Management Science

Publication details, including instructions for authors and subscription information:
http://pubsonline.informs.org

Counterfactual Decomposition of Movie Star Effects with Star Selection

Angela (Xia) Liu, Tridib Mazumdar, Bo Li

To cite this article:
Angela (Xia) Liu, Tridib Mazumdar, Bo Li (2015) Counterfactual Decomposition of Movie Star Effects with Star Selection. Management Science 61(7):1704-1721. http://dx.doi.org/10.1287/mnsc.2014.1923

MANAGEMENT SCIENCE
Vol. 61, No. 7, July 2015, pp. 1704–1721
ISSN 0025-1909 (print) | ISSN 1526-5501 (online)

http://dx.doi.org/10.1287/mnsc.2014.1923

Counterfactual Decomposition of Movie Star Effects with Star Selection

Angela (Xia) Liu
School of Economics and Management, Tsinghua University, Beijing 100084, People's Republic of China, liux@sem.tsinghua.edu.cn

Tridib Mazumdar
Whitman School of Management, Syracuse University, Syracuse, New York 13244, tmazumda@syr.edu

Bo Li
School of Economics and Management, Tsinghua University, Beijing 100084, People's Republic of China, libo@sem.tsinghua.edu.cn

We investigate the effects of a movie star on the movie's opening week theater allocations and box office revenue. Because the pairing of a star and a movie involves a bilateral matching process between the studio and the star, the star (hence the nonstar) movie samples are nonrandom and the star variable is potentially endogenous. To assess the star as well as movie characteristics effects, we utilize a switching model to account for endogenous assignment of stars and nonstars into respective movie samples. In addition to controlling for selection biases, the endogenous switching model generates managerially relevant insights into the factors that influence a star's assignment to a movie. Additionally, because the star and nonstar movie characteristics (e.g., movie budget, distribution, genre, etc.) are often systematically different, we counterfactually estimate the theater allocations and revenues that nonstars (stars) *would have* generated had they acted in movies endowed with the same characteristics as the star (nonstar) movies. The decomposition analysis, conducted at different quantiles of theater and revenue distributions, shows that the presence of a star has a much stronger effect on theater allocations than the movie characteristics have. However, the revenue difference is entirely contributed by the differences in the characteristics of the star and nonstar movies. Thus, the star effects on revenue come indirectly through the theater allocations as well as from the characteristics of the movies in which they participate.

Keywords: movie; star power; star selection; endogenous switching model; counterfactual decomposition; quantile regression
History: Received March 8, 2011; accepted January 23, 2014, by Pradeep Chintagunta, marketing. Published online in *Articles in Advance* August 11, 2014.

1. Introduction

Stars are ubiquitous in industries such as entertainment, athletics, healthcare, and stock brokerage and within large corporations. Economists have proposed theories on how individual participants emerge as stars in an industry (Adler 1985, MacDonald 1988, Rosen 1981), and why only a few participants achieve star status and extract disproportionately large financial rewards (MacDonald 1988). In the movie industry, stars facilitate the "green-lighting" process and are believed to mitigate risks of largely unpredictable box office outcome (De Vany and Walls 1999, Vogel 2007, Basuroy et al. 2003). Stars, on their part, not only extract large payments for their participation but also can often ensure that the movie contains the ingredients of a potential blockbuster (Elberse 2007, Lehmann and Weinberg 2000, Vogel 2007). The financial rewards stars receive and the power they wield have raised concerns within the movie industry about the returns they generate. While some industry observers maintain that stars do influence movie outcomes positively, others highlight instances of big-star movie flops, labeling stars as "overpaid" (Ackman 2002, 2003). Despite the evidence that stars do not necessarily ensure successful outcomes, biases toward stardom exist—credit for successful movies is typically bestowed upon stars, while flops are attributed to the "bad luck" of the producers.

Whether stars indeed contribute to the success of a movie has also been the focus of several academic studies (e.g., Albert 1998; De Vany and Walls 1999, 2004; Karniouchina 2011; Luo et al. 2010; Ravid 1999). Some studies find positive relationships between the presence of a star in a movie and its financial outcomes (Luo et al. 2010, Sochay 1994, Wallace et al. 1993), while others find that, after accounting for payments to stars, stars do not necessarily increase a movie's return on investment (ROI) (e.g., De Vany and Walls 1999, Ravid 1999) or the value of the movie company

(Elberse 2007). In reviewing academic research and industry practices, Eliashberg et al. (2006) recommend additional research to answer the question: "To what extent do stars actually contribute to the success of movies?" (p. 645).

The objective of this paper is to investigate the effects a star has on the number of theaters allocated to a movie in its opening week and its opening week revenue.[1] An important consideration in the investigation of star effect is the star–movie matching process in which the studio matches a movie (or a script) with a star, and the star evaluates whether the payments and movie's outcome potentials are attractive enough to retain or enhance status as a star in the profession (Ravid 1999, Sorenson and Waguespack 2006). A star's decision to accept or reject a studio's invitation renders the sample of star movies nonrandom and the star variable potentially endogenous (Elberse 2007). Although the star–movie matching process can potentially have biasing influence on the assessment of the effects of stars and other movie covariates (Eliashberg et al. 2006, Luo et al. 2010, Ravid 1999, Vogel 2007), the process thus far has remained unexplored in investigations of star effects.

We account for the endogenous assignment of a star (hence a nonstar) to a movie by employing an endogenous switching model in which a selection model is estimated in the first stage, producing a selection correction term (i.e., inverse Mills ratio; Heckman 1979) for each movie in the star and nonstar movie samples. These corrections are appended to the second stage models estimating covariate effects in the opening week theater allocation and revenue for each subsample.[2] In addition to correcting the potential bias introduced by the endogenous star participation in a movie, this procedure produces new substantive insights into the star–movie matching process. For instance, we find that star participation is influenced by past recognitions of the director and producers, production companies, likely shooting locations, and expected Motion Picture Association of America (MPAA) ratings.

A second difficulty in the assessment of star effects arises from potentially systematic differences between the star and nonstar movie characteristics, which may also cause the movie outcomes to be different. For instance, star movies are typically associated with larger budgets and wider distributions (Ravid 1999, De Vany and Walls 1999, Vogel 2007). Stars may demand and secure special effects and exotic locations or may influence the movie's script to be more consistent with their strength in a genre or to affect the movie's MPAA ratings. Because the presence of a star in a movie and the characteristics of the movie may be intertwined, a proper assessment of star effects on movie outcomes requires isolating the star effects on a movie's outcome from the effects caused by the characteristics of the movie. Past research has not separated the star effects from the movie characteristics effects.

To isolate the stars effects from the effects of the movie characteristics, we frame the question as follows: what would the movie's outcome be if a nonstar actor/actress had instead acted in a movie with the same characteristics that a star movie possesses, or vice versa? This counterfactual approach decomposes the differences in the opening week theater count and box office revenue of movies with and without a star into two parts: one part is due to the observed differences in the distribution of characteristics of star versus nonstar movies (hereafter, "characteristics" effects) and the other part is due to the difference in the *returns* that stars can incrementally derive over nonstars from these characteristics (hereafter, "star return" effect).

We estimate the star effects on theaters and revenue and carry out the counterfactual decompositions at different percentiles of the outcome distributions by using a quantile regression (QR) framework (Koenker and Bassett 1978, Machado and Mata 2005, Melly 2005). It is well-known that box office returns are heavily skewed, with a large proportion of movies achieving only a modest return at the box office (De Vany and Walls 2004). This distributional feature makes the ordinary least square estimates at the mean misleading (Walls 2005). A QR approach is less susceptible to outliers and heavy tails. Also, the counterfactual decompositions using a QR approach provide a more complete picture of how the contributions of the star and the movie characteristics might differ at different levels of the theater and revenue distributions, and therefore a QR approach is of greater managerial significance.

In summary, this paper adds to prior star power literature in three areas. First, we control for the star selection bias as well as other endogeneities (e.g., budget and theater allocations) commonly present in the movie research when estimating the theater and revenue models. Second, we take a distributional perspective and employ QR to investigate the star effects across the entire movie outcome distributions. Finally, we counterfactually decompose the star versus nonstar

[1] Although opening week performance presents an appropriate setting for assessing star effect (because other experience-based quality signals, e.g., word of mouth, or adaptive adjustments of theaters are not in play), a more complete assessment of star effect might include the performance of a movie in its entire theatrical run, international sales, and sales of home entertainment products such as DVDs, etc. We leave these examinations for future research.

[2] As we describe later, to check for robustness of the results produced by the endogenous switching model, we also employ an analogous method in which the first stage model produces a control function that is included in a pooled sample of star *and* nonstar movies, with an endogenous star dummy and all its interactions with other covariates.

movie outcome differences into movie characteristics effects and star return effects, after controlling for other endogeneities. We now present a brief literature review and a conceptual framework for our research.

2. Literature Review and Conceptual Framework

2.1. Selected Literature

In general there are two themes in researching star effects. The first addresses the questions of why studios appoint a star and why stars agree to appear in a movie. Ravid (1999) investigates two competing economic rationales for star–movie matching. The first is a signaling explanation in which risk-averse studio executives employ stars as a signal for movie quality. The second is a "rent capture" argument in which stars extract most of their "worth," thereby increasing the movie budget and reducing the ROI. Based on mean comparison, Ravid finds that movie revenues of star movies are significantly greater than those of nonstar movies but that, once the movie budget is introduced in the regression model, the effects of stars on domestic, international, and video revenues become nonsignificant. Thus, supporting the "rent capture" hypothesis, Ravid finds that stars do not contribute to movie ROIs.

The second stream of research investigates the effects of having a star on different outcomes of a movie during its theatrical run in the United States as well as in international markets. Focusing on number of screens, revenue, and profits, De Vany and Walls (1999) propose a method to investigate whether the employment of stars shifts the respective outcome *distributions*. The study shows that the star's involvement significantly increases the number of theaters allocated to a movie throughout its theatrical life, and that this effect becomes even more pronounced in the later weeks. Similar to Ravid (1999), they also find that having a star does not raise the probability of the movie being a "hit," and the positive effect of stars on profitability is limited to only a handful of stars. This result leads De Vany and Walls (1999) to state that "the booking agents do not always share the tastes or perceptions of the audience" (p. 305), which points to the need for distinguishing star effects on exhibitors from the effects on the audience. The authors also recognize the issue of stars securing "better projects and bigger budgets," (p. 310), leading to problems of establishing causality.

