THE JOURNAL OF INDUSTRIAL ECONOMICS 0022-1821
Volume LIII March 2005 No. 1

THE INFLUENCE OF EXPERT REVIEWS ON CONSUMER DEMAND FOR EXPERIENCE GOODS: A CASE STUDY OF MOVIE CRITICS*

David A. Reinstein† and Christopher M. Snyder‡

An inherent problem in measuring the influence of expert reviews on the demand for experience goods is that a correlation between good reviews and high demand may be spurious, induced by an underlying correlation with unobservable quality signals. Using the timing of the reviews by two popular movie critics, Siskel and Ebert, relative to opening weekend box office revenue, we apply a difference-in-differences approach to circumvent the problem of spurious correlation. After purging the spurious correlation, the measured influence effect is smaller though still detectable. Positive reviews have a particularly large influence on the demand for dramas and narrowly-released movies.

I. INTRODUCTION

There is an extensive theoretical literature (see, for example, Nelson [1970], Hey and McKenna [1981], Wiggins and Lane [1983], and Wolinsky [1995]) on consumer behavior in the presence of experience goods, i.e., goods for which the quality is uncertain prior to consumption. The related empirical literature has studied the impact of information about a product's quality on consumer demand from a variety of sources including advertising (Ackerberg [2003]), voluntary or mandatory product labeling (Teisl and Roe [1998], Foreman and Shea [1999], Mathios [2000], Jin and Leslie [2003]), social learning from peers (McFadden and Train [1996]), branding

*We are grateful to Roger Ebert for his prompt and insightful replies to our questions regarding the review process; to Darlene Chisholm, Maura Doyle, Robert Goldfarb, Jonathan Gruber, George Norman, Gityesh Pandya, Robert Porter, S. Abraham Ravid, W. David Walls, Anthony Yezer, and seminar participants at M.I.T. and the 2002 American Economic Association Winter Meetings for valuable comments; and to Genevieve Brittain and Neal Toomey for excellent research assistance. The paper was substantially improved following the suggestions of the editor and two anonymous referees. Some of this research was conducted while Snyder was on leave at the George G. Stigler Center for Study of Economy and State at the University of Chicago, the support of which is gratefully acknowledged. We sadly report that Gene Siskel, one of the two film critics on whom we focus of this study, died during the time we were writing the paper.

†Authors' affiliations: University of California, Department of Economics, 549 Evans Hall #3880, Berkeley, CA 94720, U.S.A.
e-mail: drein@berkeley.edu
‡Department of Economics, George Washington University, 1922 F. Street N.W., Washington, DC, 20052, U.S.A.
e-mail: csnyder@gwu.edu

(Montgomery and Wernerfelt [1992]), and indirect signals from firms' price, quantity, or advertising decisions (Nelson [1974], Caves and Greene [1996]).

In this paper, we contribute to the empirical literature by studying an additional source of product information: expert reviews. It is common to see books, concerts, movies, plays, restaurants, television shows, and other products of the entertainment industry reviewed by professional critics. Many other experience goods are also critically reviewed, whether in publications devoted to the whole range of consumer products (such as *Consumer Reports*) or to more narrow product classes (such as *PC Magazine*).

We have several motives for studying the influence of expert reviews on consumer demand. First, even if one considers expert reviews a close substitute for the other sources of information mentioned above, it is useful to study them to get a comprehensive picture of the aggregate flow of information that might influence consumers' demand for experience goods. Second, these other sources of information are not likely to be perfect substitutes for expert reviews in any event, making expert reviews worthy of independent study. The distinctive feature of expert reviews is that they are issued by a private party rather than the firm itself. On the one hand, the independence of the expert may reduce the bias in the information provided, increasing the influence on consumer demand. On the other hand, the expert may not have the same incentive to circulate the information to consumers, reducing the influence on demand. Of course, if the expert turns out to have a substantial influence on demand, the firm will have an incentive to 'capture' the expert through bribes or other means.[1] Third, new econometric problems are raised in measuring the influence of expert reviews as opposed to other sources of consumer information. As discussed in the following paragraphs, we address the econometric problems by exploiting a quasi-natural experiment in the particular industry we study, movies.

The inherent problem in measuring the influence of expert reviews on demand is that products receiving positive reviews of course tend to be of high quality, and it is difficult to determine whether the review or the quality is responsible for high demand. In formal econometric terms, the coefficient from the regression of demand on reviews will be biased upward due to the omission of quality variables. In principle the bias could be removed by accounting for quality; but quality is hard to measure for any product, especially for products whose quality is uncertain enough to merit critical

[1] The movie industry provides a recent example: the *New York Times* reported that Sony was fined by the Connecticut attorney general for inventing critic David Manning and quoting his fake reviews (e.g., 'Another Winner!') in the ads for at least four of its movies (Zielbauer [2002]). Michaely and Womack [1999] study stock analysts' biases toward firms with which their employers have other business dealings. Ravid and Wald [2002] study possible movie critics' biases toward certain production companies.

appraisals. In Eliashberg and Shugan's [1997] terms, the causal effect of reviews on demand holding quality constant is the *influence effect*; the spurious correlation between reviews and demand induced by their mutual correlation with quality is the *prediction effect*.

We propose a novel approach for distinguishing the influence and prediction effects of reviews on demand. The particular case we study is movies, an industry in which demand is readily measured by box office revenue. We consider the reviews of Siskel and Ebert, two movie critics who arguably had the greatest potential for influence through their nationally-syndicated television show. Our approach hinges on the timing of their reviews relative to a movie's release. Reviews that come during a movie's opening weekend can influence box office revenue for the remainder of the opening weekend; such reviews have both an influence and a prediction effect. Reviews that come after a movie's opening weekend cannot influence opening weekend revenue; such reviews have only a prediction effect. By taking a difference in differences—the difference between a positive and negative review for movies reviewed during their opening weekends and movies reviewed after—the prediction effect can be purged and the influence effect isolated. Our approach requires that the process by which the critics select movies to review during opening weekend and those to review after is independent of quality signals including the positiveness of their reviews. We provide tests suggesting that such selection effects are not substantial.

We find that a positive review has an influence on opening weekend box office revenue even after purging the prediction effect. The results for the combined sample of movies are only marginally statistically significant. The results are much stronger when broken down by subsample. We find an economically and statistically significant influence effect on opening weekend box office revenue for narrowly-released movies and for dramas. We find no influence effect for widely-released movies, or for genres such as action movies or comedies. Intuitively, critics' reviews are more important for 'art' movies than for 'event' movies, perhaps because, for this latter type of movie, consumers already have sufficient quality signals from press reports and advertising or consumers have a different view of quality than critics.

Results from additional regressions flesh out the model of consumer demand for movies. We find that a positive review during a movie's opening weekend does not merely steal business from later in the movie's run but in fact increases its total box office revenue. This increased revenue appears to come at the expense of competing movies showing during that weekend, although this effect is imprecisely estimated. Taken together, these results are consistent with a model in which quality-sensitive consumers have infrequent opportunities to see movies; they see high-quality movies when they have the opportunity, but do not have the opportunity to see all high-quality movies. In this model, consumers use quality information to make the secondary decision of which movie to see rather than the primary decision of whether to go out to the movies.

Our finding of a significant influence effect, at least for some types of movies, is in contrast with Eliashberg and Shugan [1997], the one previous study of box office revenue that attempts to separate influence from prediction effects. Using a sample of 56 long-running movies released in the early 1990s, the authors regress weekly box office revenue on the movie's percentage of positive reviews for each of the first eight weeks of a movie's run. They find that the percentage of positive reviews is only marginally significant during the first four weeks of the movie's run; the effect becomes larger and more significant during the next four weeks. Based on their maintained assumption that the influence effect declines during a movie's run, the authors conclude that the influence effect cannot be important and must be dominated by the prediction effect. In fact, we also find a similar pattern of increasing correlation between reviews and box office revenue over the course of a movie's run in our data, so cannot dispute their conclusion about the relative importance of the prediction and influence effects. That we still find a positive influence effect on opening weekend revenue may be due to our use of more powerful statistical tests—including over ten times the number of observations and a different measure of reviews (reviews of two influential critics rather than an average of hundreds of critics' reviews)—than Eliashberg and Shugan [1997].[2]

Besides our paper and Eliashberg and Shugan [1997], the rest of the literature on the relationship between movie reviews and box office revenue does not attempt to purge the prediction effect.[3] The studies tend to find a positive effect (Litman [1983]; Litman and Kohl [1989]; Wallace, Seigerman, and Holbrook [1993]; Sochay [1994]).[4,5] Relating our paper to the broader literature on the influence of reviews on consumer demand for a variety of

[2] Research contemporaneous with ours by Basuroy, Chatterjee, and Ravid [2003] supports this view. They apply Eliashberg and Shugan's [1997] methodology to a larger sample of 200 movies. They find that negative reviews have a significantly larger negative effect on box office revenue early in the movie's run than later, suggesting—if one maintains Eliashberg and Shugan's [1997] assumptions about the dynamics of the influence and prediction effects—a significant influence effect.

