THE JOURNAL OF INDUSTRIAL ECONOMICS 0022-1821
Volume LXIV June 2016 No. 2

BOX-OFFICE DEMAND: THE IMPORTANCE OF BEING #1*

Luís Cabral†
Gabriel Natividad‡

We propose a theoretical framework to understand the effect on a movie's eventual theatrical success of leading the box office during the opening weekend. We consider two possible channels: a positive shock to the utility from watching the movie and a greater awareness of the movie's existence. We derive a series of testable predictions, which we test on U.S. box office data. The results suggest that being #1 in sales during the opening weekend has an economically and statistically significant effect on the movie's total demand; and that the primary channel for this effect is through the greater awareness induced by being #1.

I. INTRODUCTION

How easy is it to Predict Movie Demand? In reference to this question, screenwriter William Goldman once famously quipped that 'nobody knows anything' (Goldman [1983]). One thing industry participants do know, however: winning the first competitive battle at the box office — the very first weekend of a film's theatrical life — can be a strong predictor of a movie's eventual success.

In this paper, we propose a theoretical framework to understand the relation between being the overall winner in box office revenues during a film's opening weekend and the film's subsequent economic success. We consider two possible channels. A first one is that being anointed as box office winner implies a positive shock to consumer utility for watching that movie: for example, being #1 might work as a coordination device for moviegoers with a strong social consumption motivation, that is, moviegoers who want to watch the movies that others watch. A second effect is that some moviegoers are 'inattentive,' so their consideration set places a

*We are grateful to the Editor, two referees, and various seminar participants for helpful comments and suggestions. The usual disclaimer applies

†Authors' affiliations: Paganelli-Bull Professor of Economics and International Business, Stern School of Business, New York University; PSPS and IME Research Fellow (IESE); and Research Fellow, CEPR (London). New York, New York, U.S.A.
e-mail: lcabral@stern.nyu.edu.
‡Professor of Economics, Universidad de Piura, Lima, Peru.
e-mail: gabriel.natividad@udep.pe.

disproportionate weight on #1 movies. In other words, being #1 increases awareness of a movie's existence.

The two channels we consider parallel the classical persuasion-information dichotomy of the effects of advertising on demand. We derive theoretical results that allow us to identify the presence of both effects jointly and of information-mediated effects in particular. One possible test is that that being #1 increases the slope of the regression of box office revenue on movie quality. A second one is that increased consumer exposure to media promotion of films (by their actors and directors) negatively impacts the joint effect of being #1 and movie quality on total box office revenue (in other words, media exposure is a substitute for the awareness effect of being #1).

We test these predictions using U.S. box office data. Controlling for all variables that we are able to control for — including in particular movie quality, the competitive strength of contemporary rivals, and a variety of fixed effects — our results suggest that being #1 has an economically and statistically significant effect on a movie's eventual performance. On average, being #1 is associated with an increase of $68 to 75 million to a movie's total box office sales. Considering that the mean total sales of our 1,380 #1 movies is $93 million, this is a very large number indeed. Moreover, our regressions are consistent with the effect of being #1 appearing in interaction with the movie's quality; that is, being #1 is more beneficial for movies of higher quality. This result is consistent with the theoretical prediction that being #1 affects box office revenues by creating greater awareness of the movie's existence. Further evidence of this information effect is given by our finding that, for movies that were widely featured in the media prior to release, the effect of being #1 is smaller.

Our paper relates to a literature that focuses on estimating the demand for movies (Eliashberg and Shugan [1997], Sawhney and Eliashberg [1996], Elberse and Anand [2007], Elberse [2007], Moul [2007], Einav [2007], Natividad [2013]). Particularly relevant for our research is the work by Moretti [2011], who suggests that the opening weekend matters for a movie's subsequent performance through peer consumer effects triggered by either positive or negative initial surprises. To the best of our knowledge, ours is the first paper to specifically address the effect of being #1 in the box office on a movie's eventual success. Moreover, we show that the effect of a #1 ranking remains significant even when we consider the effect of opening weekend Moretti-like surprises.

Regarding the effect of ordinal rankings, our paper is related to the work by Sorensen [2007], who finds that appearing on the *New York Times* bestseller book list leads to a modest increase in sales for the average book. Similarly, Carare [2012] finds a positive impact of being among the the top 25 apps of the Apple store, though app sales data are unavailable and rankings are truncated. By contrast to books and apps, in the movie industry all new releases are ranked, as the number of new films every week is not large. This leads us to focus on the distinct effect of #1 ranking. Our work also differs in that we provide additional evidence regarding the possible channels through which #1 rankings affect demand.[1]

Another related literature studies the impact of a product's exposure on demand. Typically, this literature finds that, when faced with such lists, individuals often show a disproportionate tendency to select options that are placed at the top (Baye *et al.* [2009], Smith and Brynjolfsson [2001]).

