the championship games instead won those games and competed in the Super Bowl.

Consider for example, Pittsburgh's 2005 loss. This meant that 161,000 fewer households watched the ad in Pittsburgh than would have watched had Pittsburgh won. However, compared to the total viewership for the Super Bowl that year of 86 million this is only 0.2 percent, a tiny factor in an advertiser's decision.

The largest city in our sample is New York, but even in this case, the impact of the counterfactual is only 1.2 percent. Nationwide the average absolute difference in viewers across all DMAs and years was 0.4 percent of national viewership. Would the choice of ad to be shown in the Super Bowl depend on a 20 percent boost in viewership for 0.4 percent of the population? We believe that this effect is insignificant from an economic viewpoint and unlikely to affect studio decisions.

A similar argument applies to the host cities which are known in advance. However, the population of the host cities comprise only 1.6 percent on average of the DMAs in our sample. It seems implausible that choosing which movie ad to show nationwide in the Super Bowl would be influenced by a 0.2 percent boost in viewership (1.6 percent of the population times a 15 percent boost).

5.3 Expected fans

Even though advertisers do not know with certainty who will play in the Super Bowl game, they can form judgments about who will play. Our contacts in the movie business tell us that decisions on which movies to advertise are made far in advance of the playoffs, and they would be highly unlikely to substitute at the last minute based on which teams were playing due to the major investments they have made in planning, publicity, and production of the movie ad. Furthermore, as we have seen, the effect on viewership of the movie ad is tiny.

Nevertheless, let us take this critique seriously and see how plausible it is. Consider the Vegas odds for the AFC and NFC Championship games.[3] We converted these odds to probabilities using the method described in Stern [1986] and calculated the expected fans for each city, where the expectation is made using the Vegas odds just prior to the championship game. We then used the expected fans as control variables in the regressions described earlier.

We did not use the host city as an instrument since we thought that if advertisers were so sophisticated that they considered expected fans in their decisions, they would certainly take into account the host city in those decisions, which would make the host city an invalid instrument. The expected fans specification made no essential difference in the results.

Let us summarize the argument. In our baseline specification, the instrument is whether a city's team qualified for the Super Bowl. If advertisers were highly sophisticated and picked advertisements based on which teams were performing well up to the point they chose to advertise this could be a biased instrument. By controlling for the probability a team makes it to

[3] These are available at `http://www.vegasinsider.com/nfl/afc-championship/history/`.

the Super Bowl at the time of the Championship games, we ensure that our instrument is "as good as random."

5.4 Impact of outcome on subsequent ad spend

If advertisers choose their subsequent ad spend on a movie based on the associated Super Bowl ratings, our instrument would not be valid. To check this possibility we ran a regression to see if local ad spend was associated with home and host cities. Our data on local ad spend is from Kantar Media, and these data were only available to us starting in 2009.

The results of this regression are shown in Column 2 of Table 4. These estimates should be compared to those in Column 1 which is the first-stage regression from Table 1 but restricted to data from 2009 onward. Both dependent variables, the Nielsen ratings and ad spend per capita, are expressed in logs. Hence, the regression coefficients can be interpreted as percentage response. The impact of host and home cities on Nielsen ratings is large and statistically significant while the corresponding coefficients for local ad spend are small and statistically insignificant.

6 Variations on the baseline model

Here we consider a few variations on the baseline model.

Table 4: Local ad spend compared to Nielsen ratings

	log(Nielsen Ratings)	log(Ad Spend PC+1)
	(1)	(2)
City of AFC Championship Game Winner	0.107***	-0.003
	(0.021)	(0.008)
City of NFC Championship Game Winner	0.133***	-0.011
	(0.020)	(0.008)
Super Bowl Host City	0.093***	0.005
	(0.020)	(0.008)
Adjusted R-squared	0.80	0.58
Observations	336	336
Fixed Effects	City and Year	City and Year

* $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

Notes: These regressions include only observations for 2009 onward due to data availability. For each year and city, we add up local television spending across all Super Bowl movies. This gives one observation for each year and city, making the data directly comparable to Nielsen ratings data. There are a small number of zeros in local ad spend for a few small cities and niche movies, which is why we took log of `adspend` + 1. Note that these cities may well have seen some movie ads through *national* advertising campaigns.