[3] There are several studies in communications literature (Faber and O'Guinn [1984], Wyatt and Badger [1984,1987]) that ask questions of focus groups regarding the importance of reviews relative to other ways of generating interest in a movie (advertising, word of mouth, etc.). There is much larger literature that forecasts box office revenue leaving aside critics' reviews, including Anast [1967], Austin [1984], Smith and Smith [1986], Austin and Gordon [1987], Dodds and Holbrook [1988], Prag and Casavant [1994], Sawhney and Eliashberg [1996], De Vany and Walls [1996,1997], Albert [1998], Neelamegham and Chintagunta [1999], and Moul [2004].

[4] An exception is Ravid [1999], which finds no significant effect.

[5] In many of these studies, the bias due to the prediction effect can be expected to be large since the source for reviews is an annual movie guide with ex post ratings rather than contemporaneous reviews.

products in addition to movies, much of the literature does not attempt to purge the spurious prediction effect. The papers that do so tend to focus on expert reviews which contain objective information, ranging from summaries of user reliability surveys as published by *Consumer Reports* for used cars (Hollenbacher and Yerger [2001]) to summaries of health plan performance indicators by various agencies (Spranca *et al.* [2000] and Jin [2002]). In a sense, these papers bear closer resemblance to the literature on product labeling cited above. Our paper differs from these in that the expert reviews we consider are more subjective, being the personal opinion of the expert. Whether such subjective reviews are more or less influential than more objective ones is an empirical question: on the one hand, there may be more new information in the subjective review than the publication of objective statistics that may already be commonly known; on the other, consumers may put less stock in soft information. Our paper also employs a methodology for separating the influence from the prediction effect which differs from these other papers.

II. MODEL

Let R_i be the box office revenue for movie $i = 1,\ldots, I$ measured over the time period T_i, for example movie i's opening weekend or entire run. Because movies are experience goods, consumers may seek signals of quality in advance of attending, such as the positiveness of a critic's review, denoted C_i, or other signals contained in advance publicity, marketing, word of mouth from others who have already seen the movie, etc., denoted S_i. Assume these signals influence consumer demand, and thus box office revenue, according to the following equation:

$$\ln R_i = \alpha + \beta D_i C_i + \delta S_i + e_i, \tag{1}$$

where α, β, and δ are coefficients, presumably with $\beta,\delta \geq 0$, and where e_i is an error term. The variable D_i is a dummy equal to one if the review C_i was published before the end of the period T_i and equal to zero if C_i was published after. Equation (1) indicates that C_i can influence consumer demand during the period T_i only if it was published before the end of T_i, i.e., only if $D_i = 1$.

We are primarily interested in estimating the coefficient β, which captures the influence of the critic's review on box office revenue. The presence of S_i in equation (1) poses an econometric problem since most of the components of S_i are likely to be unobservable to the econometrician. Letting $\hat{\beta}'$ be the ordinary least squares estimate of β from (1) omitting S_i, it can be shown (see, e.g., Wooldridge [2002], p. 62) that

$$\text{plim}\hat{\beta}' = \beta + \delta \frac{\text{Cov}(D_i C_i, S_i)}{\text{Var}(D_i C_i)}. \tag{2}$$

Since C_i and S_i are both signals of the movie's quality, they are likely to be positively correlated, implying the second term in (2) is positive, in turn implying $\hat{\beta}' > \beta$. This upward bias in $\hat{\beta}'$ is an instance of the standard omitted variables problem.

The traditional methodology (Litman [1983]; Litman and Kohl [1989]; Wallace, Seigerman, and Holbrook [1993]; Sochay [1994]; Eliashberg and Shugan [1997]; Ravid [1999]; Basuroy, Chatterjee, and Ravid [2003]) produces an estimate of β which is related to $\hat{\beta}'$. Rather than regressing ln R_i on D_iC_i omitting S_i (as with $\hat{\beta}'$), though, the traditional methodology can be roughly characterized as regressing ln R_i on C_i omitting S_i. That is, the traditional methodology includes all reviews as a right-hand side variable, without regard to their timing. Let $\hat{\beta}''$ be the resulting estimated coefficient on C_i. It can be shown that

$$\text{(3)} \qquad \operatorname{plim}\hat{\beta}'' = \bar{D}\beta + \delta\frac{\operatorname{Cov}(C_i, S_i)}{\operatorname{Var}(C_i)},$$

where $\bar{D}$ is the average value of the dummy variable D_i, or equivalently the fraction of movies reviewed before the end of revenue period T_i rather than after. There are two sources of bias in the traditional methodology. There will be the upward bias due to the omitted variable problem, given by the second term of (3). There will be another source of bias if some movies were not reviewed until after period T_i, for then $\bar{D} < 1$, biasing $\hat{\beta}''$ downward. It is impossible to tell ex ante if the net effect of the two biases is positive or negative.

Our approach to obtain a consistent estimate of β is to run a regression along the lines of

$$\text{(4)} \qquad \ln R_i = \alpha + \beta D_i C_i + \gamma C_i + u_i.$$

This regression is similar to equation (1) except C_i has been substituted as a proxy for the variable which is unobservable to the econometrician, S_i. For (4) to be identified, D_iC_i cannot be perfectly colinear with either the constant or C_i, which in turn requires there to be some movies that are reviewed before the end of T_i and some after.

It can be shown (see, e.g., Wooldridge [2002], pp. 63–64) that C_i is a good proxy for S_i, meaning that the ordinary least squares estimate $\hat{\beta}$ of the coefficient β in equation (4) will be consistent, if two conditions hold. First, C_i has to be redundant in the sense that it does not contribute to the conditional expectation of $\ln R_i$ once D_iC_i and S_i are known:

$$\text{(5)} \qquad \mathrm{E}(\ln R_i | D_i C_i, S_i, C_i) = \mathrm{E}(\ln R_i | D_i C_i, S_i).$$

Redundancy of C_i is true by definition since movies reviewed after period T_i cannot influence revenue R_i during the period. Second, letting v_i be the error in the linear projection of S_i on C_i,

$$\text{(6)} \qquad S_i = \lambda_0 + \lambda_1 C_i + v_i$$

where $\mathrm{E}(v_i) = \mathrm{E}(C_i v_i) = 0$, then $\mathrm{Cov}(D_i C_i, v_i) = 0$. A sufficient condition for $\mathrm{Cov}(D_i C_i, v_i) = 0$ is for D_i to be independent of C_i and v_i:

$$\text{(7)} \qquad \mathrm{Cov}(D_i C_i, v_i) = E(D_i C_i, v_i)$$

$$\text{(8)} \qquad = E(D_i)E(C_i v_i)$$

$$\text{(9)} \qquad = 0$$

where equation (7) holds since $\mathrm{E}(v_i) = 0$, (8) holds using the independence of D_i from C_i and v_i, and (9) holds since $\mathrm{E}(C_i v_i) = 0$. If the two conditions for C_i to be a good proxy for S_i hold, it is obvious that the right-hand side variables in equation (4) will be uncorrelated with the error u_i, and so $\hat{\beta}$ will be a consistent estimate of β.

To summarize, the preceding analysis highlights two key conditions that must hold for equation (4) to produce a consistent estimate of the influence effect. First, the data must contain variation in the timing of reviews, represented by D_i. That is, some movies must have been reviewed before the end of T_i and some after. As we will see in Section III, this requirement is satisfied by our data: we will take T_i to be the movie's opening weekend; about 80 percent of the movies in our data were reviewed before the end of opening weekend and about 20 percent after. Second, D_i must be independent of C_i and v_i, which is equivalent to D_i being independent of C_i and S_i. That is, the timing of the critic's review must be independent of the positiveness of the review and of other quality signals. A violation of this requirement suggests the presence of a selection effect, whereby the critic selects which movies to review when based on the quality of the movie. We will provide evidence that our data satisfy this independence requirement in Section VI.