Going beyond the geographic boundary, Neelamegham and Chintagunta (1999) find that the presence of stars has a positive and significant impact on the viewership in the U.S. market, but less so in other countries. To investigate the simultaneous and dynamic relationship between movie revenue and screen allocations, Elberse and Eliashberg (2003) model the supply (i.e., number of screens allocated for the movie) and demand (i.e., movie revenue) in which screening decisions are based on exhibitors' adaptive expectations of audience demand, and the viewing behavior of the audience depends on the number of screens allocated for the movie. Their study shows that, in the U.S. market, the presence of a star cast has no effect on the number of screens but has a significant positive effect on opening week revenue. Stars are found to have significant positive effects on screening decisions in France, Germany, and Spain but have no effect on revenue in any of the foreign markets.

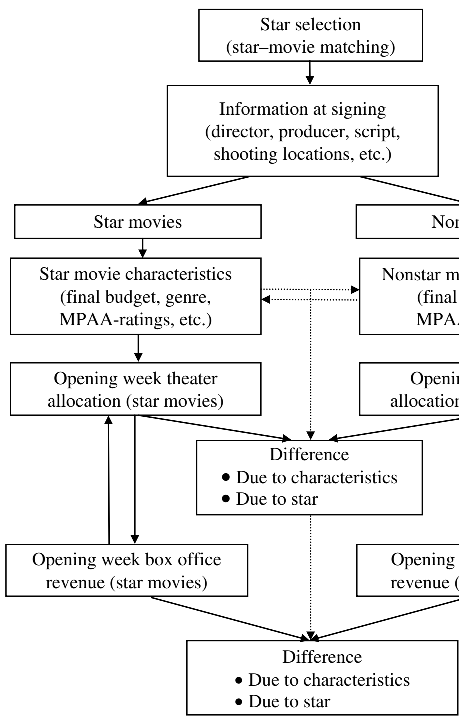
In an effort to control for the effects caused by the screen–revenue interplays and the star's ability to secure promising projects, Elberse (2007) employs an event study methodology to examine the effect of casting announcements on the stock prices (rather than revenue or screens) of movies in HSX Movie Stocks, a simulated stock market for movies. The study shows that the announcement of a star cast increases simulated movie stock prices on the day of announcement and thereafter during the event window, although the average cumulative abnormal returns (ACAR) vary across stars. We should note that the focus here is the investor reactions to casting announcements and not the revenue and costs of having a star.

A limitation of past research is that it designates a movie as a star movie or a nonstar movie without considering how a star is associated with a movie in the first place. Although several researchers have noted the existence of a reverse causality problem arising from a star's receiving preferential movie offers from studios and a star's proclivity to choose movie projects that have the elements needed for success (De Vany and Walls 1999), this problem has not been adequately addressed in past research. Elberse (2007) states this problem as follows: "It is difficult to control for a selection bias introduced by the possibility that the most powerful stars ... are able to choose the most promising movie projects ... *and* [emphasis added by the authors] to isolate the effect of an individual actor or actress on a movie's performance ..." (p. 110). We seek to address this gap by explicitly modeling the star selection process and counterfactually separating the star effects from the effects of the characteristics of the movies in which stars (versus nonstars) appear. The conceptual model that guides our study is presented in Figure 1.

2.2. A Conceptual Framework

As noted, engaging a star in a movie involves a bilateral decision of the studio and the star. Although this is a critical first step in the ultimate success or failure of a movie, not much is known about the star-contracting process, presumably because the negotiations take place privately between the studio and the "talent agents" representing the stars. From the studio's perspective,

Figure 1 A Conceptual Framework for Investigating Star Effects



despite the high cost of engaging a star, stars help to secure financing for the project and serve as a risk-mitigation strategy for the studios (Eliashberg et al. 2006, Vogel 2007, Basuroy et al. 2003). Indeed, it is not unusual to have a script written with a specific star in mind. From the star's perspective, talent agents negotiate with the studios to secure the best terms possible for their clients (Vogel 2007). Stars may prefer a movie with reputable producers, one that has the backing of a large production company, or one directed by an award-winning director. Stars may also avoid projects that require shooting in foreign or demanding locations or movies with the ingredients of an R rating. Anecdotally, Mel Gibson turned down the lead role of Maximus in *Gladiator* because of its demanding shooting schedule in foreign locations; the role landed Russell Crow an Oscar and made him a star. Conversely, Julia Roberts accepted her role in *Pretty Woman* only after Molly Ringwald, then a nonstar, turned down the offer because she was not comfortable with the script's tone and content.

Beyond the star-contracting process, stars may also use their clout to influence certain movie characteristics (Elberse 2007, Lehmann and Weinberg 2000, Vogel 2007), which may result in systematic differences in the star and nonstar movie characteristics. As noted, star movies on average have higher budget and receive wider distributions. Stars may also demand special effects or script changes to alter the movie genre or to obtain certain MPAA ratings. Reportedly, Angelina Jolie wanted changes to the script of *Salt* to make it a more action-oriented movie. To quote the Oscar-winning actor Michael Caine, "The difference between a movie star and a movie actor is this—a movie star will say, 'How can I change the script to suit me?' and a movie actor will say, 'How can I change me to suit the script?' "[3] When the characteristics of the movie are influenced by a star, the star effect becomes intertwined with the effects of the movie characteristics. A method utilized in labor economics to disentangle the two sources of difference is to employ a counterfactual decomposition approach (e.g., see Albrecht et al. 2003, Arulampalam et al. 2007). Applied in our context, the procedure empirically swaps the star and nonstar movie characteristics so that the star and nonstar movie outcome differences can be decomposed into those caused by the movie characteristics and those by the star.

As noted, our focus is on the star effects on the opening week theater counts and revenues. The theater allocation decision is made in expectation of revenues, and revenues are also influenced by the number of screens on which the movie opens (Elberse and Eliashberg 2003). Since past revenue information is unavailable in the opening week, we invoke rational expectation (Muth 1961), which postulates that distributors can accurately forecast the actual revenue and that there is no systematic over- or underallocation of theaters. Distributors utilize available information on factors that may influence revenue (e.g., budget, other movie characteristics) to determine the opening week number of theaters. This framework leads to the opening week theater and revenue models and the star selection model, which are presented next.

3. Models with Star Selection

3.1. Theater and Revenue Models

As noted, theatrical releases are based on revenue expectations, which are based on movie budget, the presence of a star in the movie, and the usual movie characteristics known at the time of the opening week screening decision (Eliashberg et al. 2006, Sorenson and Waguespack 2006). We therefore specify the opening week theater model for movie ij as follows:

$$\mathrm{Log}(Open\_Theater_{ij}) = \alpha_{0j} + \alpha_{1j}\mathrm{Log}(Budget_i) + \alpha_{2j}X_{ij} + \varepsilon_{ij}. \quad (1)$$

where $i =$ Movie, $1, 2, \ldots, N_j$; $j = 1$ if the movie contains a star, and 0 otherwise; X_{ij} denotes a vector of characteristics of movie ij that includes major studio, Friday, sequel, seasonality, genre, and MPAA ratings;

[3] See http://www.imdb.com/name/nm0000323/bio. To the extent that star influences exist, they occur in private, leaving us with only anecdotal support for the possibility that star and movie characteristics are not entirely independent.

and α_{1j} and α_{2j} denote respective returns on budget and movie characteristics that nonstar and star movies generate.

Likewise, our opening week revenue model is specified as follows:

$$\mathrm{Log}(Open\_Revenue_{ij}) = \beta_{0j} + \beta_{1j}\mathrm{Log}(Open\_Theater_i) + \beta_{2j}W_{ij} + \phi_{ij}. \quad (2)$$

where β_{1j} and β_{2j} denote respective returns on opening week theaters and movie characteristics that nonstar and star movies generate; and W_{ij} is similar to X_{ij} except that this vector additionally includes critics' reviews, usually available at the time of the movie's theatrical release, as an additional covariate that may influence the audience, hence revenue.

Budget effect on revenue comes via screen allocations (Elberse and Eliashberg 2003). As previously noted, the parameters in Equations (1) and (2) for each quantile of the *Open_Theater* and *Open_Revenue* distributions are estimated by using quantile regression (Koenker and Bassett 1978).[4]

3.2. Star Selection Model

Ordinarily, the returns on movie characteristics and budget (or opening week theater allocation) could be estimated by running quantile regressionsfor Equations (1) and (2) separately for the star and nonstar movie subsamples. Such an approach, however, may lead to inconsistent and biased estimates because the assignment of a star (or conversely a nonstar) to a movie is nonrandom. Specifically, when the effect of a covariate in the movie outcome equation is estimated with the star movie sample, the nonstar movies (which may either not have been offered to a star or have been rejected by stars while in the preproduction or script stage) are excluded and vice versa. Such a nonrandom sample caused by the selection might lead to a biased estimation of the covariate effects, unless a correction is made that considers the factors influencing a star's decision to participate in a movie. We employ a two-step Heckman (1979) correction in an endogenous switching regression framework. The first step is the star selection model that produces the correction terms (functions of the inverse Mills ratio) for each observation in the star subsample and nonstar subsample. These corrective terms are then added to the movie theater allocation and revenue equations for the star and nonstar samples in the second stage.[5]

The star selection model is specified via a latent variable of the following form:

$$E_k^* = Z_k\gamma + \eta_k, \quad (3)$$

where E_k^* is the latent utility that drives an actor to participate in movie project $k = ij$ ($j = 1, 0$; $i = 1, \ldots, N_j$), and Z_k is the movie-related information available at the time that may make the movie project attractive to a star; η_k includes the unobservables. We only observe whether a movie involves a star cast or not (i.e., $E_k = 1$ or 0). A star is postulated to accept a movie project when E_k^* exceeds a certain threshold, i.e., a reservation utility for the star. We estimate Equation (3) using the semiparametric likelihood approach of Klein and Spady (1993).