More generally, our paper is related to an extensive economics and marketing literature on the effects of advertising. In economics, one frequently contrasts the persuasive and informative roles of advertising. More recently, this has been analyzed in the context of advertising models of demand that explicitly control for consideration sets (Andrews and Srinivasan [1995], Bronnenberg and Vanhonacker [1996], Goeree [2008], Haan and Moraga-Gonzalez [2011], Horowitz and Louviere [1995], Siddarth, Bucklin and Morrison [1995]). The relation with our research is that one can think of being #1 in the box office as a form of advertising. Given that, the question arises as to whether its effect on demand — which we estimate to be quite significant — occurs through channels similar to persuasive advertising (shocks to consumer utility) or rather through information effects (shocks to the consumers' consideration sets). Our empirical evidence is consistent with information playing an important role, that is, #1 rankings' making moviegoers aware of a movie's existence.

The paper is structured as follows. In Section II we lay down a theoretical model that features #1 effects through two different channels: utility shock and consumer awareness. We derive a series of theoretical propositions that imply specific empirical predictions. In Section III, we describe our data. The main empirical results are presented in Section IV. Section V concludes the paper.

II. THEORETICAL FRAMEWORK

As a motivation to our theoretical analysis, we first consider a simple plot of the data, depicted in Figure 1. On the horizontal axis, we measure opening weekend performance, defined as the difference in first weekend box office (in millions of 2009 dollars) between a movie and the second-best movie. Thus, movies with a positive value are box office winners, whereas

[1] Still regarding ordinal rankings, our work is related to the economics and strategy literature that studies ordinal vs. cardinal measures of success, in particular the importance of being #1 (Lazear and Rosen [1981], Podolny [1993], Cabral [2014]). A strand of the strategy literature has employed ordinal ranking methodologies to assess sustained competitive advantage over multi-year periods (Ruefli and Wilson [1987], Powell and Reinhardt [2010]); by contrast, we focus on the microdynamics of weekly sales that is typical of information goods in competitive markets.

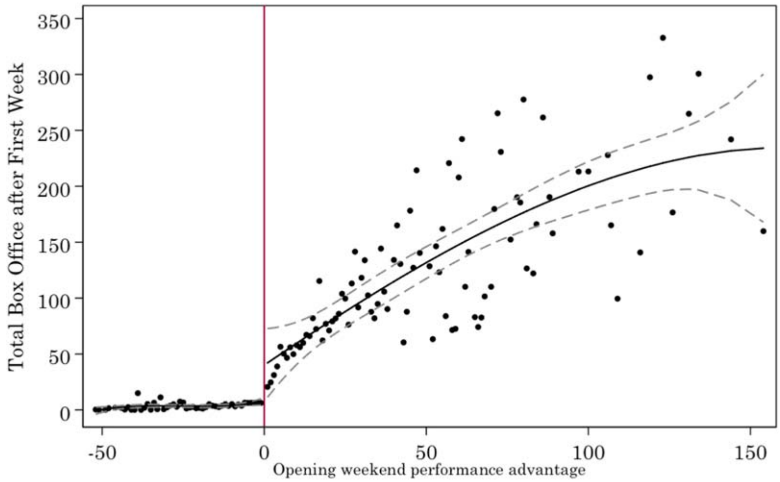

Figure 1
Relation between Performance Advantage at $t = 0$ and Eventual Performance Using Box Office Data

movies with a negative or zero value are non-box-office winners. Each unit in the horizontal axis is a bin of width equal to one million dollars. The vertical axis, in turn, shows the average total box-office performance of all movies in a given bin after their full theatrical run in the United States considering only weeks after the opening week. The dashed lines are 95% confidence intervals of third-degree polynomial fits of the opening week performance variable in explaining total box-office performance.

Figure 1 suggests that there is a very small discontinuity at zero; however, there is a considerable shift in the relation's slope at zero. In other words, (a) being #1 during the opening weekend has a positive effect on a movie's eventual performance, though small if the movie barely makes it to #1; and (b) the effect is greater the better a winner does during the opening weekend. We next consider a simple theoretical framework to better understand the possible effects of being #1 on a movie's success.

We consider two possibilities for the effect of being #1, that is, two channels through which being #1 affects theatrical performance. The first channel corresponds to an increase in consumer utility from watching a movie. This could result, for example, from social consumption benefits (e.g., watching a movie that most people will be talking about). In other words, it could be that being #1 during the opening weekend acts as a coordination device for some consumers who care primarily about discussing a movie with other moviegoers. A second possible channel is that topping the box office creates a special awareness for the movie. For example, the movie's title will appear on many headlines and thus enter the consideration set of consumes who would otherwise not know of its existence.

To consider these two possibilities, we propose a model whereby the lifespan of a movie's theatrical run can be divided into two periods, 0 and 1. We think of 0 as opening weekend and 1 as the rest of a movie's theatrical run. We will refer to $t=0$ consumers as early moviegoers and $t=1$ consumers as late moviegoers. Each week n new movies are released. For simplicity, we assume that movies are vertically differentiated. Specifically, let $r_i \in R^+$ be movie i's quality level.[2]

A central feature of our model is that we assume there are two types of consumers: 'informed' (type a) and 'inattentive' (type b). Fully informed consumers observe $\{r_i\}_{i=1}^{n}$, whereas inattentive consumers only observe the value of r_i of box office leaders.[3]

Movie demand evolves as follows. At $t=0$, only informed consumers (type a) watch movies. Each consumer watches each movie if and only if the utility from watching is greater than the outside option. We assume that utility is given by $u_i=r_i$.[4] For simplicity, we assume that each consumer views as many movies as there are values of r that exceed the outside option.[5] Also for simplicity, in what follows we omit the movie index i, with the understanding that our analysis applies to each movie.