Including the proxy C_i in the regression (4) makes our estimator of the influence effect a difference-in-differences estimator: it is the difference between the effect of a positive and negative review between movies reviewed before the end of period T_i and those reviewed after. Reviews before the end of T_i will have both the influence effect we are trying to measure plus the upward bias due to the omission of S_i (labeled the prediction effect in the Introduction). Reviews after the end of T_i will have only the omitted variable bias/prediction effect. Differencing the two purges the omitted variable bias/prediction effect.

Our final empirical specification is slightly richer than equation (4):

$$\text{(10)} \qquad \ln R_i = \alpha_0 + \alpha_1 D_i + \beta D_i C_i + \gamma C_i + X_i \theta + Z_i \mu + u_i.$$

This regression includes a vector of movie characteristics that are observable to the econometrician such as genre, producer identity, season of release, etc., denoted X_i. It includes D_i directly, allowing the intercept to depend on the timing of the review. We will sometimes add further proxies for movie quality, Z_i, such as other critics' reviews in certain specifications to reduce the error v_i in the linear projection of S_i on quality proxies in equation (6).

III. DATA

Our study focuses on the influence of two critics, Gene Siskel and Roger Ebert, on opening weekend box office revenue. Siskel and Ebert are ideal candidates for study because they were regarded as the most influential movie critics.[6] Their influence was due in large part to their nationally-syndicated television show (first titled *At the Movies*, later titled *Siskel & Ebert*) in which they each rendered their opinion on about four movies each week, a 'thumbs up' for a positive and a 'thumbs down' for a negative opinion.

Records were kept on the day movies were reviewed on their television show, allowing us to apply the estimation methodology from the previous section, which relies on the timing of the review relative to the movie's opening weekend. Consider Figure 1. The Friday, Saturday, and Sunday during the first week of a movie's run constitute its opening weekend.[7] For many observations in our data set, the movie was reviewed on Saturday morning during the opening weekend,[8] in which case we set the dummy variable *DURING* equal to one. For these observations, there is some potential for the review to influence box office for the remainder of the weekend. Even if consumers did not see Siskel and Ebert's television show itself before making their decision, positive reviews were often quoted in the movie's advertisements.

Most other movies were not reviewed until the week or several weeks after. There is no potential for these reviews to influence opening weekend box office revenue, though there will likely still be a positive correlation between them due to the prediction effect. A small number of movies were reviewed before the opening weekend. We omitted them from the final data set; pooling them with the $DURING = 1$ movies, or indeed treating them as a separate category, did not materially affect the results. In sum, our final data set consists of two groups of movies: those reviewed during their opening weekend, indicated by $DURING = 1$, and those reviewed after, indicated by $DURING = 0$.

[6] They were ranked among Smith's [1998] list of the 100 most influential people in the history of the movies, the only critics to make the list. Smith writes: "...[their] five-time Emmy Award winning program is aired on 180 of the country's broadcasting stations ...reachable through a staggering 95% of all television sets."

[7] In the case of a Monday holiday, the data typically fold the Monday figures into opening weekend.

[8] Siskel and Ebert's television show aired on Sundays in a few markets but aired on Saturdays in most markets including the largest ones (New York, Chicago, Los Angeles, etc.).

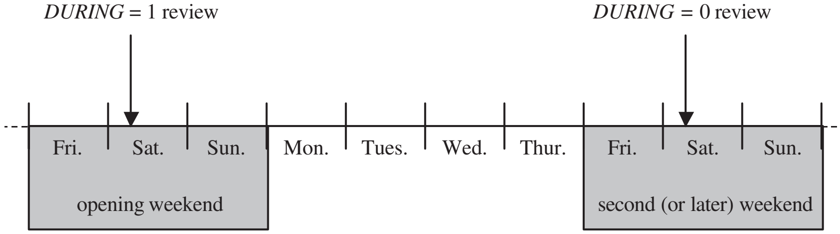

Figure 1
Timing of Review Relative to Movie's Release

Table I lists the main variables we will employ, together with descriptive statistics, and sources.[9] Our box office revenue variables (corresponding to R_i from Section II) are *TOTREV* and *OPENREV*. Our review variables (corresponding to C_i from Section II) are *SISKEL UP* and *EBERT UP*. It will sometimes be useful to combine Siskel's and Ebert's reviews into a single review variable, ranked as follows in terms of increasing quality: no thumbs up (implying $ONE\ UP = 0$ and $TWO\ UP = 0$), exactly one thumb up (implying $ONE\ UP = 1$ and $TWO\ UP = 0$), or two thumbs up (implying $ONE\ UP = 0$ and $TWO\ UP = 1$). Our dummy for the timing of the review relative to the opening weekend (corresponding to D_i in Section II) is *DURING*. Our other controls (corresponding to X_i in Section II) include *SCREENS* and *FOURDAY* reported in Table I, as well as dummies for year of release, month of release, genre,[10] and production company.[11] Our additional quality proxy (corresponding to Z_i in Section II) is *MALTIN*, the movie's rating on a 1–4 scale by Leonard Maltin, another popular film critic.[12] We also collected data on weekend box office revenue for all movies

[9] The final data set has 609 movie/observations. Merging data from the three sources—the Box Office Guru web site, the Siskel and Ebert web site, and Leonard Maltin's book of reviews (Maltin [1999])—resulted in 806 observations. We dropped 13 that were reviewed on a day other than Saturday, 137 that opened on fewer than 50 screens (indicating a 'narrow release strategy' which might confound our results), eight that were reviewed more than 40 days after opening, and 39 which were reviewed before the opening weekend.

[10] Movies are allowed to fall into more than one of our genres, which include adventure, animated, children's, comedy, crime, documentary, drama, fantasy, film noir, horror, musical, mystery, romance, science fiction, thriller, war, and western.

[11] We grouped production companies together with subsidiaries to form nine dummies: Disney (including Buena Vista and Miramax), Sony (including Sony, Columbia, Sony Classics, and TriStar), Fox (including Fox and Fox Searchlight), MGM/UA (including MGM/UA, MGM, Goldwyn, and United Artists), Gramercy, Orion, Universal, Warner Bros. (including Warner Bros., New Line, and Fine Line), and Paramount. The remaining movies were mostly produced by small companies (independents).

[12] Based on the work of Litman [1983] and later authors, who included Academy Awards as regressors in their revenue equations, we added information on Academy Awards—major awards such as best film, director, actor, and actress, and the other minor awards—to our data set. The data was taken from Maltin [1999]. For brevity, we do not report the regressions we ran including Academy Award variables because all alternative forms of the Academy Award variables we tried were insignificant, and their inclusion/exclusion had no effect on the other coefficients.

TABLE I
DEFINITIONS OF VARIABLES AND DESCRIPTIVE STATISTICS

Variable	Definition	Units	Mean	Standard Deviation	Maximum	Minimum
TOTREV	U.S. box office revenue for movie's entire run	million 1999 $	30.6	39.6	312.0	0.4
OPENREV	U.S. box office revenue for movie's opening weekend	million 1999 $	6.6	6.9	53.9	0.1
SISKEL UP	Siskel positive review ('thumbs up')	dummy	0.32	—	1	0
EBERT UP	Ebert positive review ('thumbs up')	dummy	0.41	—	1	0
ONE UP	Thumb up by exactly one of Siskel or Ebert	dummy	0.31	—	1	0
TWO UP	Thumbs up by both Siskel and Ebert	dummy	0.21	—	1	0
DURING	Siskel and Ebert review during opening weekend	dummy	0.81	—	1	0
SCREENS	Number of screens exhibiting on opening weekend	thousands	1.49	0.64	3.70	0.05
FOURDAY	Four day weekend (Monday holiday)	dummy	0.07	—	1	0
MALTIN	Rating by Leonard Maltin, a popular film critic	1–4 scale	2.29	0.56	3.50	1.00