Because Equations (1) and (2) are estimated by using quantile regressions, the selection corrections are made at each quantile. Also, we allow for possible nonlinearity in the star–movie selection bias by including higher-order terms of the inverse Mills ratios, $\lambda_{i1}(\gamma) = \phi(Z_{i1}\gamma)/\Phi(Z_{i1}\gamma)$ and $\lambda_{i0}(\gamma) = -\phi(Z_{i0}\gamma)/[1-\Phi(Z_{i0}\gamma)]$ (where ϕ and Φ are the probability density function and cumulative distribution function of the standard normal distribution, respectively) for each movie in the star and nonstar group (see Bollinger et al. 2011 and Buchinsky 1998 for similar treatments). Based on Equation (3), we specify the conditional θth quantile of the error term in Equations (1) and (2), respectively, as

$$\mathrm{Quantile}_\theta[\varepsilon_{ij} \mid X_{ij}, \mathrm{Log}(Budget_{ij})] = \sum_{m=1}^{M} \delta_{mj}^{\theta}\lambda_{ij}^{(m)}(\gamma), \quad (4)$$

$$\mathrm{Quantile}_\theta[\phi_{ij} \mid W_{ij}, \mathrm{Log}(Open\_Theater_{ij})] = \sum_{m=1}^{M} \eta_{mj}^{\theta}\lambda_{ij}^{(m)}(\gamma). \quad (5)$$

To operationalize the model, we first estimate the selection model, which yields estimated parameter $\hat{\gamma}$. In the second step, we construct the terms in

[4] In a quantile regression, the conditional θth quantile of the dependent variable (e.g., the opening week revenue of a movie will be greater than 60% of the movies in the sample) is expressed as a (linear) function of the independent variables such as budget, MPAA ratings, etc. The coefficients of the θth regression quantile are estimated by minimizing the weighted sum of the absolute residuals, where positive residuals receive a weight of θ and negative residuals receive a weight of $(1-\theta)$. Because the number of positive residuals is greater for quantiles below the median, these receive greater weights than negative residuals, and the opposite is true for quantiles above the median. Because the weighted sum of residual is not differentiable, the linear programming method is used to obtain the estimates.

[5] The Heckman correction in an endogenous switching regression framework has been widely utilized in labor economics to generate segment-level estimates, where the population decides to participate in one or the other segment. For instance, the method has been utilized to investigate wage differentials for union and nonunion workers (Lee 1978) and for private and public employees (Gyourko and Tracy 1988, Heitmueller 2006) and those due to migration (i.e., "stayers" versus "movers") in Appalachia (Bollinger et al. 2011). In marketing, the endogenous switching regression framework has been employed to investigate strategy-specific covariate effects on stock returns when a firm decides to employ either a "proactive" or a "passive" product recall strategy (Chen et al. 2009).

Equations (4) and (5) with the estimated first-stage parameters and append them to Equations (1) and (2). However, because we are interested in the star effect across the whole distribution instead of at the mean level, we estimate selection-corrected conditional quantiles, proposed by Buchinsky (1998), as follows:

$$\mathrm{Log}(Open\_Theater_{ij}) = \alpha^{\theta}_{0j} + \alpha^{\theta}_{1j}\log(Budget_{ij}) + \alpha^{\theta}_{2j}X_{ij} + \sum_{m=1}^{M}\delta^{\theta}_{mj}\lambda^{(m)}_{i}(\gamma) + \mu_{ij}, \tag{6}$$

$$\mathrm{Log}(Open\_Revenue_{ij}) = \beta^{\theta}_{0j} + \beta^{\theta}_{1j}\log(Open\_Theater_{ij}) + \beta^{\theta}_{2j}W_{ij} + \sum_{m=1}^{M}\eta^{\theta}_{mj}\lambda^{(m)}_{i}(\gamma) + \nu_{ij} \tag{7}$$

for each quantile θ and on the sample of star and nonstar movies that yields the coefficients $(\hat{\alpha}^{\theta}_{j}, \hat{\beta}^{\theta}_{j}, \hat{\delta}^{\theta}_{j}, \hat{\eta}^{\theta}_{j})$, where $\alpha^{\theta}_{j} = (\alpha^{\theta}_{0j}, \alpha^{\theta}_{1j}, \alpha^{\theta}_{2j})$ and others are defined similarly. To capture the heterogeneity in the distribution of dependent variables, we estimate Equations (6) and (7) for 99 quantiles ranging from 0.01 to 0.99. We set $M = 2$ in the series estimators in Equations (4) and (5) to achieve reasonable accuracy without making the model too complicated (Bollinger et al. 2011).

In the switching regression framework, the information about the nonstar movies is included in the examination of the "target" star movie equation, and vice versa. When estimated with star movies (as in Equation (6) or (7), with $j = 1$), the parameter estimates can be interpreted as the hypothetical returns that a star can generate for the specific characteristics on *any* movie; this includes the actual star movies as well as all *potential* star movies that actually turned out to be nonstar movies, controlling for stars endogenously selecting in which movies to act. Similarly, when estimated with nonstar movies (as in Equation (6) or (7), with $j = 0$), the coefficients are the hypothetical returns that nonstars can generate on the corresponding covariates for all possible nonstar movies (which includes movies that actually turned out to be star movies), controlling for star selection bias.

An Alternative Approach. To check for robustness, we also employ another method that is analogous to the selection correction method described above. Here we estimate a single model for opening week theater allocation (and opening week revenue) with a star dummy and with all interactions of covariates with star. The star variable is suspected to be endogenous because the decision to accept a movie offer might be correlated with unobservables that may affect the movie's theater (or revenue) and therefore the respective error term. The endogeneity correction utilizes a control function approach (CF approach), which is similar to the Heckman sample selection correction. The first stage utilizes a probit model, which relates the variable suspected to be endogenous ($D = 1$ if the movie contains a star, 0 otherwise) to appropriate exogenous variables including instruments and the intercept (Z), as follows (ignoring subscripts):

$$D = \phi(Z\eta) + V, \tag{8}$$

where $\phi(\cdot)$ is the cumulative distribution function of a standard normal distribution. The residuals are used to construct a control function, which is then incorporated in the QR in the second stage. Lee (2007) derives that

$$P(Y \leq X\phi + DX\delta + \lambda_{\theta}(V) \mid X, Z) = \theta, \tag{9}$$

where Y is the second-stage dependent variable, i.e., Log(*Open_Theater*) or Log(*Open_Revenue*); X denotes movie characteristics, including Log(*Budget*) in the open theater model or Log(*Open_Theater*) in the open revenue model, plus 1; and V denotes the error in the first-stage star selection equation, and V is replaced by a residual from the first stage to construct the control function, $\lambda_{\theta}(V) = F^{-1}_{Y-X\phi-DX\delta \mid X, Z}(\theta) = F^{-1}_{Y-X\phi-DX\delta \mid V}(\theta)$.

Analogous to the selection correction method described earlier, $\lambda_{\theta}(V)$ is approximated by series approximation. As before, we retain the first two terms for simplicity. The terms $(\phi + \delta, \phi)$ (we suppress their dependence on θ) are the analogs of the two sets of coefficients, $(\alpha^{\theta}_{1}, \alpha^{\theta}_{0})$ or $(\beta^{\theta}_{1}, \beta^{\theta}_{0})$ in Equations (6) and (7), corresponding to the star group and nonstar group, and can be consistently estimated by running a partially linear quantile regression of Y against (X, D, and $\lambda_{\theta}(V)$).

In summary, although both methods described above are two-stage methods, the endogenous switching regression approach generates estimates for each segment with different inverse Mills ratios added to the separate star and nonstar outcome equations. The CF approach produces a single control function added to a single outcome equation with a star dummy and all possible star covariate interactions to estimate differential star returns on the covariates.[6] As will be presented later in §5, the two methods produce similar results.

3.3. Counterfactual Decomposition

Under the counterfactual decomposition approach, the stars and nonstars are "given" the *other group's* movie characteristics (i.e., drawn from their respective distributions) to determine how many theaters would be allocated and how much revenue would be generated if their *own* returns on these characteristics were

[6] We also considered the instrumental variable quantile regression (IVQR) methodology (Chernozhukov and Hansen 2005, 2006), which uses the method of moments generated with instruments in the QR model to account for endogenous explanatory variables. However, IVQR was difficult in this setting because we have several endogenous variables as well as all the star covariate interactions.

applied. For instance, let Y^S and Y^N be the revenues of a star movie (S) and a nonstar movie (N), respectively, and the difference between the two depends on the differences in the characteristics of star versus nonstar movies (i.e., X^S and X^N) as well as the return that a star (versus a nonstar) can generate on this characteristic (i.e., β^S and β^N). The difference in mean revenues of the two groups can be decomposed as follows (Blinder 1973, Oaxaca 1973):

$$\begin{aligned} E(Y^S) - E(Y^N) &= E(X^S)\beta^S - E(X^N)\beta^N \\ &= [E(X^S) - E(X^N)]\beta^S \\ &\quad + [E(X^N)(\beta^S - \beta^N)]. \end{aligned} \tag{10}$$

The first component is the characteristic effect, caused by the difference in the characteristics of the star (versus the nonstar) movies, and the second component is the star return effect, which is the incremental return that a star (versus a nonstar) can generate on the movie characteristic of a nonstar movie.

Machado and Mata (2005) and Melly (2005) extend the Oaxaca–Blinder mean decomposition method to achieve decompositions at different levels of the outcome distribution. They employ a sieve of quantile regressions to generate counterfactual distributions; in our case, these include (1) a distribution of a movie outcome that would arise if a star were to appear in a movie with the same characteristics as a nonstar movie but produced the same returns on these characteristics that a star would, and (2) a distribution of a movie outcome that would arise if the star were to appear in a movie with the characteristics of a star movie but produced returns on these characteristics that a nonstar would.

Formally, we denote $\hat{Q}_\theta(X^S, \beta^S_\theta)$ as the unconditional θth quantile of the distribution of the outcomes of the movies that have the star movies' characteristics profile (distribution) and the star movies' returns on the characteristics. Similarly, $\hat{Q}_\theta(X^N, \beta^N_\theta)$ is defined for the nonstar movies and nonstar returns. Finally, $\hat{Q}_\theta(X^S, \beta^N_\theta)$ represents the θth quantile of the counterfactual distribution of the revenues of the movies that have the star movies' characteristics profile but have returns resembling those of nonstar movies. Our desired decomposition can be expressed as the following:

$$\begin{aligned} &\hat{Q}_\theta(X^S, \beta^S_\theta) - \hat{Q}_\theta(X^N, \beta^N_\theta) \\ &\quad = [\hat{Q}_\theta(X^S, \beta^S_\theta) - \hat{Q}_\theta(X^S, \beta^N_\theta)] \\ &\qquad + [\hat{Q}_\theta(X^S, \beta^N_\theta) - \hat{Q}_\theta(X^N, \beta^N_\theta)]. \end{aligned} \tag{11}$$

Again, the first part of the right-hand side of the above equation is associated with the differences in the coefficients for the two groups receiving the same movie characteristics, i.e., the star return effect at the θth quantile. The second part of the right-hand side of the equation is the characteristic effect at the θth quantile, i.e., the effect due to the differences in the characteristics between the star and the nonstar movies. In this paper, we use the decomposition method of Melly (2005). A technical summary of the econometric methods for MM decomposition can be obtained from the authors upon request.