At $t=1$, a fraction $1-\lambda$ of consumers is informed, whereas a fraction λ is inattentive. Informed consumers act as their $t=0$ counterparts with one difference: their utility from watching a movie is now given by $u=\alpha q^e+r$ if that movie topped the box office during the opening weekend, where q^e is the consumer's expectation regarding the value of q, the movie's total box office sales; and where we assume $\alpha > 0$.[6] In other words, αq^e measures the positive utility shock created by being #1 during the opening weekend. As we mentioned earlier, a natural interpretation is that αq^e measures the 'social' component of movie watching. Throughout the paper, we assume that consumers' expectation are fulfilled, so $q^e=q$.[7]

If the movie is not a top seller, then utility continues to be as before: $u=r$. Also at $t=1$, a fraction λ of consumers is inattentive. These

[2] In our empirical implementation, we will consider various dimensions of horizontal differentiation as well, such as genre.

[3] We do not need to determine the reason for consumer inattentiveness. It could be rational (consumers have a high cost of becoming informed), or it could result from some non-optimizing behavior pattern. Also, a more general version of the model could consider a continuum of types, with different probabilities of awareness of a given movie's existence. However, we believe the main qualitative features of the model would remain the same.

[4] More generally, we could assume that utility is proportional to movie quality r, and with no additional loss of generality simply assume that utility is given by r. In other words, by an appropriate change of units we make the utility coefficient equal to 1.

[5] An extension of our basic framework might consider the problem of selecting one or a number m of movies to watch at $t=0$.

[6] We also assume that α is sufficiently small so as to avoid the possibility of multiple self-fulfilling equilibria. If there are multiple equilibria, then our comparative statics results apply to sets of equilibrium points and thus lose 'bite' considerably.

[7] Also, while we assume utility αq, we could also include a fixed component, so that social utility would be $\alpha_0+\alpha q$. The same qualitative results follow in this alternative formulation.

consumers behave like $t=1$ informed consumers except that they are only aware of the existence of the #1 movie.

Suppose that type k's outside option ξ_k is distributed according to $F_k(\xi_k)$. Let μ be the ratio of late consumers with respect to early consumers. By an appropriate change of units, we assume that the number of early consumers equals 1. Finally, let 1 be an indicator variable that equals 1 if the movie was ranked number 1 at $t=0$, zero otherwise.

It follows from the preceding analysis that a movie's total box office sales are given by

$$(1) \qquad q = F_a(r) + \mu\Big((1-\lambda)\,F_a(r+\alpha\,q\,1) + \;1\,\lambda\,F_b(r+\alpha\,q\,1)\Big)$$

Notice that the indicator variable 1 appears three times in (1). Twice it appears as the argument of $F_k(\cdot)$. This represents the utility boosting effect of being number one, an effect that is measured by the parameter α interacting with movie sales. The third appearance of 1 is multiplying $\mu\,\lambda\,F_b(r+\alpha\,q\,1)$. This represents the awareness boosting effect of being number one, the fact that the movie now belongs to the inattentive consumers' consideration set.

We next develop a series of theoretical results which provide us with testable empirical implications from the model.

Proposition 1. Being #1 has a positive effect on a movie's theatrical performance:

$$q|_{1=1} > q|_{1=0}$$

Proof. See Appendix.

This is the most basic of our results: we expect that being number 1 implies a boost to a movie's eventual box office performance. Regardless of the channel that the #1 effect takes place, we expect the effect to be positive.

Our next two results pertain to the effect of being #1 on the relation between movie quality and movie performance. ■

Proposition 2. If most late moviegoers are inattentive, then being #1 increases the sensitivity of theatrical performance with respect to movie quality. Formally, there exists a $\lambda' \in [0,1)$ such that, if $\lambda > \lambda'$, then

$$\frac{\partial q}{\partial r}\bigg|_{1=1} > \frac{\partial q}{\partial r}\bigg|_{1=0}$$

Proof. See Appendix.

Intuitively, if late moviegoers are inattentive, then the primary effect of being #1 is to make late moviegoers aware of the movie's existence. If a movie is not #1, then late moviegoers are not aware of its existence, and the movie's total sales equal its sales during the opening weekend. To the extent that, conditional on awareness, late moviegoers are more likely to watch a movie if it is of higher quality, then being #1 not only increases attendance but also increases the sensitivity of attendance to movie quality.

Notice that the qualifier $\lambda > \lambda'$ is a sufficient, though not necessary, condition. In other words, Proposition 2 derives an empirical implication of our model but should not be seen as a separating test between the two channels we consider ('persuasion' and 'information'). ■

Our next result follows a similar line of reasoning. This one, we believe, provides a direct test for the 'information' or 'awareness' channel of being #1:

Proposition 3. Suppose that $f_a(r)$ is decreasing in r when $1=1$. If the fraction of uninformed consumers, λ, is sufficiently high, then an increase in λ implies an increase in the interaction term being #1 × movie quality, that is, $1\, r$:

$$\frac{d}{d\lambda}\left(\frac{\partial q}{\partial r}\bigg|_{1=1} - \frac{\partial q}{\partial r}\bigg|_{1=0}\right) > 0$$

Proof. See Appendix.