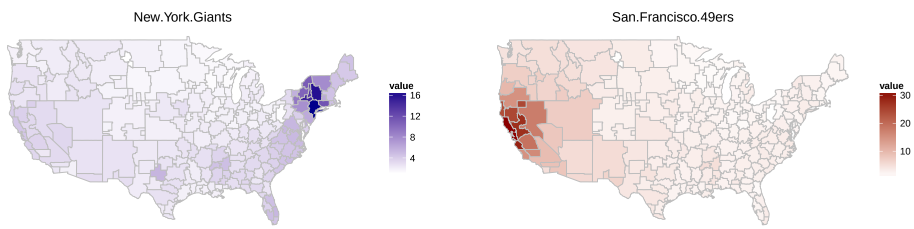

Figure 3: Heat map of estimated fan density for New York Giants and San Francisco 49ers using method described in text.

6.1 Other definitions of fans

In our baseline model we use dummy variables for the home cities of the two participating teams. However, some major cities do not have an NFL team, but football fans in those cities may identify with teams from other cities. We use Google entity search data from Google Trends in each NFL city for each NFL team to measure the local interest in that team. See Figure 3 which shows the distribution of searches for the New York Giants and the San Francisco 49ers. The geographic pattern suggests that this is a plausible measure for the fan distribution. Our results using this definition of fans are shown in Column (2) of Table 5.

Table 5: Variations on baseline model for opening week searches

	log(Google Searches on Release Week)				
	(1)	(2)	(3)	(4)	(5)
Nielsen ratings	0.762**	0.684**	0.687*	0.705	0.721**
	(0.318)	(0.333)	(0.355)	(0.620)	(0.360)
log(Pre search)	0.069***	0.069***	0.069***	0.069***	0.078***
	(0.017)	(0.017)	(0.017)	(0.017)	(0.017)
Adj. R^2	0.89	0.89	0.89	0.89	0.92
Observations	3,080	3,080	3,080	3,080	3,080
Specification	+Host	Trends	−Host	+Vegas	Weighted

* $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

Notes: Robust standard errors clustered at the city-year level are shown in parentheses. City and movie fixed effects are included in all specifications. Nielsen ratings correspond to the percent of households watching the Super Bowl in an average half hour. Column (1) uses home and host cities as instruments, Column (2) uses the Google Trends data to measure fans, Column (3) omits the host variable, Column (4) uses the expected fans measure based on Vegas odds, Column (5) uses the original specification with population weighting. Data sources are discussed in more detail in Section 3.

6.2 Opening weekend box office revenue

As mentioned above, we have two measures of outcome: Google searches on the movie title and opening weekend revenue.

The movie sales data we have is only available for a subset of cities. In particular, we only have data for movies that advertised in the Super Bowl and cities that were the home cities for teams that qualified for a Super Bowl or were the runners-up.

Despite the smaller sample, there is evidence of a significant positive effect of Super Bowl ratings on movie sales as shown in Table 2, Columns (3) and (4). Note, though, that the effect on ticket sales is smaller than the effect on Google searches. This is true even if we use only the sub-sample of cities for which we have box office data. Table 6 reports regressions using the alternate definition of fans.

6.3 Placebo analysis

It is conceivable that Super Bowl ratings could influence subsequent movie attendance for all movies. We consider this possibility highly implausible, but decided to check it anyway.

One could look at city-by-city movie attendance following the Super Bowl, but a better test is to look at movies that were *similar* to those advertised in the Super Bowl. Accordingly, we constructed a placebo set of movies. If watching the Super Bowl is correlated with subsequent overall movie atten-

Table 6: Variations on baseline model for opening week box office

	log(Box Office PC)				
	(1)	(2)	(3)	(4)	(5)
Nielsen Ratings	0.771**	0.705**	0.507	1.401***	0.444
	(0.362)	(0.342)	(0.352)	(0.527)	(0.283)
log(Pre Search)	0.035***	0.035***	0.035***	0.038***	0.055***
	(0.012)	(0.012)	(0.012)	(0.012)	(0.016)
Adj R^2	0.96	0.96	0.96	0.96	0.97
Observations	1,088	1,088	1,088	1,088	1,088
Specification	+Host	Trends	−Host	+Vegas	Weighted

* $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

Notes: Robust standard errors clustered at the city-year level are shown in parentheses. City and movie fixed effects are included in all specifications. Nielsen ratings correspond to the percent of households watching the Super Bowl in an average half hour. Column (1) uses home and host cities as instruments, Column (2) uses Google Trends data to measure fans, Column (3) omits the host variable, Column (4) uses the expected fans measure based on Vegas odds, Column (5) uses the original specification with population weighting. Data sources are discussed in more detail in Section 3.

dance, we would expect to see it affect both those movies that were advertised and similar movies that weren't advertised.