4. Variable Definitions, Data, and Descriptive Statistics

Our data sample includes 1,010 movies, and it consists of all movies that (1) were released between 2001 and 2005 in the United States and (2) have publicly available production budget and box office revenue data.[7] During 2001–2005, a total of 2,607 movies were released in the United States (http://www.the-numbers.com). Thus, our sample covers about 39% of these movies, which compares quite favorably with prior empirical work.

4.1. Variables in the Star Selection Model

At the time a star is offered a movie project, only limited information is available. Nevertheless, "for funding to be obtained, a project must already be outlined in terms of story line, director, producer, location, cast, and estimated budget" (Vogel 2007, p. 116). Accordingly, we speculate that the involvement of powerful production companies or recognized producers would impact a star's participation decision. Recognized producers or large production companies provide assurance that the resources needed for the making and distribution of the movie will be available. The involvement of an eminent director may be another key consideration for a star. Beyond attracting public attention, a recognized director may provide assurance to the star of a higher-quality movie. In terms of story line, stars might shy away from movie scripts that contain too much violence or sexual content that could result in the movie's being rated R by the MPAA, leading to more restricted releases and hence less revenue. Stars may also avoid foreign movies to conserve time and avoid shooting in foreign locations.

The variables included in the selection model and their respective operationalizations are (1) whether the movie is produced by someone who has been the producer of an Oscar Best Picture Award movie or who has won an Annual Producers Guild Award (*PowerProducer*), (2) whether the movie is directed by someone who has won the Oscar Best Director Award or an Annual Directors Guild Foundation Award (*PowerDirector*), (3) whether the movie is produced by

[7] We cross-reference http://www.imdb.com, http://www.the-numbers.com, and http://www.boxofficemojo.com to include all the movies meeting these criteria.

the "big six" production companies or their affiliates[8] or by some production companies who produced the Oscar Best Picture Award movies (*PowerProduction*), (4) whether the movie is a foreign movie (*Foreign*), and (5) whether the movie is rated R (assuming that the star can accurately predict if MPAA will rate the movie as R). In addition, to account for unobserved factors influencing a star's selecting a movie, we construct an instrument-like variable that captures whether stars flock to (or avoid) certain movie genres, measured by the average numbers of stars in the same genre movies as the focal movie in our sample (*IV-Star*).

4.2. Variables in the Theater and Revenue Models

For each of the movies in our sample, we collect both the number of opening week theaters and opening week box office revenue from multiple sources duly cross-referenced. Because our movie sample spans a five-year period from 2001 to 2005, we adjust the box office revenue and production budget data by the Consumer Price Index published by the U.S. Bureau of Labor Statistics.

Open_Theater Model. We include the following movie characteristics as explanatory variables: (1) whether the movie is a *sequel*; (2) whether the movie is released on *Friday*; (3) MPAA rating of R (following Basuroy et al. 2006 and Karniouchina 2011, we only differentiate rated R movies from the rest); (4) whether the movie belongs to one of the six major movie genres (*action* (base), *adventure*, *comedy*, *drama*, *romance*, and *thriller*); (5) *seasonality* (following Elberse and Eliashberg 2003 and Luo et al. 2010, we use the seasonality index, which is based on the aggregate weekly U.S. box office revenues from 1969 to 1984 (Vogel 2007), to control for the seasonal changes of box office revenue (Einav 2007)); (6) *major studios* (to account for "studio power," a phenomenon whereby a movie produced by a "major" studio has a wider distribution than a movie that is not, we collect the name of the distributors and use a dummy variable to indicate whether the movie is distributed by a major studio (Elberse and Eliashberg 2003)).

With respect to *Budget*, we use data on production budget because we do not have data on marketing budget for all the movies in our sample. Reliable information on marketing budget allocated for post-production activities is difficult to obtain from public sources (Ravid 1999, p. 470, Footnote 19). Besides, advertising budgets are usually fixed proportion of the production budget (Chintagunta et al. 2010, Vogel 2007). *Budget* is suspected to be endogenous because certain unobserved factors such as special effects, exotic locations, marketing expenditures, etc., that are correlated with budget may also affect the expected movie revenue, which in turn may influence the number of theaters, hence the error term in the theater equation (Elberse and Eliashberg 2003, Shugan 2004).

We use the *average* of the log-transformed production budget of the movies in the same genre and with the same MPAA rating as the focal movie (for separate star movie and nonstar movie subsamples) as the instrumental variable for Log(*Budget*) in the *Open_Theater* equation. The rationale behind this choice of instrument is analogous to that provided by Chintagunta et al. (2010). The budgets of movies in the same genre and with the same MPAA rating are likely to have similar magnitudes and hence are correlated with each other. However, the budgets of other movies are not likely to be correlated with the focal movie's idiosyncratic characteristic, or the error term.

Open_Revenue Model. In addition to the variables in the open theater model, we include critics' reviews in the revenue model. We use "*Metascore*" (http://www.metacritic.com), a weighted average of the individual critic's scores[9] for each movie on a 0–100 point scale, which is then normalized to a 0–1 point scale. Despite the lack of transparency about how it computes its scores, *Metascore* has been widely used in the movie literature as a measurement for critics' reviews (Hennig-Thurau et al. 2006, 2009; Kamakura et al. 2006). Because not all movies are rated on the metacritic.com website, we only have 916 observations on this variable.

We treat *Open_Theater* in the revenue model as endogenous because the number of theaters allocated is based on revenue expectation, which in turn is based on unobserved characteristics of the movie. Similar to the instrumental variable for *Budget*, we use the average of the log-transformed opening week theater numbers of the movies in the same genre and with the same MPAA rating as the focal movie in our subsamples (star movies and nonstar movies) as the instrumental variables for Log(*Open_Theater*) in our open revenue equation.[10] We do not include budget in the revenue equations because the budget's impact on revenue occurs indirectly through screen allocation (see Elberse and Eliashberg 2003).

[8] We define "big six" according to the article "Major film studio" (http://en.wikipedia.org/wiki/Major_film_studio, accessed June 2012).

[9] More details on the source and calculation of *Metascore* can be found at http://www.metacritic.com/about-metascores (accessed May 2012). The website states that it "...carefully curate(s) a large group of the world's most respected critics, assigning scores to their reviews, and applying a weighted average to summarize the range of their opinions."

[10] Although literature has identified movie budget and theaters as potentially endogenous, there may be other variables that could be endogenous. Because of the difficulty in finding appropriate instruments for all potential endogenous variables and in gathering related data, we restrict our analysis to the two main endogenous variables identified in the literature.

4.3. Operationalization of "Star"

The definition of a "star" is central to the estimations of the theater and revenue models as well as the counterfactual decomposition analysis. Past research has used different criteria, such as the actor/actress's box office history, award history, or appearances in the industry's "power" or "hit" lists. Although identification of the sources of stardom is in itself a fertile area of inquiry, we rely on published lists of stars that are well-recognized in the industry. When compiling the star power list, the editors of these publications usually evaluate stars in a comprehensive manner that goes beyond the star's box office and awards history. In this study, we use the annual trade magazine *Premiere*'s "Power 100 list" (De Vany and Walls 1999, 2004; Liu 2006; Walls 2009), published typically in the May or June issue and listing Hollywood's most powerful people, including prominent actors/actresses, directors, producers, agents, and lawyers. Any actor/actress who appears on that year's list is regarded as a "star" for the rest of that year *after the list's publication* and the year after that. We consider an actor/actress as a "star" for two successive years after which they lose the status unless they are relisted as a star during this period.

There are 71 movie actors/actresses in total who make up our final *Premiere* star lists between year 2001 and 2005. Some of the stars whose names appear several times are Adam Sandler, Ben Affleck, Ben Stiller, Cameron Diaz, George Clooney, Matt Damon, Nicolas Cage, and Renée Zellweger. If a movie is released during an actor/actress's tenure as a star, the movie is classified as a star movie. This procedure produces 233 star movies (of which 38 movies have multiple stars[11]) and 777 nonstar movies in our data. Some of the successful star movies include *Mr. and Mrs. Smith* (Brad Pitt), *Pearl Harbor* (Ben Affleck), *Signs* (Mel Gibson), and *Bruce Almighty* (Jim Carrey). Successful nonstar movies include *Harry Potter and the Sorcerer's Stone* (Daniel Radcliffe had not yet appeared in the star list), *The Lord of the Rings*: *The Fellowship of the Ring*, *Spider Man*, and *Star Wars Episode* 3.[12]

4.4. Descriptive Statistics

Table 1 provides the variable operationalizations, the data sources, and the summary statistics for these variables, broken down by stars and nonstars. As conjectured, the proportions of movies with stars relative to those without a star are greater for movies directed by award-winning directors, produced by award-winning producers, and produced by big six production companies. Also as expected, there are fewer foreign movies with stars than with nonstars.

The average budget of star movies is twice as large as that of nonstar movies, which is consistent with the data reported by Ravid (1999). The highest-budget star movie in our sample is *King Kong* (Jack Black, Naomi Watts: $207 M) and the lowest-budget star movie is *Lisa Picard is Famous* (Sandra Bullock: $1 M). The highest-budget nonstar movie is *The Chronicles of Narnia: The Lion, the Witch, and the Wardrobe* ($180 M), and the lowest is a war documentary, *Return to the Land of Wonders* ($5,000).

The mean number of theaters in the opening week for star movies is 2,356 (median: 2,711) and the same for nonstar movies is 1,470 (median: 1,621). The star movies in the sample with the largest and the smallest numbers of opening week theaters are *Shrek* 2, with 4,163 theaters and *Lisa Picard is Famous*, that opened in only one theater. The nonstar movies with the largest and the smallest numbers of opening week theaters are *The Incredibles* (closely followed by *Harry Potter and the Goblet of Fire*) with 3,933 theaters, and *Return to the Land of Wonders*, opening in one theater. The mean opening week revenue for star movies is $18.2 M (median: $13.1 M) and that for nonstar movies is $9.7 M (median: $4.7 M). The star movies in the sample with the largest and the smallest opening week revenues are *Shrek* 2 ($108 M) and *Lisa Picard is Famous* ($14,129). The nonstar movies with the largest and the smallest opening week revenues are *Spiderman* ($115 M) and *The Dark Hours* ($423).