The intuition for Proposition 3 is similar to that for Proposition 2. The greater λ is, the greater the awareness effect of being #1. This effectively increases the sensitivity of a movie's performance to its quality. In other words, a good movie does well at the box office, but a good movie that is #1 during the opening weekend does particularly well. In fact, being #1 implies that many more consumers are aware of its existence; and being good implies that awareness turns into sales.

Propositions 2 and 3 are similar in that both refer to the interaction of a movie's quality and its sales ranking. However, the precise mathematical statement is different. In Proposition 2, we assume λ is high and posit that #1 movies have a higher $\partial q / \partial r$. That is, Proposition 2 corresponds to a difference in derivatives. By contrast, Proposition 3 corresponds to a derivative of a difference in derivatives: the greater λ is, the greater the difference between the derivatives $\partial q / \partial r$ between #1 and not #1 movies.

In terms of economics, Proposition 2 derives an implication of our model that holds true for both the 'persuasion' and 'information' channel. By contrast, Proposition 3 provides a direct test of the presence of an information effect.

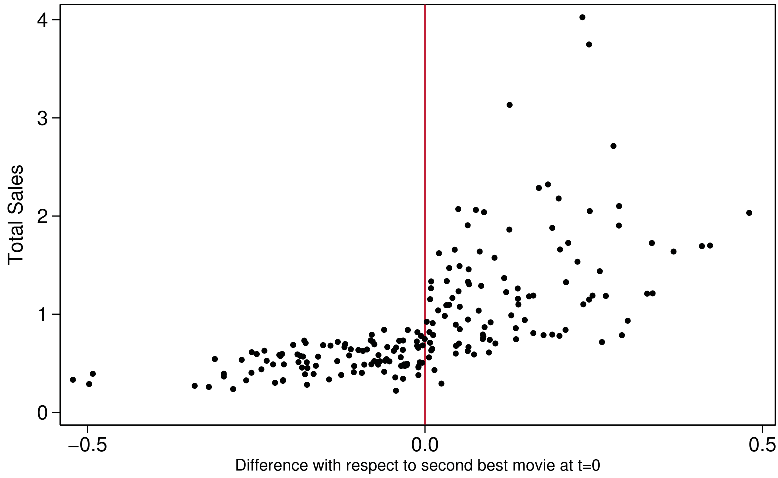

Figure 2
Relation between Performance at $t = 0$ and Eventual Performance Using Pseudo-Data

To conclude this section, we illustrate the model's mechanics by generating pseudo-data on its primitives. Suppose that $\lambda = 1$, $\mu = 6$ and that the outside option of type k consumer is normally distributed with parameters $\mu_a = \sigma_a = \sigma_b = 1$ and $\mu_b = 5$. By generating 100 observations of $n = 10$ movies each, we obtain the values in Figure 2. As can be seen, this theoretically-driven pattern is broadly consistent with the empirical features of Figure 1. The similarity between Figures 1 and 2 provides some support for our theoretical model as an explanation for the effect of #1 effects.

Propositions 1, 2 and 3 provide a series of empirical predictions regarding the effect of a #1 ranking. We examine these in the next sections by regressing q on r, 1, and a series of covariates. Specifically, testing Proposition 1 corresponds to estimating the effect of 1, that is, a dummy variable that is equal to 1 if the movie was #1 at the box office during the opening weekend. In order to test Propositions 2, we consider the interaction variable 1 r. Finally, in order to test 3 we consider the triple interaction between $1, r$ and a proxy for λ. ■

III. DATA

To investigate the effect of being #1, we assembled a database on the population of feature films released in U.S. theaters between 1 January, 1982 and 31 December, 2009. Our dataset draws from different sources: *Variety*, the leading industry periodical, and AC Nielsen EDI, a market research provider, report weekly and weekend box office revenue, weekly screens, and other movie characteristics such as genre. Studio System and *Variety* provide distribution company information. IMDb, an online database owned by *Amazon.com*, contains film- and person-level data not only on each feature film released but also on other appearances of cast members on TV shows and in the printed press, recording the exact date of these appearances. Proprietary information on production budgets through 2009 was acquired from Baseline Intelligence, a *New York Times* company; this provider is a well-trusted source used by industry decision-makers.

Our analysis is conducted at the movie level. From the population of all films released in the United States, we drop those that are re-releases of existent films, and we also drop 45 instances in which there was only one film released in a given week, resulting in a total of 9,933 distinct movies. Information on production budget is available for 79% of these films (7,854). We follow Moretti's [2011] procedure to fill in the missing production budget information on the 21% remainder using the industry average, thus keeping all feature films for estimation.[8] Our sample comprises 1,388 distinct weeks, each of which had an average of 7.16 opening films.

Each of the weekends in our sample period has a box office winner, yet not always is that film an opening film. Our dummy variable *Being #1* is therefore defined as equal to one when a film is the absolute winner of a weekend's box office during its first weekend in the marketplace across all opening and existing films that weekend. Only 8% of films achieve that #1 status.