Notes: 609 observations. *TOTREV* and *OPENREV* deflated using urban CPI index. *OPENREV* includes Friday, Saturday, and Sunday revenue for all cases except movies opening on four day weekends; these also include Monday revenue. *TOTREV*, *OPENREV* and *SCREENS* from the Box Office Guru web site (www.boxofficeguru.com). *SISKEL UP* and *EBERT UP* from Siskel and Ebert web site (www.tvplex.com/BuenaVista/SiskelandEbert). *MALTIN* from Maltin [1999]. *DURING* calculated by authors using opening date from the Box Office Guru web site and review date from the Siskel and Ebert web site. The rest of the variables computed by authors. Variables in table collected at movie level of observation. We also collected, from the *Variety* web site (www.variety.com), total U.S. weekend box office revenue for all movies for each of the 346 weekends in the data set, *MKTREV*, deflated using urban CPI index. The mean of *MKTREV* in million 1999 dollars is 59.7 and the standard deviation is 19.7

TABLE II
RAW CORRELATIONS AMONG SELECTED VARIABLES

	ln(*TOTREV*)	ln(*OPENREV*)	*MALTIN*	*SISKEL UP*	*EBERT UP*
ln(*OPENREV*)	0.95^{***}				
MALTIN	0.35^{***}	0.25^{***}			
SISKEL UP	0.16^{***}	0.13^{***}	0.27^{***}		
EBERT UP	0.17^{***}	0.09^{**}	0.33^{***}	0.35^{***}	
DURING	0.04	0.03	0.18^{***}	0.04	0.08^{**}

Note: Correlation coefficient significant at the *ten per cent level, **five per cent level, ***one per cent level in a two-tailed test.

in the U.S. market, labeled *MKTREV*, which will be used in the analysis of competitive effects using market-level data in Section VI. The notes to Table I provide descriptive statistics on *MKTREV*.

Table II presents correlations among revenue measures, Siskel and Ebert reviews, and the *MALTIN* quality proxy. It is tempting to conclude that there is an influence effect from the positive correlation between the reviews and the revenue measures. Since both influence and prediction effects are combined in the correlation, such a conclusion would be unwarranted. Indeed, the raw correlation between ln(*OPENREV*) and *SISKEL UP* is higher for movies reviewed after than during the opening weekend, and similarly for the correlation between ln(*OPENREV*) and *EBERT UP*, impossible if the influence effect were the only effect present.

From Table II it appears that *MALTIN* should serve as a good proxy for unobserved quality. The pattern of correlations between critics' reviews and revenue measures, namely higher with ln(*TOTREV*) than with ln(*OPENREV*), indicates that the critics' reviews are more correlated with revenue later in a movie's run than earlier, consistent with the findings of Eliashberg and Shugan [1997]. *MALTIN* is even more highly correlated with box office revenue than *SISKEL UP* and *EBERT UP*, but this is due in part to Maltin's rating scale being more refined than Siskel and Ebert's.

The correlations in the *DURING* row provide ambiguous evidence on the existence of selection effects which might cause our difference-in-differences estimator to be inconsistent. There is essentially no correlation between the revenue measures and *DURING*. On the other hand, certain of the review variables are positively correlated with *DURING*, raising the possibility of selection effects. Because of the importance of the selection issue for the consistency of our estimator, we explore it in more detail in the next section.

IV. EVIDENCE ON SELECTION EFFECTS

As noted in Section II, a sufficient condition for the consistency of our difference-in-differences estimator of the influence effect is for the timing of reviews D_i to be independent of their positiveness C_i or other quality signals S_i. In this section, we provide evidence that the process by which Siskel and Ebert selected which movies to review when was largely independent of quality, providing some confidence in the consistency of our results.

Our personal correspondence with Roger Ebert suggested that the decision to a review a movie during its opening weekend, rather than after, was random in most cases.[13] Siskel and Ebert's general policy was to review a movie during its opening weekend.[14] The most common reason for reviewing a movie after its opening weekend was that it simply did not fit in the previous show. In other cases, they needed to review a backlog of movies accumulated during a hiatus for attending film festivals. Studios only rarely prevented Siskel and Ebert from screening movies in advance of opening weekend, contradicting Smith's [1998] claim that studios often did this to keep two thumbs down from ruining the movie's opening.

To provide more formal evidence on the randomness of the selection process, we ran several specifications of a probit with *DURING* as the dependent variable and with revenue, the nature of Siskel and Ebert's reviews, and other controls used in the regressions in Section V below as right-hand side variables. The results are reported in Table III. For these and all subsequent regressions throughout the paper, we report White [1980] heteroskedasticity-robust standard errors in parentheses, adjusted to account for the possible correlation in the errors for movies released in the same weekend. Whether the review variables used are *SISKEL UP* and *EBERT UP* as in column (1) or *ONE UP* and *TWO UP* as in column (2), the review variables are not significant in the regression. Furthermore, there is little explanatory power in the probit. The pseudo R^2 is at most 22 percent. What explanatory power there is does not come from the revenue or Siskel and Ebert review variables: as shown in column (5), the log likelihood and pseudo R^2 remain essentially unchanged if the review variables, along with other quality proxies and revenue, are omitted from the regression entirely. Rather, most of the explanatory power comes from variables relating to the release date: *FOURDAY*, year dummies, and month dummies. The coefficients on these variables show that movies were more likely to be reviewed late when there was a large number of releases: during four day holiday weekends, during more recent years, and during the months of January, May, June, August, and December. This is consistent with

[13] Email correspondence on September 14 and September 17, 2000.

[14] The few cases in which they reviewed a movie before its opening weekend may have involved non-random selection. They sometimes issued early reviews for films which they wanted to boost or for films that were particularly newsworthy. In other cases, the selection process was more random, for example when they issued early reviews to avoid a backlog when a large number of movies were set to be released. In any event, the possible non-random selection of movies to be reviewed prior to opening supports our decision to drop those 39 observations.

TABLE III
PROBIT EVIDENCE ON SELECTION EFFECTS

	(1)	(2)	(3)	(4)	(5)
ln(*OPENREV*)	0.00	0.00	0.11	0.11	—
	(0.14)	(0.14)	(0.13)	(0.13)	
SISKEL UP	0.00	—	0.05	—	—
	(0.16)		(0.16)		
EBERT UP	0.08	—	0.19	—	—
	(0.15)		(0.14)		
ONE UP	—	0.17	—	0.22	—
		(0.16)		(0.16)	
TWO UP	—	0.03	—	0.21	—
		(0.22)		(0.20)	
MALTIN	0.43***	0.44***	—	—	—
	(0.13)	(0.13)			
SCREENS	– 0.06	– 0.06	– 0.21	– 0.22	—
	(0.20)	(0.20)	(0.20)	(0.20)	
FOURDAY	– 0.94***	– 0.94***	– 0.99***	– 1.00***	– 0.98***
	(0.28)	(0.28)	(0.27)	(0.27)	(0.27)
Constant	0.30	0.24	– 0.37	– 0.48	1.23***
	(1.81)	(1.81)	(1.78)	(1.79)	(0.39)
Year dummies	31.57***	31.32***	31.74***	31.57***	28.90***
Month dummies	31.03***	30.69***	33.29***	33.13***	26.57***
Genre dummies	17.11	17.54	18.60	18.85	—
Producer dummies	9.04	9.47	9.05	9.58	—
Log likelihood	– 227	– 226	– 231	– 231	– 233
Pseudo R^2	0.21	0.22	0.20	0.20	0.15

Notes: Dependent variable is *DURING*, the dummy variable indicating a movie that is reviewed by Siskel and Ebert during its opening weekend. Regressions involve 573 observations. The sole movie in the documentary genre, the sole movie in the film noir genre, and the 34 movies opening in December had to be dropped since their associated categories were perfect predictors of the dependent variable. White [1980] heteroskedasticity-robust standard errors, adjusted to account for possible correlation in the errors for movies opening in the same weekend, reported in parentheses below coefficient estimates. Entries for dummy variables are χ^2 statistics for test of joint significance. Significantly different from zero in a two-tailed test at the *ten per cent level, **five per cent level, ***one per cent level.

the claim above that Siskel and Ebert mainly reviewed movies after opening when they had too many movies to review in a given week on their show.

The one piece of evidence suggesting that the timing of Siskel and Ebert's reviews may not be completely independent of quality signals is the significance of *MALTIN*. High-quality movies—as gauged by Maltin's review—tended to be reviewed earlier by Siskel and Ebert. This is an odd result given that quality—gauged by the reviews of Siskel and Ebert themselves—had little effect on the timing of their reviews. We checked whether the inclusion of the *MALTIN* quality proxy might be masking the significance of the review variables—*SISKEL UP* and *EBERT UP* in column (3) and *ONE UP* and *TWO UP* in column (4)—by running the previous regressions omitting *MALTIN*. The review variables remained insignificant in all cases. In any event, the combined selection effect produced by *MALTIN* and variables besides those relating to release date cannot be too large. As shown in column (5), omitting all of these variables only reduces the probit's pseudo R^2 by 30 percent, or about six percent of the total variance of *DURING*.