Other descriptive statistics in Table 1 suggest that major studios tend to employ more stars. There are greater proportions of movies in the action and adventure genres for the star movie subsample, and a smaller proportion of drama and thrillers. Release dates, sequels, critical ratings, and seasonality index are not much different between the star and nonstar movies in our sample. Overall, the movie sample covers a wide range of budgets, theater allocations, and revenues.

4.5. Distributions of the Dependent Variables

In Table 2, we present the distributions of the number of open theaters and opening week revenue for all movies, nonstar movies, and star movies. As suspected, both distributions are skewed, with open theater mean and median for all movies being 1,674 and 2,038, respectively, and open revenue mean and median being $12 M and $7 M, respectively. A large number of movies (about 20% of the sample), mostly made up of nonstar movies, are allocated ≤ 10 theaters. More

[11] Because the number of multistar movies is so small, we did not create a separate category for analysis. Besides, creating a subcategory within the set of star movies would produce a different number of variables in the star and nonstar movie samples, which is not compatible with the counterfactual analysis.

[12] To check for robustness of our star definition, we also consulted another authoritative star power list, the annual "Celebrity 100" list published by *Forbes* (e.g., see Joshi and Hanssens 2009). This star list produced nearly identical distributions of the dependent variables and similar QR estimates and decompositions. These results are available from the authors. All subsequent reporting is based on *Premiere* star classification.

Table 1 Descriptive Statistics

			Nonstar movies ($n = 777$)		Star movies ($n = 233$)	
Variable	Description	Data source	Mean	SD	Mean	SD
PowerDirector	Whether the movie's director had previously won awards	Calculated based on data from imdb.com	0.08	0.28	0.22	0.41
PowerProducer	Whether the movie's producer had previously won awards	Calculated based on data from imdb.com	0.15	0.36	0.34	0.47
PowerProduction	Whether the movie is produced by the "big six" production companies/affiliates	Calculated based on data from imdb.com	0.47	0.50	0.76	0.45
Foreign	Whether the movie is a foreign movie	imdb.com	0.23	0.42	0.16	0.37
Budget	Production budget	the-numbers.com	$28 M	$30 M	$58 M	$36 M
Open_Theater	Total numbers of theaters in the opening week	boxofficemojo.com	1,470	1,297	2,356	1,146
Open_Revenue	Box office revenue in the opening week	boxofficemojo.com	$9.7 M	$15.4 M	$18.2 M	$17.6 M
Major_Studio	Whether the movie is released by a major studio	http://en.wikipedia.org/wiki/Major_film_studio	0.55	0.50	0.64	0.48
Friday	Whether the movie is release on Friday	boxofficemojo.com	0.86	0.35	0.82	0.38
Sequel	Whether the movie is a sequel	imdb.com	0.07	0.25	0.08	0.27
Metascore	Average critical rating of the movie	metacritic.com	51.78	18.71	51.83	15.81
Seasonality	Seasonality index	Vogel (2007)	0.63	0.14	0.65	0.15
Action	Whether the movie genre is action	the-numbers.com	0.09	0.29	0.16	0.37
Adventure	Whether the movie genre is adventure	the-numbers.com	0.09	0.29	0.12	0.32
Comedy	Whether the movie genre is comedy	the-numbers.com	0.26	0.44	0.29	0.46
Drama	Whether the movie genre is drama	the-numbers.com	0.35	0.48	0.25	0.44
Romance	Whether the movie genre is romance	the-numbers.com	0.06	0.24	0.08	0.27
Thriller	Whether the movie genre is thriller	the-numbers.com	0.15	0.36	0.10	0.30
R	Whether the movie is rated R	the-numbers.com	0.40	0.49	0.36	0.48

than 70% of star movies were allocated 2,000 or more theaters. The revenue distribution is also skewed, with 32% of all movies, mostly nonstar movies, having less than $1 million in revenue in the opening week.

In the bottom part of Table 2, we present the means, and the standard deviations of the log-transformed dependent variables (DVs) (i.e., *Open_Theater* and *Open_Revenue*), as well as the 10th, 25th, 50th, 75th, and 90th quantile statistics. We also include a measure of the dispersion, the 0.9–0.1 spread for both types of movies. First, we present the statistics for the full sample and the two subsamples (i.e., nonstar and star movies). Star movies have higher values than nonstar movies across whole distributions for both DVs, suggesting that star movies outperform nonstar movies on both metrics. Moreover, generally there are wider gaps between the star movies and the nonstar movies at the lower tail than those at the upper tail of the distributions, which is supported by the wider 0.9–0.1 spread for nonstar movies. Overall, the distributions of both DVs are skewed, which justifies the use of QR. We now report the results for star selection, followed by the QR and counterfactual decomposition results.

5. Results

5.1. Star Selection Model

Table 3 provides the results of the first-stage star selection equation. Overall, the model is significant. All the regressors have significant coefficients with expected signs. The involvement of an award-winning director and producer positively influences stars' choice of movie projects. Stars are also attracted to movies produced by the big six production companies. However, the likelihood of a star's accepting a movie decreases if a movie is a foreign (with respect to the United States) movie or if it is expected to receive an R rating by the MPAA. Interestingly, as indicated by the significant *IV-Star* parameter, stars, as a group, tend to exhibit a propensity to accept (or avoid) movies of certain genres.

5.2. Star and Movie Characteristics Effects on Theater and Revenue

In Tables 4 and 5, we report the estimated covariate effects for the 15th, 25th, 50th, 75th, and 85th quantiles using the Heckman selection correction and the control function approaches (CF approaches), respectively. Recall that the former method produces separate estimates and the inverse Mills ratios for the star and the nonstar samples. It also includes control functions to correct the budget or open theater endogeneities. The latter method uses a pooled sample of star and nonstar movies with a fully interacted model, and a single set of control function estimates for the endogenous star variable and other endogeneities. For each table, the left panel shows the results for the *Open_Theater* equation, and the right for the *Open_Revenue* equation.

Open_Theater Results. The parameter estimates produced by the two methods for the nonstar movies are

Table 2 Distributions of the Dependent Variables

Panel A: Distributions of star and nonstar movie frequencies

No. of open theaters	Total movies ($N = 1{,}010$) (%)	Nonstar movies ($N = 777$) (%)	Star movies ($N = 233$) (%)
1	3	4	0
1–10	18	20	8
10–100	10	12	4
100–500	2	2	2
500–1,000	3	3	1
1,000–1,500	4	5	2
1,500–2,000	8	8	6
2,000–2,500	13	13	13
2,500–3,000	20	17	30
3,000–3,500	14	11	25
3,500–4,000	4	3	7
>4,000	0	0	2
Open revenue (million)	**Total movies (%)**	**Nonstar movies (%)**	**Star movies (%)**
<$1	32	38	13
$1–$2	2	2	2
$2–$4	6	7	3
$4–$8	14	14	13
$8–$16	21	20	26
$16–$32	15	12	25
$32–$64	6	4	15
>$64	2	2	3

Panel B: Distributions of log-transformed dependent variables

					Quantiles				
	n	Mean	SD	IQR	0.1	0.25	0.5	0.75	0.9
Full sample									
Log(*Open_Theater*)	1,010	5.9	2.77	6.69	1.39	3.09	7.62	7.94	8.08
Log(*Open_Revenue*)	1,010	9.41	2.6	6.51	5.45	7.22	10.48	11.32	11.96
Nonstar sample									
Log(*Open_Theater*)	777	5.55	2.86	6.66	1.39	2.4	7.39	7.87	8.05
Log(*Open_Revenue*)	777	9.03	2.68	6.52	5.19	6.6	10.16	11.12	11.71
Star sample									
Log(*Open_Theater*)	233	7.07	2.05	5.44	2.71	7.66	7.91	8.05	8.15
Log(*Open_Revenue*)	233	10.71	1.8	4.71	7.62	10.48	11.19	11.79	12.33

quite similar (left panels of Tables 4 and 5), and the star selection parameters (inverse Mills) are significant at all quantiles (Table 4), suggesting the presence of significant selection bias. The selection effect is concave, as indicated by the negative second-order inverse Mills parameter. For the pooled sample, the star control functions are significant at the 25th and the 50th (median) quantiles, and again the second-order control function estimates are negative.

Table 3 Star Selection Model

	Coefficients[a]	Std. error	Z	Prob > Z
PowerDirector	1.895	0.576	3.290	0.001
PowerProducer	1.654	0.518	3.190	0.001
PowerProduction	1.403	0.401	3.500	0.000
Foreign	−1.067	0.376	−2.840	0.005
R	−0.584	0.251	−2.330	0.020
IV-Star	8.679	3.000	2.890	0.004

Notes. $N = 1{,}009$; log-likelihood $= -494.5412$; Wald χ^2 (df $= 6$) $= 13.80$; Prob $> \chi^2 = 0.0319$. Parameters: DV—Star's presence (1) or absence (0) in the movie.

[a]The semiparametric maximum likelihood estimator (Klein and Spady 1993) is used to estimate the parameters.