Our data sources allow us to construct proxys for movie quality and movie exposure to media and other promotional events. Our first movie quality variable is given by 'star power.' This has been widely documented as as a signal of quality driving movie sales (Elberse, [2007]); in our context, it is operationalized as the sum of box office revenue of films of each team member over the three calendar years prior to film's release, divided by the number of team members; this variable, which we construct based on data from *Variety*, AC Nielsen EDI, and IMDb, is fixed at the moment of release.

A second measure of movie quality is given by average user ratings at IMDb, a proxy that has been used in previous research (e.g., Natividad, [2013]). It should be noted that this variable is only available after the film is released, thus requiring the assumption that the average consumer's judgment about the film reflects underlying quality already present at the moment of release.

Regarding media exposure prior to the release of a movie, we consider the sum of all printed press articles, interviews, magazine covers and magazine pictorials, as well as the sum of all TV show appearances of actors and directors in the four weeks prior to the release of the film; all this information is from IMDb. Alternatively, we measure participation in film festivals prior to the release date, also available from IMDb.

[8] For robustness, we also corroborated that our results hold in the sample that does contain production budget information without filling in for missing values.

TABLE I
DESCRIPTIVE STATISTICS (MOVIE LEVEL)

Variable	Obs	Mean	Std. Dev.	Min	Max
Total box office revenue	9929	22.22	49.71	0.00	802.92
Being #1	9929	0.08	0.27	0.00	1.00
Star power	9923	0.31	0.34	0.00	4.63
Quality rating	9891	6.18	1.21	1.10	9.80
Media exposure	9923	0.09	0.19	0.00	2.50

Table I provides basic summary statistics of the main variables used in the regressions that follow.

Controls and Fixed Effects. Our tests take the form of movie-level regressions of total box office revenue on proxys capturing the main features of our theory, as outlined above. To assuage concerns that the results may be driven by unobserved movie heterogeneity, we introduce the following control variables: the production budget of the film in millions of 2009 dollars, movie genre dummies, movie MPAA rating dummies, movie distribution company fixed effects, and date of release fixed effects. Importantly, while these granular controls and fixed effects do not fully account for potential omitted variables that may carry significant empirical weight, they constitute the broadest set of variables we can think of.

Varying levels of competition could be seen an alternative explanation for #1 effects, for example if films choose a particularly attractive week in which their appeal may be heightened, or if weak-competition weeks make #1 status a correlate of others' weakness rather than a booster of the film's own future performance. Our specification addresses the competition issue in two reinforcing ways. First, by introducing date-of-release fixed effects, we are effectively comparing each given film with all other films that chose the same week for release, thus narrowing the range of explanations for being #1 effects. Second, by introducing a control for the sales-weighted average of the production budget[9] of the five most expensive competing films other than the focal film playing during the first five weeks of its release, we essentially reduce the variability across weeks that may be confounding the effect of being #1 with a competitive story.

IV. EMPIRICAL RESULTS

As mentioned in Section II, if we consider a regression with box office revenues as a dependent variable, Proposition 1 implies that expected value conditional on $1=1$ is greater than expected value conditional on $1=0$. Proposition 2, in turn, implies a positive value for the interaction

[9] Alternatively, we use star power instead of production budget to create this moving average control variable; the untabulated results are essentially the same.

TABLE II
BEING #1 AND MOVIE QUALITY

	Dependent Variable: Total Box Office	
Being #1 × Star power	22.267**	
	(9.54)	
Being #1 × Quality rating		24.248***
		(2.76)
Being #1	62.655***	74.786***
	(6.34)	(15.79)
Star power	8.052***	
	(1.96)	
Quality rating		4.269***
		(0.40)
Controls and Fixed Effects	Yes	Yes
R^2	0.70	0.72
Sample size	9923	9891
N. of clusters (release date)	1384	1384

Notes: ***, **, * significant at the 1%, 5% and 10% level. Robust clustered standard errors in parentheses.

coefficient $1r$. Finally, Proposition 3 implies a test of a triple interaction. We report our tests in slightly different order from our theoretical section, given our empirical specification.

Testing Proposition 2. The two regressions in Table II correspond to two possible r variables: star power and movie quality ratings, both described in Section III. In both cases, the coefficient on $1r$ is positive, as predicted by Proposition 2. The coefficients are statistically significant (p value lower than 1%) and economically significant as well. Computed at mean variable values, the elasticity of box office revenues with respect to star power is 0.43 for #1 movies but only 0.11 for non #1 movies. For movie quality ratings, the elasticity increases from 1.19 to 7.93.

Alternatively, we can measure economic significance by the product of the estimated coefficient by the ratio of the standard deviations of independent and dependent variable. For star power, we get 0.21 for #1 movies and 0.055 for non #1 movies. In other words, an increase of one standard deviation in star power is associated with a 0.21 standard deviation increase in box office revenues for #1 movies, but only a 0.055 standard deviation increase in box office revenues for non #1 movies. For movie quality ratings, the numbers are 0.692 and 0.104.

In sum, the results are consistent with Proposition 2, both in terms of coefficient sign and statistical significance. Moreover, the effects seem economically meaningful.