Specifically, we used nearest-neighbor matching based on the movie budget, movie category (comedy, action, etc.), distributor, critic ratings, and year and month of release. We used the `matchit R` package which is specifically designed for this purpose and described in detail in Ho et al. [2007a,b]. We provide lists of the advertised and matched movies in the online appendix. In our view, these two lists appear to be similar.

The results are shown in Table 7 for our baseline specification and a few of the variations considered above. What is noteworthy is that the coefficient on `Nielsen ratings` is insignificant for all specifications. Of course, the movies advertised in the Super Bowl were chosen for that distinction and our matching is far from perfect, so this analysis cannot be considered definitive evidence. Nevertheless, it is suggestive.

We can test to see whether the estimated coefficient on ad views (Nielsen ratings) is different for the advertised and placebo movies. To do this we combine the two datasets and add an interaction term for Nielsen ratings and the advertised movies. This is denoted by `Nielsen` $\times$ `Super Ad` in Table 8. The interaction effect is significant at the 10 percent level in our baseline specification (Column 2) and at the 5 percent level when we use the Google Trends measure for fans (Column 3).[4]

[4] Another question is whether placebo movies do worse than they would have if the Super Bowl ads had not run. That is, does advertising for Super Bowl movies cause substitution away from placebo movies? The relevant coefficient to test this is the first one in Table 8, `Nielsen Super Bowl Ratings`. Unfortunately, we get different answers

Table 7: Effects of Advertising: Placebo movies

	log(Google Searches on Release Week)			
	(1)	(2)	(3)	(4)
Nielsen Ratings	-0.091	-0.373	0.059	0.198
	(0.374)	(0.387)	(0.444)	(0.876)
log(Pre-Super Search)	0.083***	0.083***	0.084***	0.084***
	(0.019)	(0.018)	(0.019)	(0.019)
Adjusted R-squared	0.87	0.87	0.87	0.87
Observations	2,747	2,747	2,747	2,747
Specification	2SLS	2SLS (Trends fans)	2SLS (-Host)	2SLS (+Vegas)

* $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

Notes: Column (1) shows the baseline IV estimates from Table 2 using the placebo data. Columns (2)-(4) illustrate variations on the baseline model that we consider elsewhere in the paper, such as other definition of fans (Section 6.1), excluding the host city as an instrument, and using Vegas odds to compute expected fans (Section 5.3). The notes from Table 2 apply here as well.

Table 8: Placebo and advertised movies

	log(Google Searches on Release Week)			
	(1)	(2)	(3)	(4)
Nielsen Super Bowl Ratings	-0.483**	-0.091	-0.373	0.059
	(0.238)	(0.374)	(0.387)	(0.444)
Nielsen X Super Ad	0.797***	0.853*	1.057**	0.628
	(0.305)	(0.448)	(0.461)	(0.483)
log(Pre-Super Search)	0.083***	0.083***	0.083***	0.084***
	(0.019)	(0.019)	(0.018)	(0.019)
Adjusted R-squared	0.88	0.88	0.88	0.88
Observations	5,827	5,827	5,827	5,827
Specification	OLS	2SLS	2SLS (Trends fans)	2SLS (-Host)

* $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

Notes: City and movie fixed effects are used in all specifications. See the notes to the previous table for definitions. Coefficients for other specifications are available in the online appendix.

6.4 Interpretation

The results suggest that an increase of 100 ratings points raises weekend ticket sales for a movie advertised on the Super Bowl by at least 50 percent. Note that 100 ratings points means a switch from 0 percent of people watching to 100 percent of people watching. In other words, it measures the difference from a hypothetical situation in which everybody watched the ad to a hypothetical situation in which nobody watched the ad.

Since the Super Bowl averages about 42 ratings points overall, this implies that a Super Bowl ad increases release-week ticket sales by about 21 percent. In other words, the coefficient suggests there are 21 percent more ticket sales when 42 percent of the country watched the Super Bowl than there would have been if nobody watched the Super Bowl. The average movie in our sample took in $40 million on the opening weekend. Thus the incremental ticket revenue from the Super Bowl ad were roughly $8.4 million on average. Since a Super Bowl ad cost is about $3 million, this means an overall return of 2.8 to 1.