For the nonstar subsample, the effects of budget on open theater are positive and significant at all levels of the theater distribution under both methods. However, the budget effect on theater allocation is greater at the lower quantiles than at the upper quantiles of the open theater distribution. The effect of production by major studios on opening week theater allocations of nonstar movies is positive and significant at all quantiles under the pooled analysis, although the effect declines at higher quantiles. For the separate sample analysis, again the major studio effect on open theater is stronger at the lower quantiles. Relative to the base (action) genre, the effects of *Comedy*, *Romance*, and

Table 4 QR Results with Star Selection and Budget/Theater Control Functions (Split Samples)

Nonstar movies	Log(*Open_Theater*) ($n = 715$)					Log(*Open_Revenue*) ($n = 696$)				
Percentage below →	15%	25%	50%	75%	85%	15%	25%	50%	75%	85%
Constant	**−13.82**	**−9.39**	**−11.51**	**−9.63**	**−3.75**	1.27	1.45	**3.54**	**3.36**	**3.81**
Log(*Budget*)	**1.21**	**1.03**	**1.19**	**1.12**	**0.78**					
Log(*Open_Theater*)						**1.01**	**1.01**	**0.91**	**0.95**	**0.99**
Major studio	**0.67**	**0.52**	**0.43**	0.12	0.07	0.07	0.06	0.09	0.07	0.02
Friday	0.29	0.10	**0.25**	0.31	0.14	−0.14	−0.26	**−0.31**	−0.11	−0.20
Sequel	**0.74**	0.37	0.03	−0.09	−0.05	**0.48**	0.36	**0.23**	**0.36**	0.27
Seasonality	−0.81	−0.22	−0.44	−0.19	−0.25	−0.02	−0.02	−0.21	0.16	0.01
Metascore						**0.03**	**0.03**	**0.02**	**0.02**	**0.02**
Adventure	0.48	−0.04	−0.15	0.25	0.24	0.16	0.18	0.14	**0.40**	0.36
Comedy	−0.07	0.11	**0.64**	**0.81**	**0.59**	0.13	0.20	0.04	−0.02	−0.06
Drama	**−1.39**	**−1.90**	**−1.44**	0.08	0.18	**0.41**	0.51	0.1	0.12	0.11
Romance	**1.32**	**0.74**	**0.8**	**0.82**	**0.53**	0.20	0.08	−0.06	−0.28	−0.19
Thriller	**1.30**	**0.65**	**0.92**	**0.98**	**0.67**	0.29	0.25	0.04	0.07	−0.06
R	**−1.33**	**−0.99**	**−0.19**	0.13	0.08	0.04	0.19	**0.18**	**0.19**	0.21
InvMills (*star*)	**7.67**	**5.98**	**6.63**	**5.35**	**2.90**	0.21	0.68	0.07	−0.20	−0.82
InvMills2 (*star*)	**−2.85**	**−2.39**	**−2.55**	**−1.98**	**−0.99**	−0.19	−0.42	−0.25	−0.06	0.24
CntrlFn (*budget/theater*)	−0.35	−0.23	**−0.34**	**−0.34**	**−0.30**	−0.06	−0.09	−0.04	**−0.13**	**−0.16**
CntrlFn2 (*budget/theater*)	**0.05**	−0.03	**0.08**	0.03	**0.03**	−0.01	0.00	0.00	**0.01**	**0.02**
Star movies	Log(*Open_Theater*) ($n = 215$)					Log(*Open_Revenue*) ($n = 217$)				
Constant	−10.34	−1.66	**2.9**	**5.61**	**4.47**	2.52	−0.12	1.80	**3.32**	2.47
Log(*Budget*)	**1.34**	**0.71**	**0.36**	**0.22**	**0.29**					
Log(*Open_Theater*)						**0.91**	**1.09**	**0.96**	**0.83**	**0.90**
Major studio	−0.33	−0.28	−0.04	−0.06	**−0.14**	−0.09	−0.01	**0.32**	**0.36**	0.25
Friday	0.46	0.34	0.02	−0.02	−0.01	−0.26	−0.23	**−0.25**	−0.09	−0.21
Sequel	0.86	0.42	0.13	0.11	0.02	0.44	0.33	0.29	0.38	0.13
Metascore			0.03			**0.03**	**0.03**	0.57	**0.02**	**0.03**
Seasonality	0.39	−0.14		−0.13	−0.15	0.56	0.75	**0.03**	0.20	0.42
Adventure	−1.15	−0.14	−0.07	0.04	−0.01	−0.57	**−0.58**	−0.33	−0.45	−0.35
Comedy	−0.44	−0.16	−0.02	0.05	0.07	−0.60	−0.36	**−0.34**	−0.36	−0.23
Drama	**−2.67**	**−2.67**	**−0.31**	−0.13	−0.06	−0.64	−0.34	**−0.39**	**−0.54**	−0.41
Romance	0.29	0.12	0.03	0.02	−0.02	−0.37	−0.19	−0.19	−0.34	−0.13
Thriller	0.02	0.45	**0.14**	0.03	−0.02	−0.22	−0.26	−0.28	−0.26	−0.23
R	−0.62	−0.22	**−0.24**	**−0.12**	**−0.10**	−0.32	−0.24	−0.2	**−0.26**	−0.23
InvMills (*star*)	0.70	1.18	**0.80**	−0.24	0.27	0.86	1.66	0.87	1.60	1.66
InvMills2 (*star*)	−0.32	−0.80	−0.30	0.10	−0.05	−0.46	−0.60	−0.27	−0.70	−0.59
CntrlFn (*budget/theater*)	0.15	0.61	0.09	0.06	−0.01	−0.11	−0.21	−0.15	−0.05	−0.19
CntrlFn2 (*budget/theater*)	−0.11	**−0.28**	**−0.11**	**−0.11**	**−0.08**	−0.01	0.00	0.01	0.01	−0.01

Notes. Entries in bold are statistically significant at $p < 0.05$. *InvMills* (*star*) and *InvMills*2 (*star*) are the first two terms of the series estimators correcting the star selection bias. *CntrlFn* (*budget/theater*) and *CntrlFn*2 (*budget/theater*) are the first two terms of the control function series estimators for correcting the endogeneity of budget and theater, respectively.

Thriller are positive at nearly all quantiles of theater distributions for nonstar movies. However, the drama genre is less widely distributed than the base genre. As expected, rated R movies are less widely distributed than other MPAA rated movies below the median.

We now consider the star effects on the opening week theater allocations (the bottom left panels of Tables 4 and 5). For the star main effects, the two methods produce somewhat different results, but the pattern is quite similar. The star effects are negative or nonsignificant for movies below the median but become positive and significant at the median and upper quantiles. Thus, stars do help in wider releases of movies at the median to upper quantiles of the theatrical releases. However, stars generally do not generate greater returns on the movie budget and other movie characteristics than do nonstars. In fact, relative to nonstar movies, star movies generate lower returns on budget for movies at the upper quantiles. Likewise, stars do not add to the effects of major studio on opening week theatrical releases. With regard to genre effects, the coefficients for the star sample (Table 4) are generally nonsignificant for romance, comedy, and thriller at all quantiles. In the pooled analysis, the star interactions with the comedy, romance, and thriller genres are generally negative and significant. Thus, overall, star movies are unable to derive any additional benefit from these genres to obtain wider releases relative to nonstar movies of the same genres.

Open_Revenue Results. Again, the results produced by the two methods are very similar (see the right panels of Tables 4 and 5). The inverse Mills ratio (and the star control function) parameters are mostly nonsignificant, suggesting that star selection bias is not a concern in

Table 5 QR Results with Star and Budget/Theater CF (Pooled Sample with Interactions)

	Log(*Open_Theaters*)					Log(*Open_Revenue*)				
Percentage below →	15%	25%	50%	75%	85%	15%	25%	50%	75%	85%
Constant	**−10.44**	**−7.14**	**−6.93**	**−5.48**	0.40	1.41	0.71	2.02	**2.10**	**2.68**
Log(*Budget*)	**1.14**	**0.87**	**1.01**	**0.98**	**0.59**					
Log(*Open_Theater*)						**0.98**	**1.02**	**0.97**	**0.93**	**0.97**
Major studio	**0.65**	**0.66**	**0.43**	**0.21**	**0.11**	0.07	0.04	0.03	0.06	0.00
Friday	0.33	0.10	**0.25**	**0.25**	0.05	−0.19	**−0.27**	−0.26	−0.13	−0.19
Sequel	0.77	0.34	0.09	−0.11	−0.04	**0.53**	0.34	0.15	**0.35**	0.33
Seasonality	−1.13	−0.25	−0.27	−0.09	**−0.22**	−0.05	0.04	0.09	0.17	0.05
Metascore						**0.02**	**0.03**	**0.02**	**0.02**	**0.02**
Adventure	0.33	−0.13	−0.11	0.09	**0.26**	0.10	0.10	0.15	**0.40**	0.48
Comedy	−0.15	0.00	**0.56**	**0.74**	**0.43**	0.04	0.11	0.06	−0.02	−0.10
Drama	**−1.64**	**−2.44**	**−2.16**	**−0.21**	−0.01	0.31	**0.43**	0.16	0.11	0.13
Romance	**1.47**	0.50	**0.73**	**0.70**	**0.38**	0.12	0.04	−0.06	**−0.29**	−0.22
Thriller	**1.13**	**0.55**	**0.81**	**0.71**	**0.43**	0.26	0.21	0.09	0.14	0.02
R	**−1.53**	**−1.18**	**−0.35**	−0.01	0.04	−0.04	0.09	0.13	0.12	0.16
Star	**−9.49**	−2.08	**5.62**	**6.99**	**2.16**	**2.01**	0.71	1.10	0.62	0.49
Budget ∗ *Star*	**0.75**	0.21	**−0.41**	**−0.53**	**−0.17**					
Theater ∗ *Star*						**−0.15**	−0.05	**−0.10**	**−0.06**	−0.05
Studio ∗ *Star*	**−1.06**	**−0.73**	**−0.46**	**−0.26**	**−0.21**	−0.10	0.00	0.24	0.23	0.31
Friday ∗ *Star*	0.31	0.16	−0.18	**−0.26**	−0.11	−0.04	0.06	−0.05	0.05	−0.03
Sequel ∗ *Star*	−0.04	−0.25	−0.06	0.15	0.07	−0.06	0.28	0.27	0.01	−0.13
Seasonality ∗ *Star*	1.73	−0.08	0.11	−0.08	0.20	0.52	0.53	0.41	0.03	−0.11
Metascore ∗ *Star*						0.00	0.00	0.00	0.00	0.00
Adventure ∗ *Star*	−1.71	−0.14	−0.05	−0.01	**−0.22**	−0.67	−0.58	−0.47	**−0.75**	−0.75
Comedy ∗ *Star*	−0.38	0.08	**−0.44**	**−0.64**	**−0.29**	**−0.71**	−0.53	−0.35	−0.24	−0.09
Drama ∗ *Star*	−0.84	−0.29	**1.99**	0.14	0.02	**−0.99**	**−0.78**	**−0.56**	**−0.64**	−0.54
Romance ∗ *Star*	−1.11	−0.36	**−0.63**	**−0.66**	**−0.32**	−0.59	−0.36	0.04	0.04	0.12
Thriller ∗ *Star*	−1.12	−0.20	**−0.68**	**−0.70**	**−0.41**	−0.42	−0.41	−0.23	**−0.40**	−0.22
R ∗ *Star*	0.78	**0.86**	0.16	−0.11	−0.09	**−0.37**	−0.37	**−0.34**	**−0.32**	−0.36
CntrlFn (*star*)	7.63	**10.27**	**4.77**	2.31	0.28	1.93	3.21	2.23	**4.06**	1.88
CntrlFn2 (*star*)	−5.34	**−7.76**	**−3.28**	−1.32	0.17	−2.02	−2.73	−2.13	**−3.35**	−1.49
CntrlFn (*budget/theater*)	−0.24	−0.02	**−0.17**	**−0.22**	**−0.14**	−0.01	−0.11	−0.08	**−0.09**	**−0.16**
CntrlFn2 (*budget/theater*)	0.03	−0.01	**0.05**	**0.02**	0.00	−0.01	0.00	0.00	**0.01**	0.01
N	930	930	**930**	930	930	913	913	913	913	913

Notes. Entries in bold are statistically significant at $p < 0.05$. *CntrlFn* (*star*) and *CntrlFn2* (*star*) are the first two terms of the control function series estimators for correcting the star selection bias. *CntrlFn* (*budget/theater*) and *CntrlFn2* (*budget/theater*) are the first two terms of the control function series estimators for correcting the endogeneity of budget and theater, respectively.

the revenue analysis. The endogeneity of open theater is present only at the upper quantiles. For the nonstar sample, three main results are worth noting. First, the open theater elasticity to revenue remains at about 1 at all quantiles of revenue distribution. Second, sequels have a significant and positive effect on the revenues of the nonstar movies, but only at the 15th and 75th percentiles. Third, the effect of *Metascore* is positive and significant and is nearly the same (0.02–0.03) at all levels of the revenue distribution, implying that critics' reviews affect opening revenue positively and uniformly across all quantiles.