Testing Proposition 1. Regarding Proposition 1, the results from the star power independent variable in Table II are clearly consistent with theory:

both the coefficient on 1 and the coefficient on 1 r are positive, so that, controlling for quality, being #1 is associated with higher box office revenues. At the mean value of the star power variable, this corresponds to a combined effect of 62.655+22.267×.313242=69.6299. In other words, controlling for movie quality, being #1 is associated with a $69 million increase in revenues. Regarding the model with movie quality ratings as an independent variable, it is not as clear that Proposition 1 holds. In fact, while the interaction term 1 r is positive, the coefficient on 1 is negative. However, when we account for the full effect of being #1 over the two coefficients of the regression at the mean value of the independent variable movie quality ratings we obtain −74.786+24.248×6.178959=75.0596; that is, controlling for movie quality, being #1 is associated with a $75 million increase in revenues, which is consistent with Proposition 1.

Testing Proposition 3. Proposition 3 states that, the greater the measure of inattentive consumers (parameter λ), the greater the impact of being #1 through the awareness channel. In order to test this possibility, we consider two alternative proxys for λ: pre-release media exposure and pre-release festivals. Our prediction is that alternative information sources about a movie's existence lead to a decrease in the value of λ and, by Proposition 3, a decrease in the interaction effect 1 r.

As mentioned earlier, Proposition 3 is about the derivative of a difference of derivatives: a triple interaction term between 1, r and λ. Since we have two possible measures of r and two possible proxys for λ, we have four possible regressions. The results are displayed in Table III. Three of the four regressions show the expected negative sign on the triple interaction term of interest, one of which is also statistically significant. The one regression whose coefficient is opposite to Proposition 3's prediction is not statistically significant.

Some Robustness Tests. The main purpose of our empirical regressions is to assess whether the implications of the theoretical model presented in Section II are borne out in the data; we do not claim to have exogenous variation or quasi-natural experiments for a causality test for the effect of being #1. With this caveat, we performed a series of additional regressions to evaluate the robustness of the results in Table II.

First, we compared close winners to close losers of the being #1 title in a non-parametric design using an optimal bandwidth to create the 'closeness' between losers and winners and to run local regressions. Specifically, we took only those films that were #1 by just a slight difference with respect to #2 films during their opening weekend, and compared them

TABLE III
MOVIE PRE-RELEASE INFORMATION AND THE EFFECT OF BEING #1

	Dependent Variable: Total Box Office	
Being #1 × Star power	23.728** (10.78)	
Being #1 × Quality rating		25.705*** (3.28)
Being #1 × Star power × Pre-media	−41.631** (18.98)	
Being #1 × Quality rating × Pre-media		−7.579 (10.19)
All other interaction terms and levels variables	Yes	Yes
Controls and Fixed Effects	Yes	Yes
R^2	0.70	0.72
Sample size	9923	9889
N. of clusters (release date)	1384	1384

Notes: ***, **, * significant at the 1%, 5% and 10% level. Robust clustered standard errors in parentheses.

with #2 films. This untabulated model yielded a positive and significant effect of being #1 even in a very narrow sample of winners vs. losers.

Second, we considered a set of matched pairs of films based on their quality proxys (i.e., star power and quality reviews). Specifically, we considered a subsample made up of each week's #1 film and a matched companion film released in the same week, the nearest neighbor using the Mahalanobis distance based on quality reviews and star power. Of all these selected pairs of films, where one is #1 and the other is its same-weekend match, we only took the 25% closest quality pairs and ran an OLS specification to see whether being #1 mattered for these otherwise identical quality films. As predicted by our theory, we find that it did.

Third, we also created matched pairs using alternative observable characteristics for matching. For example, we used either the first weekend sales of films; alternatively, we also used the number of opening screens. In these models, we found that being #1 always mattered positively for box office results and that the interaction between being #1 and quality was either positive and significant, or insignificant, depending on the proxy for quality. These supplementary results are useful, even if our theory does not focus on narrowly similar films to argue for the effects of being #1 in competitive markets.

Positive Surprises and Rankings. As mentioned earlier, Moretti [2011] proposes an alternative theory of why the opening weekend is so important for a movie's eventual box office performance: movies subject to positive surprises during the first weekend perform significantly better than movies not subject to positive surprises (or subject to negative surprises). In order to estimate whether the effect of being #1 goes beyond the positive surprise effect, we run a series of regressions that follow Moretti's [2011]

TABLE IV
BEING #1 VS. POSITIVE SURPRISES

	Dependent Variable: Total Box Office				
Positive surprise	25.557***		11.003***	9.848***	14.320***
	(0.97)		(0.79)	(0.79)	(0.65)
Being #1		119.763***	114.837***	33.975***	9.330**
		(1.45)	(1.48)	(6.57)	(4.32)
Interaction				85.098***	62.649***
				(6.74)	(5.31)
Controls and Fixed Effects	No	No	No	No	Yes
R^2	0.07	0.33	0.35	0.39	0.71
Sample size	9933	9933	9933	9933	9929
N. of clusters (release date)					1384

Notes: ***, **, * significant at the 1%, 5% and 10% level. Robust clustered standard errors in parentheses.

specification but add the #1 indicator variable. Specifically, as in Moretti [2011] we construct a proxy for movie surprises using a multivariate regression model, the residual of which leads to a 'positive surprise' dummy whenever it is greater than zero.