V. MEASURING INFLUENCE AND PREDICTION EFFECTS

The difference-in-differences methodology for estimating the influence effect, embodied in equation (10), called for regressing ln R_i on D_i, D_iC_i, C_i, other controls X_i, and additional quality proxies Z_i. Translating these variables into their empirical counterparts defined in Section III, we will regress ln(*OPENREV*) on *DURING*, the interaction of *DURING* with our review variables *ONE UP* and *TWO UP*, the review variables *ONE UP* and *TWO UP* entered directly, other controls including *SCREENS*, *FOURDAY*, and dummies for year, month, genre, and production company, and the additional quality proxy *MALTIN*.

The basic regression is given in column (1) of Table IV. The variables of main interest are *DURING* × *ONE UP* and *DURING* × *TWO UP*. Given that *ONE UP* is included as a separate regressor, the coefficient on *DURING* × *ONE UP* provides a difference-in-differences estimate of the influence effect of one thumb up as a per cent of opening weekend box office revenue. That is, *DURING* × *ONE UP* is the marginal effect of having the review come during opening weekend rather than after on the marginal effect of one thumb up relative to two thumbs down. Similarly, *DURING* × *TWO UP* is the influence effect of two thumbs up as a per cent of opening weekend box office revenue. As the table shows, the influence effect of one thumb up is 11 per cent of opening weekend box office revenue, though statistically insignificant. The influence effect of two thumbs up is 25 per cent, marginally significant at the ten per cent level.

As equation (2) and the surrounding text suggests, the coefficients on *ONE UP* and *TWO UP* can be used to determine the direction and significance of the prediction effect, the spurious correlation between reviews and revenue caused by their mutual covariance with unobservable quality. Column (1) of Table IV shows that the prediction effect is positive, from four to seven per cent, but statistically insignificant.

The results for the ancillary variables are all in line with expectations. The coefficient on *DURING* is small and insignificant, showing that the regression's intercept does not vary with the timing of reviews. The coefficient on the additional quality proxy *MALTIN* is positive and highly significant. The coefficient on *SCREENS* is large and highly significant. The coefficient on *FOURDAY* is positive and marginally significant. The table does not report details on the fixed effects for year, month, genre and producer, reporting only the *F*-statistic from a joint test of each set's significance. Looking more carefully at the month fixed effects, the seasonal pattern of box office revenue emerges as expected: movies in the spring and fall tend to earn less than summer and winter, with revenues in

TABLE IV
OPENING WEEKEND BOX OFFICE REVENUE REGRESSIONS, BASIC SPECIFICATION, FOR VARIOUS SUBSAMPLES

	Breakdown by *SCREENS*			Breakdown by Genre		
Subsample:	All movies (1)	Wide Release (2)	Narrow Release (3)	Drama (4)	Action (5)	Comedy (6)
DURING	– 0.10	– 0.10	– 0.05	– 0.28	– 0.08	– 0.06
	(0.07)	(0.09)	(0.11)	(0.18)	(0.14)	(0.10)
DURING × *ONE UP*	0.11	– 0.06	0.13	0.51**	0.07	– 0.01
	(0.12)	(0.15)	(0.18)	(0.24)	(0.22)	(0.14)
DURING × *TWO UP*	0.25*	– 0.04	0.37*	0.65*	0.08	– 0.37
	(0.15)	(0.20)	(0.21)	(0.34)	(0.24)	(0.30)
ONE UP	0.07	0.26**	0.02	– 0.27	0.22	0.23*
	(0.10)	(0.13)	(0.15)	(0.22)	(0.21)	(0.12)
TWO UP	0.04	0.35*	– 0.15	– 0.42	0.14	0.72**
	(0.14)	(0.19)	(0.20)	(0.32)	(0.21)	(0.28)
MALTIN	0.27***	0.24***	0.33***	0.25***	0.32***	0.28***
	(0.04)	(0.05)	(0.06)	(0.08)	(0.07)	(0.06)
SCREENS	1.22***	1.15***	1.44***	1.17***	1.04***	1.13***
	(0.05)	(0.10)	(0.10)	(0.08)	(0.10)	(0.07)
FOURDAY	0.12*	0.08	0.18**	0.24	0.15	0.15
	(0.06)	(0.08)	(0.12)	(0.15)	(0.14)	(0.11)
Constant	12.22***	12.66***	11.79***	12.28***	12.17***	12.60***
	(0.18)	(0.29)	(0.25)	(0.38)	(0.36)	(0.29)
Year dummies	3.27***	1.60	1.08	1.32	1.21	0.96
Month dummies	3.93***	3.33***	1.12***	3.00***	3.17***	3.03***
Genre dummies	5.36***	3.36***	3.06***	—	—	—
Producer dummies	4.41***	2.57***	2.81***	3.21***	2.08**	2.15**
R^2	0.78	0.69	0.70	0.79	0.82	0.75
Observations	609	304	305	198	148	267

Notes: Ordinary least squares regressions. White [1980] heteroskedasticity-robust standard errors, adjusted to account for possible corelation in the errors for movies opening in the same weekend, reported in parentheses below coefficient estimates. Entries for sets of dummy variables are F statistics for test of joint significance. Significantly different from zero in a two-tailed test at the *ten per cent level, **five per cent level, ***one per cent level.

June and July significantly higher than the rest of the months. Other work, including Radas and Shugan [1998] and Krider and Weinberg [1998] also finds similar strong seasonal patterns.[15] Among genres, animated, children's, documentaries, and film noir earned less than average, while crime, fantasy, romance, and thrillers earned more. Movies from the large studios tended to earn significantly more than from independent studios.

In the remaining columns of Table IV, we repeat the basic specification for various subsamples of movies. Columns (2) and (3) break our sample into widely and narrowly-released movies, where widely-released movies opened on more than the sample median number of screens and narrowly-released movies on the sample median or fewer. The results show no influence effect

[15] We abstract from the competitive/strategic aspects of the timing of a movie's release. See De Vany and Walls [1997] and Chisholm [2003] for work along these lines.

for widely-released movies but a positive influence effect for narrowly-released ones. The influence effect of two thumbs up for narrowly-released movies was 37 per cent, statistically significant at the ten per cent level. Columns (4), (5), and (6) break our sample into the main genres: dramas, comedies, and action movies. There appears to be no influence effect for action movies and comedies, but a large and statistically significant influence effect for dramas: 51 per cent for one thumb up and 65 percent for two thumbs up. Taken together, these results for subsamples of movies are consistent with the intuition that critics' reviews influence 'art' movies but not 'event' movies. It may be that consumers of the 'event' movies (e.g., large budget action movies and comedies) already have good signals of quality from advertising and press reports, or that they are insensitive to the qualities judged by the critics. At the other extreme, say for an art house drama, Siskel and Ebert's positive review may largely determine its success or failure.

The coefficients on *ONE UP* and *TWO UP*, which as noted above are related to the prediction effect, have an interesting pattern across the results in columns (2) through (6) for various subsamples of movies. While the influence effect as measured by the interaction terms $DURING \times ONE\ UP$ and $DURING \times TWO\ UP$ is insignificantly different from zero for larger movies, action movies, and comedies, the prediction effect as measured by the coefficients on *ONE UP* and *TWO UP* is positive, statistically significantly so in many cases. The opposite is true for small movies and dramas. For them, the influence effect is significantly positive, but the prediction effect is zero or negative (the negative coefficients on *ONE UP* and *TWO UP* for dramas are large but statistically insignificant). These results suggest that for widely-released movies, action movies, and comedies, Siskel and Ebert's criteria for quality were similar to the average moviegoer's, though the average moviegoer was insensitive to reviews. For narrowly-released movies and dramas, Siskel and Ebert's criteria for quality may have only matched that of a smaller segment of the population ('high brow' consumers), and these 'high brow' consumers were more sensitive to critics' reviews than the average moviegoers.