To assess the effects of stars on the opening week revenue, we compare the differences between the intercepts of the split-sample analysis (Table 4) and examine the star dummy coefficients in the pooled analysis (Table 5). Both methods suggest that the star effect is nonexistent, except at the lowest quantile. There is also no evidence of stars producing any incremental returns on allocated theaters and movies characteristics. The theater–star interactions are either nonsignificant or even slightly negative. The presence of a star actually depletes the revenues of movies in the drama genre and those rated R.

In summary, star selection bias is present in the open theater analysis, requiring correction. The presence of a star positively influences the allocations of theaters in the opening week for movies at the median and the upper quantiles of the distributions. There is practically no star effect on opening week revenue. Also, there is no evidence that stars produce any incremental returns on movie budget, theatrical releases, or other movie characteristics in either the opening week theater allocations or audience-generated revenues. It is useful to note, however, that the analyses that produce these results do not address the question of whether the returns that stars (or nonstars) would have generated if the movie characteristics, including budget and theater allocations, of the two groups were swapped. The counterfactual decomposition helps to separate the star and the characteristics effects, which we discuss next.

Table 6 Decomposition QR with Star Selection Corrections

		Log(*Open_Theater*)			Log(*Open_Revenue*)		
Quantiles	Effect sources	Coefficients	Std. err.	*p*	Coefficients	Std. err.	*p*
0.1	Total difference	6.499	1.475	0.000	1.460	1.394	0.295
	Movie characteristic	1.204	0.488	0.014	2.461	1.351	0.069
	Star return	5.295	1.648	0.001	−1.001	0.942	0.288
0.2	Total difference	5.716	1.354	0.000	3.403	0.678	0.000
	Movie characteristic	0.946	0.353	0.007	4.409	0.669	0.000
	Star return	4.770	1.451	0.001	−1.005	0.858	0.241
0.3	Total difference	5.232	1.365	0.000	2.789	0.713	0.000
	Movie characteristic	0.857	0.305	0.005	3.718	0.594	0.000
	Star return	4.375	1.372	0.001	−0.928	0.786	0.238
0.4	Total difference	4.784	1.345	0.000	0.135	0.654	0.836
	Movie characteristic	0.788	0.241	0.001	1.134	0.332	0.001
	Star return	3.996	1.392	0.004	−0.999	0.641	0.119
0.5	Total difference	4.374	1.322	0.001	−0.223	0.606	0.713
	Movie characteristic	0.621	0.219	0.005	0.728	0.140	0.000
	Star return	3.753	1.378	0.007	−0.950	0.622	0.127
0.6	Total difference	4.111	1.304	0.002	−0.356	0.600	0.554
	Movie characteristic	0.524	0.198	0.008	0.605	0.118	0.000
	Star return	3.586	1.356	0.008	−0.961	0.612	0.117
0.7	Total difference	3.854	1.290	0.003	−0.436	0.603	0.470
	Movie characteristic	0.452	0.170	0.008	0.511	0.102	0.000
	Star return	3.402	1.330	0.011	−0.946	0.608	0.120
0.8	Total difference	3.639	1.273	0.004	−0.471	0.602	0.434
	Movie characteristic	0.380	0.136	0.005	0.431	0.091	0.000
	Star return	3.260	1.302	0.012	−0.902	0.605	0.136
0.9	Total difference	3.336	1.249	0.008	−0.515	0.598	0.390
	Movie characteristic	0.226	0.098	0.022	0.371	0.109	0.001
	Star return	3.11	1.265	0.014	−0.885	0.595	0.137

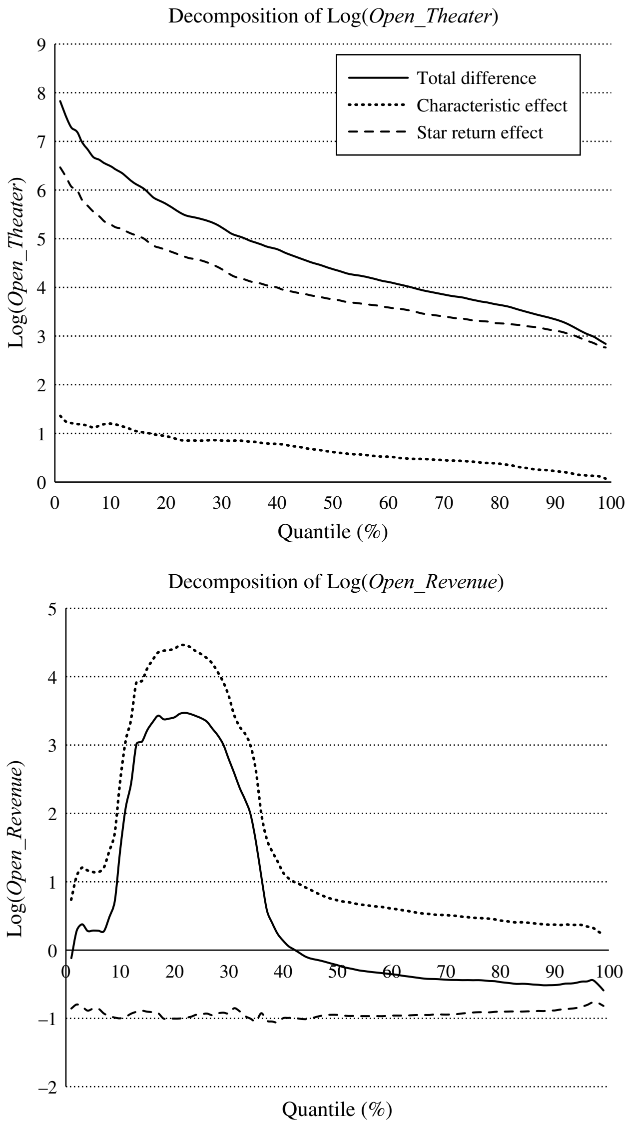
5.3. Counterfactual Decomposition

As discussed earlier, the counterfactual decomposition requires applying the returns on the movie characteristics of one group (i.e., the star or nonstar) to the distributions of the movie characteristics endowed to the *other* group. By replacing the characteristics of nonstar movies with those of star movies (and vice versa), we obtain the quantiles of the distribution that we would observe if the nonstar movies had the characteristics of star movies (i.e., the counterfactual distribution). Because of the need for this substitution and to be consistent with the counterfactual decomposition literature, we carry out this analysis with the results for the split-sample results with star selection corrections. The decompositions of the difference in opening week theaters along the 10th, 20th, …, 90th percentiles are presented in the left half of Table 6. Figure 2 (upper graph) shows the decompositions visually.

The estimated star and nonstar theater allocation gap is positive and significant across the entire distribution, which implies that the number of opening week theaters allocated for star movies is higher than that for the nonstar movies at all the corresponding quantiles. Second, the star versus nonstar gap in opening week theater allocations is wider for movies in the bottom quantiles, and the difference monotonically declines as the movies move up the open theater quantiles. Third, while both the star return and characteristics effects are statistically significant in all the quantiles, the gap between star and nonstar is mostly explained by the star return effect up across the entire *Open_Theater* distribution. The contribution of movie characteristics to the total differences is only about 18% at the 10th quantile, declining to about 7% at the 90th quantile. The remaining difference is explained by the star returns.

The counterfactual decompositions of Log(*Open_Revenue*) at each percentile are reported in the right half of Table 6 and are plotted in the lower graph of Figure 2. Overall, the total difference in the star versus nonstar effects on opening week revenues is positive and significant between the 20th and 30th percentiles, suggesting that star movies only outperform nonstar movies in the middle to low quantiles. The difference is mostly due to the differences in the characteristics of the star and nonstar movies. The star return effect is not significant at any of the quantile levels, suggesting that stars do not impact Log(*Open_Revenue*) directly. It is, however, important to note that star status does matter in the opening week theater allocation, which is a significant covariate in the QR analysis of the *Open_Revenue* at all levels of its distribution (Table 6). Thus, the effect of stars on audience is channeled

Figure 2 Counterfactual Decomposition of Star Effects on Theaters and Revenues



through screen allocations. We now summarize our findings and discuss their implications.

6. Summary and Discussion

The large financial rewards that movie stars receive and the influences they exert on the participants in the movie industry have led industry analysts, as well as academic researchers, to raise the question of what stars really produce in return. Answering this question is not straightforward, primarily because stars may be preferentially selected to act in movies, and the movies they accept may possess the ingredients of a successful future. We address the selection problem by adopting an endogenous switching regression model, in which we model the star selection process that produces a correction for the selection bias in the assessment of star effects on the opening week theater allocations and revenue. To assess the robustness of the results, we also estimate a pooled (star and nonstar) model in which the star is deemed endogenous, and the endogeneity is corrected with control functions. To disentangle the star effects from the effects of the movie characteristics with which stars may be differentially endowed, we frame the question as "What would the movie outcome be if a nonstar had instead acted in a movie with the same characteristics as those in which a star had appeared?" The counterfactual method decomposes the gap between the outcomes of star and nonstar movies into two parts: (1) the characteristic effect, which arises from the differences in the movie characteristics of the star and nonstar movies; and (2) the star return effect, whicharises from the star's ability to obtain greater returns on the same movie characteristics than a nonstar can. To investigate how the characteristics and star return effects might differ across the entire range of the outcome distributions, we carry out the decomposition in a quantile regression framework. We discuss the main findings and their implications.