According to industry practice, the studio typically pays for the entire marketing costs and receives 40-50 percent of the domestic box office revenue. (The exact numbers are closely guarded secrets, but see Danzig and Hughes [2014] for some estimates.) Hence the return to the studio from the Super Bowl ad is about 1.4 to 1, or a 40 percent ROI. [5]

---

depending on the specification. It is usually negative – suggesting there is substitution – but only statistically significant in one out of four main specifications.

[5]Hartmann and Klapper [2015] estimate a 153 percent ROI for Super Bowl beer ads,

We want to emphasize four caveats in interpreting these results.

First, this back of the envelope calculation ignores future revenue streams such as ticket sales after the opening weekend and other revenue through home movie purchases, TV licensing, and so on. Some of this additional revenue may be attributable to the Super Bowl ad impressions, though we have no easy way to measure this.

However, a causal relationship between increased movie attendance and increased home entertainment sales is consistent with Choi et al. [2015] who use opening-weekend snowstorms as an instrument and find that a 10 percent rise in theatrical attendance causes an 8 percent increase in DVDs/Blu-ray sales when they are released. Cable licensing deals are also directly tied to box office success so that any increase in box office revenue will positively impact revenue from this channel.

We also do not know how the incremental revenue is divided among the various parties—how much goes to the studios, producers, writers, stars, and so on. Similarly, we don't know exactly how the costs of the Super Bowl ad are divided among the various parties. However, as indicated above, it appears that studios are the primary decision makers with respect to Super Bowl ads and bear most of the marketing costs.

Second, in calculating the return to advertising, we are assuming that the *incremental* viewers of the Super Bowl have the same response to ads as those who would watch the Super Bowl anyway. It is possible that the committed

but caution that this is a likely an overestimate.

fans pay more attention to the game and less to ads. Or perhaps they are much more engaged with the entire experience and so pay more attention to ads than the incremental viewers. It is also possible that the incremental fans have substantially different tastes in movies than the fans you would get simply by purchasing more ad slots. We provide some evidence on this in Section 7.

Third, we don't know how these results extend to other settings, as the Super Bowl has unique qualities. There are other similar events such as the World Series, basketball playoffs, the Summer and Winter Olympics, and so on. These natural experiments are not quite as clean-cut as the Super Bowl, but are certainly worthy of future study.

Fourth, one might ask why the estimated return is so high. First, it is important to understand that our results pertain to returns on movies that the studio has *chosen* to advertise on the Super Bowl. The return on advertising movies with mediocre prospects could be much lower. Second, once the network has set a market-clearing price, we would expect that the marginal ad would earn a normal, risk-adjusted rate of return. However, the *average* ad would typically earn a return higher than the marginal ad. One might then ask "if the return to the movie ad is so high, why don't the studios advertise more movies?" The answer to this question may be that they only have a few movies for which a Super Bowl ad makes economic sense. Movie theaters can only show a limited number of movies at any one time, and the conventional wisdom in the industry is that if two blockbusters are released on the same weekend, the revenues of both movies will suffer. As a result, studios typically try to stagger the release of blockbusters, so at any one time there are only a few movies that would warrant Super Bowl treatment. Whatever the explanation, we typically see only 6-8 movie ads per Super Bowl and this number does not vary much from year to year.

Finally, we want to clarify how these results fit with the Super Bowl Impossibility Theorem (Lewis and Rao [2012]). They argue that it is nearly impossible for a firm to test the effects of an individual ad campaign, even if it randomly assigned DMAs during a Super Bowl. How, then, can we find such highly statistically significant results? The answer is that the Super Bowl Impossibility Theorem refers to the question of measuring the effectiveness of a *single* campaign. But here, we study the average effect of 70 campaigns. The noise level is too high to say anything about the effects of a particular advertisement, but the average performance of all movies in our sample can be estimated reasonably precisely.

7 Heterogeneous treatment effects

We have shown that the incremental ad exposures due to the home-team effect have a causal impact on both Google queries and opening weekend revenue. This suggest that increased ad expenditure would also have an incremental impact on these outcomes. However, the incremental ad views from the home-team effect may well be different than the incremental ad views from simply spending more money on advertising.