The remainder of the section describes alternative specifications run to test the robustness of our baseline model and to provide more nuanced tests of our underlying demand model. We restrict attention to the subsample of dramas since this exhibits the strongest influence effect. In columns (1) and (2) of Table V, we investigate the importance of including additional quality proxies besides the review variables. In principle, our difference-in-differences estimator does not require additional quality proxies besides the review variables *ONE UP* and *TWO UP* for consistency, though adding them (referred to as Z_i in Section II) should improve the estimator's precision and should reduce any remaining bias in our estimator due to omitted quality signals. Column (2) of Table V removes the quality proxy

TABLE V
BOX OFFICE REVENUE REGRESSIONS FOR SUBSAMPLE OF DRAMAS

Dependent Variable:	ln(*OPENREV*) (1)	ln(*OPENREV*) (2)	ln(*OPENREV* ÷ *SCREENS*) (3)	ln(*TOTREV*) (4)
DURING	– 0.28	– 0.21	– 0.24	– 0.47
	(0.18)	(0.18)	(0.18)	(0.27)
DURING × *ONE UP*	0.51**	0.51**	0.32	0.70**
	(0.24)	(0.25)	(0.24)	(0.34)
DURING × *TWO UP*	0.65*	0.56*	0.58	0.85
	(0.34)	(0.33)	(0.42)	(0.55)
ONE UP	– 0.27	– 0.21	– 0.11	– 0.31**
	(0.22)	(0.23)	(0.22)	(0.31)
TWO UP	– 0.42	– 0.21	– 0.33	– 0.40
	(0.32)	(0.30)	(0.39)	(0.53)
MALTIN	0.25***	—	0.30***	0.52***
	(0.08)		(0.10)	(0.11)
SCREENS	1.17***	1.17***	—	1.23***
	(0.08)	(0.08)		(0.10)
FOURDAY	0.24	0.24	0.20	0.13
	(0.15)	(0.16)	(0.16)	(0.21)
Constant	12.28***	12.69***	6.66***	13.00***
	(0.38)	(0.32)	(0.40)	(0.54)
Year dummies	1.32	1.56	0.48	1.88*
Month dummies	3.00***	2.59***	2.62***	2.41***
Producer dummies	3.21***	3.52***	2.09**	3.27***
R^2	0.79	0.78	0.34	0.72

Notes: Ordinary least squares regressions for subsample of dramas, including 198 observations. White [1980] heteroskedasticity-robust standard errors, adjusted to account for possible correlation in the errors for movies opening in the same weekend, reported in parentheses below coefficient estimates. Entries for sets of dummy variables are *F* statistics for test of joint significance. Significantly different from zero in a two-tailed test at the *ten per cent level, **five per cent level, ***one per cent level.

MALTIN from the basic regression in the previous column. Removing *MALTIN*, which had a large, positive and statistically significant coefficient in column (1), from the regression in column (2) does not materially affect the coefficients of interest, i.e., those on *DURING* × *ONE UP* and *DURING* × *TWO UP*. The measured influence effect remains above 50 per cent for both one and two thumbs up and remains statistically significant. The additional quality proxy *MALTIN* takes over some of the function of the review variables *ONE UP* and *TWO UP* as quality proxies, causing the coefficients on the review variables to fall fairly uniformly.

The possible endogeneity of *SCREENS* is a concern. Producers may have opened movies they expected, for reasons unobservable to the econometrician, to earn more revenue on more screens, implying that *SCREENS* may be positively correlated with the error in the revenue equation. The true coefficient on *SCREENS* may be one, i.e., revenue is proportional to the number of screens, but the estimated coefficient may be biased upward, in turn biasing the rest of the coefficients in the regression. To check whether our results of central interest were being driven by the possible endogeneity of *SCREENS*, we re-ran the regression in column (1) constraining the coefficient on *SCREENS* to be one. We did this by using revenue per screen as the dependent variable rather than the revenue level and omitting *SCREENS* from the right-hand side. We have again restricted attention to the subsample of dramas. The results, presented in column (3) of Table V, are similar in magnitude to those in column (1). The standard errors are larger, so the estimated influence effects, though large at 32 per cent for one thumb up and 58 per cent for two thumbs up, are statistically insignificant.

The next regression in Table V fleshes out the model of consumer demand for movies. For simplicity, the model in Section II involved a static consumption decision. In practice, the consumer's decision is more complicated. It involves at least two component decisions: (a) the dynamic decision of whether to see a movie now or later in its run along with (b) the decision whether or not to substitute toward a different movie in a given weekend. The question (a) of whether a positive early review increases the demand for a particular movie or simply moves up the date at which a fixed number of consumers view it is addressed by regression (4) in Table V. (We will turn to the question (b) of whether positive reviews draw consumers away from competing movies in Section VI.) The regression is similar to that in column (1) except the dependent variable involves *TOTREV*, the total revenue over the movie's entire run, rather than opening weekend box office revenue *OPENREV*. Note that the coefficients on $DURING \times ONE\ UP$ and $DURING \times TWO\ UP$ still measure differences-in-differences. With *TOTREV* as the left-hand side variable, these coefficients are capturing the effect of earlier publication of a positive review on total revenue. On the one hand, if the influence effect expands total demand over the movie's entire run, publishing a positive review earlier will increase total revenue because it allows the influence effect to operate for a longer period of time—the first few weeks as well as subsequent weeks. The coefficients on $DURING \times ONE\ UP$ and $DURING \times TWO\ UP$ should then be positive. On the other hand, if the influence effect merely shifts demand from later in the movie's run to earlier, the coefficients on $DURING \times ONE\ UP$ and $DURING \times TWO\ UP$ should be zero. The results in column (4) of Table V again restrict attention to the subsample of dramas. The results are considerably noisier than those involving opening weekend box office, but are qualitatively similar. We find positive influence effects for one and two thumbs up, though the result is statistically significant only for one thumb up. This result implies that a positive early review does not simply shift a given demand for a movie earlier in its run but increases the movie's demand for its entire run. The point estimate of the influence effect, from 70 to 85 per cent, seems implausibly high as a percentage of a movie's box office revenue over its entire run. The size of this estimate could be due to a number of factors. First, the results in column (4) are noisier than the previous results, so we have less confidence in the point estimates. Second, the influence on opening weekend box office may indirectly increase revenue later in a movie's run through word-of-mouth effects. Third, the residual selection effects detected in Section IV may account for some of its size.

We conclude the section by comparing the results from our difference-in-differences estimator to the traditional methodology that regresses box office revenue on critics' reviews directly without differencing. As noted in Section II, the traditional estimator, labelled $\hat{\beta}''$, involves two opposing sources of bias, so it is impossible to tell *a priori* whether the resulting measure of the influence effect is positively or negatively biased. The results from applying the traditional methodology to our data set are presented in Table VI. Comparing column (1) of Table VI to column (1) of Table IV, we see that, if one does not include quality proxies, the traditional methodology overestimates the influence effect: 22 per cent compared to 11 per cent for one thumb up, and 37 per cent compared to 25 per cent for two thumbs up. In addition, the levels of statistical significance are considerably overstated by the traditional methodology. The influence effect of one thumb up is statistically insignificant with our difference-in-differences estimator but is significant at the one per cent level with the traditional methodology; the influence effect of two thumbs up is significant at only the ten per cent level with the difference-in-differences estimator but is significant at the one per cent level with the traditional methodology. Column (2) of Table VI shows

TABLE VI
REGRESSION USING THE TRADITIONAL METHODOLOGY

Subsample:	All Movies		*DURING* = 1	
	(1)	(2)	(3)	(4)
ONE UP	0.22***	0.15***	0.25***	0.18***
	(0.05)	(0.05)	(0.05)	(0.05)
TWO UP	0.37***	0.25***	0.40***	0.28***
	(0.06)	(0.06)	(0.06)	(0.07)
MALTIN	—	0.26***	—	0.27***
		(0.04)		(0.05)
SCREENS	1.22***	1.22***	1.17***	1.17***
	(0.05)	(0.05)	(0.05)	(0.05)
FOURDAY	0.07	0.13**	0.07**	0.12
	(0.06)	(0.06)	(0.08)	(0.08)
Constant	12.65***	12.14***	12.68***	12.13***
	(0.16)	(0.17)	(0.21)	(0.23)
Year dummies	3.85***	3.68***	2.72***	2.74***
Month dummies	3.75***	4.08***	3.73***	3.90***
Genre dummies	5.38***	5.17***	4.00***	4.35***
Producer dummies	5.47***	4.38***	4.10***	3.31***
R^2	0.76	0.78	0.76	0.77
Observations	609	609	492	492

Notes: Ordinary least squares regressions reflecting traditional methodology, which does not attempt to purge spurious prediction effect. White [1980] heteroskedasticity-robust standard errors, adjusted to account for possible correlation in the errors for movies opening in the same weekend, reported in parentheses below coefficient estimates. Entries for sets of dummy variables are *F* statistics for test of joint significance. Significantly different from zero in a two-tailed test at the *ten per cent level, **five per cent level, ***one per cent level.

that including quality proxies such as *MALTIN* reduces the bias associated with the traditional methodology considerably. The estimates of the influence effect are fairly close between column (2) of Table VI and column (1) of Table IV. While the point estimates are similar, the traditional methodology continues to overstate the significance levels associated with the point estimates.