Table IV displays the results of regressions where we allow both for positive surprise effects and being #1 effects. The first two regressions show the effect of each factor in isolation. A positive surprise in the first weekend is associated with an increase in total box office revenues of about $25 million, whereas being #1 is associated with an increase of $119 million. When we allow for the two effects simultaneously (third regression), we note that both are statistically significant. The surprise effect drops by about 56% to around $11 million, whereas the #1 effect remains high ($114, down from $119). These results suggest that being #1 is more important for total movie sales than being subject to a positive surprise. It should be noted that, in a given week, several newly released movies are subject to positive surprises, whereas only one or none is ranked #1 when compared to all opening and existing films. Moreover, being #1 is an easily observable movie characteristic after the first weekend, whereas a positive surprise requires fitting a multivariate regression model first, in order to obtain an above-zero residual to determine whether the movie is subject to a positive surprise.

Column 4 of Table IV adds an interaction term that combines the positive surprise dummy and being #1. The results suggest that the interaction term is quite significant, both statistically and economically. In fact, most of the effect of being #1 seems to 'require' that the movie be a positive surprise, and vice versa: a substantial portion of the effect of a positive surprise 'requires' that the movie be ranked #1. The last regression of Table IV adds the usual control variables and fixed effects: the strength of competition moving average, the production budget of the film in millions of 2009 dollars, movie genre dummies, movie MPAA rating dummies, movie distribution company fixed effects, and date of release fixed effects. The results remain qualitatively unchanged: both being #1 and being subject to a positive surprise positively affect total box office performance.

For robustness, in untabulated tests we included variants of some of the controls of Table IV. For example, in addition to introducing Moretti's surprise dummy, we also included the actual surprise residual in the regressions alongside the being #1 dummy. We also included interactions of the surprise residual and the proxys for quality. Overall, we find broad support to the role of being #1 beyond the surprise factor documented by Moretti.

Summary. As mentioned earlier, we do not claim to have exogenous variation for a causality test for the effect of being #1. In the above regressions, the #1 indicator variable may be picking up heterogeneity in movie quality not measured by our quality variables star power and movie quality ratings. Having said that, we believe the empirical results provide strong conditional correlations consistent with our theoretical prediction of the effect of leading the box office during the opening weekend, in particular the relative importance of the information channel.

V. CONCLUSION

Many models of sales dynamics have used the film industry in the United States as an empirical laboratory to understand the impact of product characteristics and firm policies on demand. We propose a theoretical model of movie consumption where #1 rankings imply both utility shocks and increased awareness of a movie's existence. Our empirical results suggest that being #1 at the box office during the opening weekend has an economically and statistically significant effect on a movie's eventual performance, and that this effect is more pronounced the higher the quality of the box office leader. Additional empirical evidence is consistent with an information-mediated effect, that is, the idea that #1 movies are more likely to be in the consideration set of potential moviegoers.

APPENDIX

Proof of Proposition 1. The right hand side of (1) is strictly increasing in q. Moreover, the right-hand side is strictly greater when $1=1$ than when $1=0$. Standard monotone comparative statics results imply that q is greater when $1=1$. ■

Proof of Proposition 2. From (1), we get

$$\lim_{\lambda\to 1} q = F_a(r) + \mu 1\, F_b(r + \alpha\, q\, 1)$$

By the implicit function theorem,

$$\lim_{\lambda\to 1} \frac{\partial q}{\partial r} = \frac{f_a(r)+\mu\, 1\, f_b(r+\alpha\, q\, 1)}{1-\mu\, 1\, f_b(r+\alpha\, q\, 1)\,\alpha}$$

It follows that

$$\lim_{\lambda\to 1} \left(\frac{\partial q}{\partial r}|_{1=1} - \frac{\partial q}{\partial r}|_{1=0}\right) = \frac{f_a(r)+\mu f_b(r+\alpha\, q)}{1-\mu f_b(r+\alpha\, q)\,\alpha} - f_a(r) > 0$$

The result follows by continuity at $\lambda = 1$. ■

Proof of Proposition 3. By the implicit function theorem,

$$\frac{\partial q}{\partial r} = \frac{f_a(r)+\mu((1-\lambda) f_a(r+\alpha\, q\, 1) + 1\, \lambda f_b(r+\alpha\, q\, 1))}{1-\mu\,((1-\lambda) f_a(r+\alpha\, q 1)+\lambda\, 1\, f_b(r+\alpha\, q\, 1)\,\alpha)}$$

Omitting the arguments of $f_k(r+\alpha\, q 1)$, we have

$$\lim_{\lambda\to 1} \frac{\partial}{\partial \lambda}\left(\frac{\partial q}{\partial r}\right) = \frac{\mu\,(1 f_b - f_a)\,(N+D)}{D^2}$$

where

$$f_a = f_a(r+\alpha\, q\, 1)$$
$$f_b = f_b(r+\alpha\, q\, 1)$$
$$N = f_a(r)+\mu\, 1\, f_b$$
$$D = 1-\mu\, 1\, f_b\,\alpha$$

Define

$$f_{a0} = f_a(r)$$
$$f_{a1} = f_a(r+\alpha\, q)$$
$$f_{b1} = f_b(r+\alpha\, q)$$