We can offer some suggestive evidence on this point. We ran a Google Consumer Survey and asked the 2,568 respondents whether they watched the Super Bowl on TV in 2013, 2014 or both years. The question of interest was whether those who watched both years were different than those who watched only one year. The dimensions on which the respondents could differ were inferred age, gender, and income.[6]

We found that those who watched the Super Bowl in both years, rather than a single year, tended to be older, more male, and live in wealthier areas. However, most of these effects tended to be statistically insignificant, with the exception of gender. We suspect that there is *some* difference between the incremental viewers from the home-city effect and the incremental viewers that would be reached by increased ad spend.

Nevertheless, we believe that our estimates can be useful in estimating the response to ad spend. Suppose that a movie advertiser targeted its ads to reflect the audience composition of the *incremental* Super Bowl viewers. This targeting could be informed by a more sophisticated version of our survey. That advertiser might well expect a response to its ad spend along the lines of that described in Section 6.2. So those estimates of the impact of spend on box office should be a *lower bound* on what ad effectiveness would be if ad targeting could be fully optimized.

[6]Inferred age and gender are based on web site visits and inferred income is based on the IP address of the respondent and Census data.

We also can test whether there are differential effects based on when a movie is released. Are ads less effective for movies released well after the Super Bowl? We divided our sample into movies with release dates more than 70 days out and those with release dates less than 70 days out. We recreated the regressions in Table 2. Somewhat surprisingly, we did not see a difference in the effects of ads on box office sales in these two groups.[7]

8 Discussion

We use a natural experiment—the Super Bowl—to study the causal effect of advertising on movie demand. Our identification strategy relies on the fact that Super Bowl ads are purchased before advertisers know which teams will play in the Super Bowl and that cities where there are many fans of the qualifying teams have substantially larger viewership than other cities do.

Within this setting we study 70 movies that were advertised during the 2004-2014 Super Bowls. We compare product purchase patterns for advertised movies in cities with fans from the qualifying teams to cities with fans of near-qualifying teams. We find a substantial increase in opening weekend revenue due to Super Bowl advertisements. On average, the movies in our sample experience an incremental increase of $8.4 million in opening weekend

[7] In general, we don't have sufficient power to break down the treatment effects. There are several other interesting questions, such as whether there are differential effects for movies with more competition, but we have to leave these questions for further research. It may be possible to investigate such issues after we accumulate a few more years of Super Bowl data.

box office revenue from a \$3 million Super Bowl advertisement.

We suggest that our methodology can be generalized to a variety of sports settings where the nature of qualifying creates a large random shock to ad viewership in a particular area, and that this methodology has notable advantages compared to the more common approach of using field experiments to determine the causal impact of advertising. The best identification comes from sporting events such as the Super Bowl in which the teams that will play are unknown at the time companies purchase advertising spot. However, even if the home cities are known it seems to us unlikely that advertisers would take this information into account when choosing its ad expenditure. So the methodology could well be applicable for a broader set of media broadcasts with differential appeal across geographies.

References

Daniel A. Ackerberg. Advertising, Learning, and Consumer Choice in Experience Good Markets: an Empirical Examination. *International Economic Review*, 44(3):1007–1040, August 2003. ISSN 0020-6598. doi: 10.1111/1468-2354.t01-2-00098. URL `http://doi.wiley.com/10.1111/1468-2354.t01-2-00098`.

Marianne Bertrand, Dean Karlan, Sendhil Mullainathan, Eldar Shafir, and Jonathan Zinman. What's Advertising Content Worth? Evidence from a Consumer Credit Marketing Field Experiment. *The Quarterly Journal of Economics*, 125(1):44, 2010. URL `http://www.ingentaconnect.com.ezp-prod1.hul.harvard.edu/content/oup/qje/2010/00000125/00000001/art00007`.

Patrick Choi, Peter Boatwright, and Michael D. Smith. The perfect storm: Using snowstorms to analyze the effect of theatrical attendance on demand for subsequently released DVDs. Technical report, Carnegie Mellon University, 2015. URL `http://ssrn.com/abstract=2639303`.

William S Comanor and Thomas A Wilson. On Advertising and Profitability. *The Review of Economics and Statistics*, 53(4):408–10, 1971. URL `http://econpapers.repec.org/RePEc:tpr:restat:v:53:y:1971:i:4:p:408-10`.