In columns (3) and (4) of Table VI, we eliminate the downward bias in the traditional methodology stemming from averaging the effect of movies reviewed during opening weekend (and thus having the possibility of an influence effect) with those reviewed after (with no possibility of an influence effect). The regressions only include movies reviewed during opening weekend, yielding the estimator labelled $\hat{\beta}'$ in Section II. Estimates of the influence effect increase are higher in columns (3) and (4) than in (1) and (2) as expected. The influence effects in the regression omitting quality proxies (column (3)) remain higher with the traditional methodology than with the difference-in-differences estimator. Including quality proxies brings the estimates from the traditional methodology closer to the difference-in-differences estimates, but they are still higher. In all cases, the traditional methodology appears to overstate the statistical significance of the influence effect.

In sum, Table VI shows that using the traditional methodology without quality proxies results in a substantial overestimate of the influence effect. Including quality proxies makes the traditional methodology less problematic, but still can give incorrect inferences, for example leading one to conclude that one thumb up has a significant influence effect.

VI. EFFECT OF REVIEWS ON COMPETING MOVIES

The results from column (4) of Table V are consistent with a model in which quality-sensitive consumers have infrequent opportunities to see movies; they see high-quality movies when they have the opportunity, but do not have the opportunity to see all high-quality movies. If this model is correct, we should see a business-stealing effect, namely a positive review for one movie during a given weekend should have a negative influence on that movie's competitors.

This hypothesis is explored in the regressions in Table VII. Our data do not have disaggregated information on these competing movies, but we do have aggregate information on box office revenue for all movies each weekend, and by subtraction can determine the sum of competitors' box office revenue. The dependent variable in the table is the natural log of national weekend box office for all movies but the ones in our data set opening on that weekend. Unlike the regressions from previous tables, the unit of observation in Table VII is a weekend rather than a movie. The right-hand side variables are the reviews and other characteristics of the movies in

TABLE VII
COMPETITORS' WEEKEND BOX OFFICE REVENUE REGRESSIONS

	f = Sum (1)	f = Mean (2)
$f(DURING)$	0.01	– 0.03
	(0.03)	(0.05)
$f(DURING \times ONE\ UP)$	0.03	0.09
	(0.04)	(0.07)
$f(DURING \times TWO\ UP)$	– 0.08*	– 0.04
	(0.05)	(0.08)
$f(ONE\ UP)$	– 0.03	– 0.08
	(0.04)	(0.06)
$f(TWO\ UP)$	0.07	0.09
	(0.05)	(0.07)
$f(MALTIN)$	– 0.01	– 0.06
	(0.01)	(0.04)
$f(SCREENS)$	– 0.04***	– 0.04
	(0.01)	(0.03)
$FOURDAY$	0.31***	0.27***
	(0.05)	(0.06)
Constant	17.36***	17.55***
	(0.06)	(0.13)
Year dummies	45.30***	46.96***
Month dummies	33.61***	27.02***
R^2	0.67	0.54

Notes: Ordinary least squares regressions with dependent variable $\ln(MKTREV–\Sigma OPENREV)$, i.e., the natural log of weekend box office for all movies except those in our data set opening on that weekend. Unit of observation is a weekend; regressions involve 346 observations. Variables indicated by f operator are sums (resp. means) of the indicated variable over all movies in our data set opening during the weekend for regression (1) (resp. (2)). White [1980] heteroskedasticity-robust standard errors reported in parentheses below coefficient estimates. Entries for sets of dummy variables are F statistics for test of joint significance. Significantly different from zero in a two-tailed test at the *ten per cent level, **five per cent level, ***one per cent level.

our data set for the given weekend. Our data set often has several movies opening on the same weekend; it was unclear how best to aggregate these movies' characteristics to form right-hand side variables, so as shown we tried several specifications, including adding up the characteristics and taking the mean.[16]

Note that the lack of disaggregated information on competing movies prevents us from restricting attention to dramas as we did in Table VI. With disaggregated data, we could examine the effect of a positive review for a drama on competing dramas' revenue. We do not know the genre of competing movies, however; so in Table VII we examine the effect of a positive review for a movie in our complete sample (all genres) on all competing movies (all genres).

One thumb up did not have a significant effect on a movie's competitors. Two thumbs up had a negative influence on competitors, statistically

[16] We also restricted the sample to weekends in which only one movie in our data set was opening, so that no aggregation in the right-hand side variables was needed. The results from this regression, not presented here, were similar to column (1) of Table VII.

significant at the ten per cent level in column (1) (where we sum movies' characteristics) but not significant in column (2) (where we average movies' characteristics).

While it should be emphasized that there is considerable noise given the aggregation in these regressions, both in the dependent and independent variables, the results in Table VII provide suggestive evidence that positive reviews for movies steal business from competitors in a given weekend. Given that reviewed movies in our sample earned an opening weekend box office of \$6.6 million on average, while competitors' weekend box was \$48 million on average, based on the results in Table VII we cannot reject the hypothesis that weekend box office revenue for all movies in the market is not increased by a positive review for one.

VII. CONCLUSION

To summarize the central results, we find some weak evidence of an influence effect in our sample of all movies. The influence of one thumb up is 11 per cent of opening weekend box office revenue, but is statistically insignificant. The influence of two thumbs up is large in magnitude, at 25 per cent, but only marginally statistically significant. We find that the influence effect differs across categories of movies, strongest for movies with a narrower release and for dramas, virtually nonexistent for movies with a wider release and for action movies and comedies. We showed that an early positive review increases the number of consumers attending a movie in total over its entire run rather than simply shifting consumers from viewing the movie later rather than earlier. This increased revenue appears to come at the expense of competing movies. The results are consistent with a model in which some consumers have an inelastic demand for attending movies in certain weekends, and use reviews as a quality signal to determine which among the available movies to see.

Comparing our estimates, which rely on a difference-in-differences approach to purge spurious prediction effects and thereby to estimate a pure influence effect, to the traditional methodology, which directly regresses box office revenue on critics' reviews and thus does not purge the prediction effect, we saw that the traditional methodology can result in considerable positive bias. Including quality proxies reduces this bias, but did not eliminate the problem. In all cases, the traditional methodology appears to overstate the results' statistical significance considerably.

Our estimates of the influence effect are surprisingly large for dramas and narrow release movies, but perhaps not implausibly large. It is consistent with a survey reported in the *Wall Street Journal*, which found that a third of moviegoers chose a film because of a favorable review, more than half of these because of a review on television (Simmons [1994]). It should also be emphasized that our reduced-form model does not limit the influence effect to the direct influence of a critic on consumers but also includes indirect effects. For example, after a positive review, a movie distributor may choose to redouble its marketing efforts, highlighting the positive review in its advertisements. A positive review may influence one consumer to view the movie, who then influences others to view the movie through word of mouth. The sum of the direct and indirect influence effects, embodied in our estimate, may plausibly be quite high.

Another explanation of our large estimate of the influence effect is that there are selection effects violating our maintained assumptions that the reviews' timing is independent of quality signals. The evidence presented in Section IV suggests that selection biases may be present but are probably not large. The evidence is only suggestive since the tests in Table III cannot rule out selection based on variables outside of our data set.

Our results suggest that expert reviews can be an important mechanism for transmitting information about goods of uncertain quality. Our results also highlight the possibility that the power to influence consumer demand may be concentrated in a few critics. Reviews can themselves be considered goods of uncertain quality, and it may be natural for critics who have established high-quality reputations to exert the most influence. This raises interesting questions of how reputations may be built, maintained, and—for venal purposes—harvested.

REFERENCES

Ackerberg, D. A., 2003, 'Advertising, Learning, and Consumer Choice in Experience Good Markets: A Structural Empirical Examination,' *International Economic Review*, 44, pp. 1,007–1,040.