Therefore,

$$\begin{aligned}\lim_{\lambda\to 1} \frac{\partial}{\partial \lambda}\left(\frac{\partial q}{\partial r}|_{1=1} - \frac{\partial q}{\partial r}|_{1=0}\right) &= \frac{\mu\,(f_{b1}-f_{a1})\,(1+f_{a1})}{(1-f_{b1})^2} + \mu f_{a0}\,(1+f_{a0}) \\ &> \mu\,(f_{b1}-f_{a1})\,(1+f_{a1}) + \mu f_{a0}\,(1+f_{a0}) \\ &> \mu f_{b1}\,(1+f_{a1}) - \mu f_{a1}\,(1+f_{a0}) + \mu f_{a0}\,(1+f_{a0}) \\ &= \mu f_{b1}\,(1+f_{a1}) + \mu\,(f_{a0}-f_{a1})\,(1+f_{a0}) \\ &> 0\end{aligned}$$

where the second and third inequalities follow from the assumption that f_a is decreasing. ■

REFERENCES

Andrews, R. and Srinivasan, T., 1995, 'Studying Consideration Effects in Empirical Choice Models using Scanner Panel Data,' *Journal of Marketing Research*, 32, pp. 30–41.

Baye, M.; Gatti, J.R.J.; Kattuman, P. and Morgan, J., 2009, 'Clicks, Discontinuities and Firm Demand Online,' *Journal of Economics and Management Strategy*, 18, pp. 935–975.

Bronnenberg, B. and Vanhonacker, W., 1996, 'Limited Choice Sets, Local Price Response and Implied Measures of Price Competition,' *Journal of Marketing Research*, 33, pp. 163–173.

Cabral, L., 2014, 'We're Number 1: Price Wars for Market Share Leadership,' working paper (New York University, New York, New York, U.S.A.).

Carare, O., 2012, 'The Impact of Bestseller Rank on Demand: Evidence from the App Market,' *International Economic Review*, 53, pp. 717–742.

Einav, L., 2007, 'Seasonality in the U.S. Motion Picture Industry,' *RAND Journal of Economics*, 38, pp. 127–145.

Elberse, A., 2007, 'The Power of Stars: Do Star Actors Drive the Success of Movies?,' *Journal of Marketing*, 71, pp. 102–120.

Elberse, A. and Anand, B.N., 2007, 'The Effectiveness of Pre-Release Advertising for Motion Pictures: An Empirical Investigation Using a Simulated Market,' *Information Economics and Policy*, 19, pp. 319–343.

Eliashberg, J. and Shugan, S.M., 1997, 'Film Critics: Influencers or Predictors?,' *Journal of Marketing*, 61, pp. 68–78.

Goeree, M., 2008, 'Limited Information and Advertising in the U.S. Personal Computer Industry,' *Econometrica*, 76, pp. 1017–1074.

Goldman, W., 1983, *Adventures in the Screen Trade*, Warner Books New York, New York, U.S.A.

Haan, M. and Moraga-Gonzalez, J., 2011, 'Advertising for Attention in a Consumer Search Model,' *Economic Journal*, 121, pp. 552–579.

Horowitz. J. and Louviere, J., 1995, 'What is the Role of Consideration Sets in Choice Modeling,' *International Journal of Research in Marketing*, 12, pp. 39–54.

Lazear, E.P. and Rosen, S., 1981, 'Rank-Order Tournaments as Optimum Labor Contracts,' *Journal of Political Economy*, 89, pp. 841–864.

Moretti, E., 2011, 'Social Learning and Peer Effects in Consumption: Evidence from Movie Sales,' *Review of Economic Studies*, 78, pp. 356–393.

Moul, C., 2007, 'Measuring Word of Mouth's Impact on Theatrical Movie Admissions,' *Journal of Economics & Management Strategy*, 16, pp. 859–892.

Natividad, G., 2013, 'Financial Slack, Strategy and Competition in Movie Distribution,' in *Organization Science*, 24, pp. 846–864.

Podolny, J.M., 1993, 'A Status-Based Model of Market Competition,' *American Journal of Sociology*, 98, pp. 829–872.

Powell, T. and Reinhardt, I., 2010, 'Rank Friction: An Ordinal Approach to Persistent Profitability,' *Strategic Management Journal*, 31, pp. 1244–1255.

Ruefli, T. and Wilson, C., 1987, 'Ordinal Time Series Methodology for Industry and Competitive Analysis,' *Management Science*, 33, pp. 640–661.

Sawhney, M.S. and Eliashberg, J., 1996, 'A Parsimonious Model for Forecasting Gross Box-Office Revenues of Motion Pictures,' *Marketing Science*, 15, pp. 113–131.

Siddarth, S.; Bucklin, R. and Morrison, D., 1995, 'Making the Cut: Modeling and Analyzing Choice Set Restriction in Scanner Panel Data,' *Journal of Marketing Research*, 32, pp. 255–266.

Smith, M. and Brynjolfsson, E., 2001, 'Consumer Decision-Making at an Internet Shopbot: Brand Still Matters,' *Journal of Industrial Economics*, 49, pp. 541–558.

Sorensen, A.T., 2007, 'Bestseller Lists and Product Variety,' *Journal of Industrial Economics*, 55, pp. 715–738.