Scott Danzig and Mark Hughes. Breakdown of movie costs, 2014. URL `http://www.quora.com/What-is-the-breakdown-of-costs-associated-with-making-a-high-budget-Hollywood-f`

Douglas Galbi. U.S. Annual Advertising Spending Since 1919, 2008. URL `http://www.galbithink.org/ad-spending.htm`.

Wesley R. Hartmann and Daniel Klapper. Super Bowl ads. Technical report, Stanford Graduate School of Business, 2015. URL `http://faculty-gsb.stanford.edu/hartmann/SuperBowl.pdf`. First version presented at INFORMS Marketing Science, June 7-9, 2012.

Joseph F. Heyse and William W. S. Wei. Modelling the Advertising-Sales Relationship through Use of Multiple Time Series Techniques. *Journal of Forecasting*, 4(2):165–181, 1985. ISSN 02776693. doi: 10.1002/for.3980040206. URL `http://doi.wiley.com/10.1002/for.3980040206`.

Daniel Ho, Kosuke Imai, Gary King, and Elizabeth Stuart. Matching as nonparametric preprocessing for reducing model dependence in parametric causal inference. *Political Analysis*, 2007a. URL http://gking.harvard.edu/files/abs/matchp-abs.shtml.

Daniel Ho, Kosuke Imai, Gary King, and Elizabeth Stuart. Matchit: Nonparametric preprocessing for parametric causal inference. *Journal of Statistical Software*, 15(3):199–236, 2007b. URL http://gking.harvard.edu/matchit/.

Jason Ho, Tirtha Dhar, and Charles Weinberg. Playoff payoff: Super Bowl advertising for movies. *International Journal of Research in Marketing*, 26:168–179, 2009. URL www.elsevier.com/locate/ijresmar.

Ye Hu, Leonard M. Lodish, and Abba A. Krieger. An Analysis of Real World TV Advertising Tests: a 15-Year Update. *Journal of Advertising Research*, 47(3):341–353, 2007.

Junsoo Lee, B. S. Shin, and In Chung. Causality Between Advertising and Sales: New Evidence from Cointegration. *Applied Economics Letters*, 3(5): 299–301, 1996. ISSN 1350-4851. URL http://econpapers.repec.org/RePEc:taf:apeclt:v:3:y:1996:i:5:p:299-301.

Randall A. Lewis and Justin M. Rao. On the Near Impossibility of Measuring Advertising Effectiveness. *Working Paper*, 2012.

Randall Aaron Lewis and David H. Reiley. Does Retail Advertising Work? Measuring the Effects of Advertising on Sales Via a Controlled Experiment on Yahoo! *SSRN Electronic Journal*, June 2008. ISSN 1556-5068. doi: 10.2139/ssrn.1865943. URL http://papers.ssrn.com.ezp-prod1.hul.harvard.edu/abstract=1865943.

Rui Miao and Yueyue Ma. The dynamic impact of websearch volume on product — an empirical study based on box office revenues. *WHICEB 2015 Proceedings*, 2015. URL http://aisel.aisnet.org/whiceb2015/14.

Reggie Panaligan and Andrea Chen. Quantifying movie magic with google search. *Google Think Insights*, 2013. URL http://ssl.gstatic.com/think/docs/quantifying-movie-magic_research-studies.pdf.

Richard Schmalensee. A Model of Advertising and Product Quality. *Journal of Political Economy*, 86(3):485–503, 1978. URL http://ideas.repec.org/a/ucp/jpolec/v86y1978i3p485-503.html.

Roger Sherman and Robert D. Tollison. Advertising and Profitability. *The Review of Economics and Statistics*, 53(4):397–407, 1971. URL http://econpapers.repec.org/RePEc:tpr:restat:v:53:y:1971:i:4:p:397-407.

Duncan Simester, Yu Hu, Erik Brynjolfsson, and Eric T. Anderson. Dynamics of Retail Advertising: Evidence from a Field Experiment. *Economic Inquiry*, 47(3):482–499, 2009. ISSN 0095-2583. URL http://econpapers.repec.org/RePEc:bla:ecinqu:v:47:y:2009:i:3:p:482-499.

Hal Stern. The probability of winning a football game as a function of the pointspread. Technical report, Department of Statistics, Stanford University, 1986.

Rama Yelkur, CHuck Tomkovick, and Patty Traczyk. Super Bowl advertising effectiveness: Hollywood finds the games golden. *Journal of Advertising Research*, 44(1):143–159, 2004. URL http://journals.cambridge.org/abstract_S0021849904040085.