Albert, S., 1998, 'Movie Stars and the Distribution of Financially Successful Films in the Motion Picture Industry,' *Journal of Cultural Economics*, 22, pp. 249–270.

Anast, P., 1967, 'Differential Movie Appeals as Correlates of Attendance,' *Journalism Quarterly*, 44, pp. 86–90.

Austin, B. A., 1984, 'Portrait of an Art Film Audience,' *Journal of Communications*, 34, pp. 74–87.

Austin, B. A. and Gordon, T. F., 1987, 'Movie Genres: Toward a Conceptualized Model and Standardized Definition,' in Austin, B. A. (ed.), *Current Research in Film: Audiences, Economics and the Law*, vol. 3 (Ablex Publishing, Norwood, New Jersey).

Basuroy, S.; Chatterjee, S. and Ravid, S. A., 2003, 'How Critical Are Critical Reviews? The Box Office Effects of Film Critics, Star-Power, and Budgets,' *Journal of Marketing*, 67, pp. 103–117.

Caves, R. E. and Greene, D. P., 1996, 'Brands' Quality Levels, Prices, and Advertising Outlays: Empirical Evidence on Signals and Information Costs,' *International Journal of Industrial Organization*, 14, pp. 29–52.

Chisholm, D. C., 2003, 'The War of Attrition and Optimal Timing of Motion-Picture Releases', Suffolk University working paper.

De Vany, A. and Walls, W. D., 1996, 'Bose-Einstein Dynamics and Adaptive Contracting in the Motion Picture Industry,' *Economic Journal*, 106, pp. 1,493–1,514.

De Vany, A. and Walls, W. D., 1997, 'The Market for Motion Pictures: Rank, Revenue and Survival,' *Economic Inquiry*, 35, pp. 783–797.

Dodds, J. C. and Holbrook, M. B., 1988, 'What's an Oscar Worth? An Empirical Estimation of the Effect of Nominations and Awards on Movie Distribution and Revenues,' in Austin, B. A. (ed.), *Current Research in Film: Audiences, Economics and the Law*, vol. 4 (Ablex Publishing, Norwood, New Jersey).

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

Faber, R. and O'Guinn, T., 1984, 'Effect of Media Advertising and Other Sources on Movie Selection,' *Journalism Quarterly*, 61, pp. 371–377.

Foreman, S. E. and Shea, D. G., 1999, 'Publication of Information and Market Response: The Case of Airline On-time Performance Reports,' *Review of Industrial Organization*, 14, pp. 147–162.

Hey, J. D. and McKenna, C. J., 1981, 'Consumer Search with Uncertain Product Quality,' *Journal of Political Economy*, 89, pp. 54–66.

Hollenbacher, A. and Yerger, D. B., 2001, 'Third Party Evaluations and Resale Prices in the U.S. Used Vehicle Market,' *Applied Economics Letters*, 8, pp. 415–418.

Jin, G. Z., 2002, 'Did Quality Information Matter? Evidence from Medicare, 1993–1998', University of Maryland mimeo.

Jin, G. Z. and Leslie, P., 2003, 'The Effects of Information on Product Quality: Evidence from Restaurant Hygiene Cards,' *Quarterly Journal of Economics*, 118, pp. 409–451.

Krider, R. E. and Weinberg, C. B., 1998, 'Competitive Dynamics and the Introduction of New Products: The Motion Picture Timing Game,' *Journal of Marketing Research*, 35, pp. 1–15.

Litman, B. R., 1983, 'Predicting Success of Theatrical Movies: An Empirical Study,' *Journal of Popular Culture*, 16, pp. 159–175.

Litman, B. R. and Kohl, L. S., 1989, 'Predicting Financial Success of Motion Pictures: The '80s Experience,' *Journal of Media Economics*, 2, pp. 35–50.

Maltin, L., 1999, *Movie and Video Guide: 2000 Edition* (Signet, New York).

Mathios, A. D., 2000, 'The Impact of Mandatory Disclosure Laws on Product Choices: An Analysis of the Salad Dressing Market,' *Journal of Law and Economics*, 43, pp. 651–677.

McFadden, D. L. and Train, K. E., 2000, 'Consumers' Evaluation of New Products: Learning from Self and Others,' *Journal of Political Economy*, 104, pp. 683–703.

Michaely, R. and Womack, K. L., 1999, 'Conflict of Interest and the Credibility of Underwriter Analyst Recommendations,' *Review of Financial Studies*, 12, pp. 653–686.

Montgomery, C. A. and Wernerfelt, B., 1992, 'Risk Reduction and Umbrella Branding,' *Journal of Business*, 65, pp. 31–50.

Moul, C. C., 2004, 'Handling Saturation in Demand: Estimates from Theatrical-Run Motion Pictures,' Washington University working paper.

Neelamegham, R. and Chintagunta, P., 1999, 'A Bayesian Model to Forecast New Product Performance in Domestic and International Markets,' *Marketing Science*, 18, pp. 115–136.

Nelson, P., 1970, 'Information and Consumer Behavior,' *Journal of Political Economy*, 78, pp. 311–329.

Nelson, P., 1974, 'Advertising As Information,' *Journal of Political Economy*, 81, pp. 729–754.

Prag, J. and Casavant, J., 1994, 'An Empirical Study of the Determinants of Revenues and Marketing Expenditures in the Motion Picture Industry,' *Journal of Cultural Economics*, 18, pp. 217–235.

Radas, S. and Shugan, S. M., 1998, 'Seasonal Marketing and Timing Introductions,' *Journal of Marketing Research*, 35, pp. 296–315.

Ravid, S. A., 1999, 'Information, Blockbusters, and Stars: A Study of the Film Industry,' *Journal of Business*, 72, pp. 463–492.

Ravid, S. A. and Wald, J. K., 2002, 'Movie Critics as Analysts: A Story of Industry Capture', Rutgers Business School working paper.

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.

Simmons, J., 1994, 'A "Thumbs Up" Pulls in the Audience,' *Wall Street Journal*, March 25, p. B1.

Smith, S., 1998, *The Film 100: A Ranking of the Most Influential People in the History of the Movies*, (Citadel Press, New York).

Smith, S. P. and Smith, V. K., 1986, 'Successful Movies: A Preliminary Empirical Analysis,' *Applied Economics*, 18, pp. 501–507.

Sochay, S., 1994, 'Predicting the Performance of Motion Pictures,' *Journal of Media Economics*, 7, pp. 1–20.

Spranca, M. E.; Kanouse, D. E.; Short, P. F.; Farley, D. O. and Hays, R. D., 2000, 'Do Consumer Reports of Health Plan Quality Affect Health Plan Selection?,' *Health Services Research*, 35, pp. 933–947.

Teisl, M. F. and Roe, B., 1998, 'The Economics of Labeling: An Overview of Issues for Health and Environmental Disclosure,' *Agricultural and Resource Economics Review*, 27, pp. 140–150.

Wallace, W. T.; Seigerman, A. and Holbrook, M. B., 1993, 'The Role of Actors and Actresses in the Success of Films: How Much is a Star Worth?,' *Journal of Cultural Economics*, 17, pp. 1–27.

White, H., 1980, 'A Heteroskedasticity-Consistent Covariance Matrix Estimator and a Direct Test for Heteroskedasticity,' *Econometrica*, 48, pp. 817–830.

Wiggins, S. N. and Lane, W. J., 1983, 'Quality Uncertanty, Search, and Advertising,' *American Economic Review*, 73, pp. 881–894.

Wolinsky, A., 1995, 'Competition in Markets for Credence Goods,' *Journal of Institutional and Theoretical Economics*, 151, pp. 117–131.

Wooldridge, J., 2002, *Econometric Analysis of Cross Section and Panel Data* (MIT Press, Cambridge, Mass.).

Wyatt, R. O. and Badger, D. P., 1984, 'How Reviews Affect Interest in and Evaluation of Films,' *Journalism Quarterly*, 61, pp. 874–878.

Wyatt, R. O. and Badger, D. P., 1987, 'To Toast, Pan or Waffle: How Film Reviws Affect Reader Interest and Credibility Perception,' *Newspaper Research Journal*, 8, pp. 19–30.

Zielbauer, P., 2002, 'Metro Briefing Connecticut: Hartford: Falsified Movie Reviews,' *New York Times*, March 13, p. B8.