6.1. Star Selection

A star's selection to appear in a movie involves a bilateral decision of the studio and the star (or the star's agent). Studios may approach stars for their suitability to the roles or for their bankability or to mitigate the risks of unfavorable movie outcomes. Stars, on their part, utilize the information available at the time of signing a movie contract to determine whether the movie has the potential for success that would retain or strengthen their star status. Although we do not fully model the equilibrium outcome of the studio–star interaction, we take the star's perspective and model the star's decision to accept or reject a role. The star selection is a critical component in the investigation into star effects on movie outcome, which has been largely ignored in past academic research.

It is not surprising that stars tend to accept movies associated with directors and producers who have been recognized by the MPAA or the Directors/Producers Guilds for their talent. Selection of movies to be produced and directed by individuals recognized by the industry obviously enhances the star's artistic reputation in addition to increasing the potential for the movie's financial success. Some of the well-known star–director pairings include Steven Spielberg and Tom Hanks and Martin Scorsese and Robert De Niro (now Leonardo DiCaprio). The star's propensity to accept movies produced by the big six production companies is likely predicated on the assurance of resources required for completion and marketing of the movie. With regard to the MPAA ratings, although such ratings are not granted at the time of signing of a star, we speculate that the scripts and the story lines would permit a star to predict fairly accurately whether the movie will be rated R. The stars' avoidance of potentially rated R movies is justified by the negative effects an R rating has on the number of theaters in which these movies open and opening week revenues. Stars also avoid foreign movies, presumably due to higher (than nonstars) opportunity costs and the difficulty of managing multiple shooting schedules across countries.

What we find interesting is that stars, as a class of actors and actresses, flock to movies with certain genres and avoid others. Comedy and adventure, the two more frequently MPAA rated genres where the participation of stars is greater than that of nonstars, have negative effects (relative to the action genre) on the opening week revenue for the star sample. Thus, the stars' affinity for these genres is not motivated by revenue expectation. The preference for comedy or adventure might be either due to the stars' beliefs about their abilities to better apply their talents to these genres or an exhibition of a herding behavior.

6.2. Is It the Star or the Movie Characteristics?

After controlling for a star's nonrandom selection for and of movie projects, we see that, relative to nonstar movies, star movies have a positive influence on the number of theaters in which they open. On the other hand, star movies have practically no incremental effect on the opening week revenues, except at the lowest quantile. The differences in theaters and revenues, however, cannot be entirely attributed to the stars if the movie characteristics of star and nonstar movies that also influence the movie outcomes are systematically different. In the separation of the star and movie characteristics effects achieved by using the counterfactual decomposition, the respective effects differ, depending on whether the outcome under investigation is theater allocations or the revenues.

In deciding the number of theaters to allocate in the opening week for the movie, the distributors rely on both the movie characteristics and the presence of a star. However, the presence of a star contributes to 80%–90% of the difference in theater allocations, and the rest is attributable to the movie characteristics. This finding is consistent with previous findings that distributors allocate more theaters to the star movies (De Vany and Walls 1999). It also lends credence to the signaling argument. Because each movie is different, distributors make decisions about opening week theater allocation with very limited information about the movie's quality and its demand, which can only be known after the opening week. Also, because distributors often work with internal finances, risk-averse distributors (as well as exhibitors) are left to rely primarily on the star cast of a movie to signal viewers about movie quality. It should be noted, however, that we only consider the theater allocated in the opening week. The exhibitors do have the opportunity to quickly adjust the theater allocations in later weeks after observing the opening week performance of the movie (Elberse and Eliashberg 2003).

With regard to the opening week revenue gaps, the results are reversed. It is the differences in the characteristics of the star and nonstar movies that explain the difference in the opening week revenue. The presence of a star does not contribute to the difference at any level. Thus, if the same movie characteristics were present for the "disadvantaged group" (i.e., nonstar movies), the revenue outcome would have been the same at all levels of the distribution (i.e., the star return effect is nonsignificant at all quantiles). This result is consistent with the findings of Ravid (1999) and De Vany and Walls (1999). Overall, relative to the distributors, the audience is less enamored of the star status of a movie and is more interested in the movie itself as well as critics' comments about the movie. The following excerpt from *MovieScope* (2011) aptly captures this conclusion:

> The basic concept of the star-driven box-office hit is clearly coming progressively under threat. The reasons are various and many, though the wane at the top end can be largely attributed to the huge mega-budget spectacle of the studio flagship special effects films that appeal to global audiences and now dominate the worldwide box office. Movies such as *Avatar, Transformers, 2012, Star Trek* et al. don't have pricey stars, because the films themselves are the pricey stars. This is a trend that looks set to stay with us.

6.3. Star's Effects on Revenue

Although the presence of a star does not explain the difference in revenues between star and nonstar movies, the star's influence on movie revenues is indirect. Combining the results from the QR and counterfactual analysis reveals that a star directly and positively affects the number of theaters allocated in the opening week at all levels of open theater distribution. For instance, at the median of the open theater distribution, about 86% of the gap in theaters allocated between the star and nonstar movies (i.e., 2,711 versus 1,621 theaters), comes from having a star in the movie. The theater allocation then influences opening week revenue and this effect remains significant at all levels of revenue distributions. For instance, the elasticity of revenue with respect to number of theaters remains at about 1.00 throughout the revenue distribution, which interestingly is about the same return that nonstar movies generate. Thus, the star's effect on revenues comes via increased theater allocations, not from the star's ability to utilize this benefit to an advantage in generating revenue. The movie revenue gap between star and nonstar movies may also be explained by the characteristics of movies in which stars appear. To the extent that stars are able to select the "right" set of movie characteristics, stars can be credited for the characteristics effects on movie revenues.

6.4. Limitations and Conclusion

Although we estimated a star selection model to account for star endogeneity, a more complete analysis would require an equilibrium modeling of the studio's decision to approach a star and the star's decision to accept or reject the offer. For instance, a studio might approach a star with high bankability to secure resources, or one whose strength in a given genre matches the genre of the focal movie. Second, the complexity associated with estimating the star effects while correcting for the star selection and star–movie characteristics intermingling required us to limit our analyses to only the opening week theater counts and opening week revenues. Although the opening week performance of a movie permits the assessment of star effects in the absence of other effects such as word of mouth of past viewers, a more complete analysis of star effect should consider a movie's performance during its entire theatrical run, which would account for the dynamic effects of theater allocations as well as word of mouth. In addition, stars may have different effects on international markets and DVD releases than on the U.S. opening week performance. Stars may also generate mediating effects such as free publicity generated by media coverage that in turn may influence movie revenues. Some of these issues have been investigated individually, but future research that integrates the different star effects is needed.

We also did not distinguish between movies with a single star and multiple stars. As previously noted, our sample had very few movies containing multiple stars. Besides, the restriction of the counterfactual method to have the same covariates in both star and nonstar estimations did not permit us to carry out this analysis. Aside from the methodological constraints, our substantive understanding of the studio–star negotiations, even with one star, is inadequate. With more than one star, possibly with different past achievements and genre strengths as well as the complementary or substitutive skills that multiple stars may bring, such an investigation is likely to be complex but would be of interest to studios.

In conclusion, the counterfactual methodology combining with QR with the star-selection correction utilized in this study is, to the best of our knowledge, the first to be used in the marketing literature. Yet, this approach offers great promise to evaluate outcomes or performances of two groups (or time periods) and separate the effects caused by endowments directed to the groups from their true abilities to derive better returns on these endowments. For instance, does a group of stores following one pricing strategy derive higher sales than a group following another strategy because of different assortments, store size, etc.? Or is one strategy better at deriving higher sales from the same characteristics than the other strategy? Do the star territory managers derive better results due to the territory characteristics and resources allocated to their territories, or due to their ability to derive better results from the same resource? Given the availability of data to generate empirical distributions for both groups, the counterfactual approach can answer such questions, which are of theoretical as well as managerial significance. Hopefully, this study will stimulate more applications of this empirical methodology to answer important marketing questions.

Acknowledgments

The authors thank the department editor, the associate editor, and the anonymous reviewers for their constructive comments. The authors also acknowledge the helpful comments provided by Amber Anand, Natarajan Balasubramanium, Yubo Chen, Scott Fay, Eunkyu Lee, Yong Liu, Ishani Majumder, Madhura Sengupta, and Joyce Zadzilka. Zailong Wan provided valuable programming assistance. Financial support from the Beijing Higher Education Young Elite Teacher Project and the National Natural Science Foundation of China [Grants 71272029, 71328202, 71102007, and YETP0135] and from Snyder Innovation Management Center is gratefully acknowledged.

References

Ackman D (2002) The myth of stars. *Forbes* (June 12), http://www.forbes.com/2002/06/12/0612stars.html.

Ackman D (2003) Hollywood's star power failure. *Forbes* (June 19), http://www.forbes.com/2003/06/18/cx_da_0619stars.html.

Adler M (1985) Stardom and talent. *Amer. Econom. Rev.* 75(1):208–212.

Albert S (1998) Movie stars and the distribution of financially successful films in the motion picture industry. *J. Cultural Econom.* 22(4):249–270.

Albrecht J, Bjorklund A, Vroman S (2003) Is there a glass ceiling in Sweden? *J. Labor Econom.* 21(1):145–177.

Arulampalam W, Booth A, Bryan ML (2007) Is there a glass ceiling over Europe? Exploring the gender pay gap across the wage distribution. *Indust. Labor Relations Rev.* 60(2):163–186.

Basuroy S, Chatterjee S, Ravid SA (2003) How critical are critical reviews? The box office effects of film critics, star power and budgets. *J. Marketing* 67(4):103–117.

Basuroy S, Desai KK, Talukdar D (2006) An empirical investigation of signaling in the motion picture industry. *J. Marketing Res.* 43(2):287–295.

Blinder AS (1973) Wage discrimination: Reduced form and structural estimates. *J. Human Resources* 8(4):436–455.

Bollinger C, Ziliak JP, Troske KR (2011) Down from the mountain: Skill upgrading and wages in Appalachia. *J. Labor Econom.* 29(4):819–857.

Buchinsky M (1998) The dynamics of changes in the female wage distribution in the USA: A quantile regression approach. *J. Appl. Econometrics* 13(1):1–30.

Chen Y, Ganesan S, Liu Y (2009) Does a firm's product-recall strategy affect its financial value? An examination of strategic alternatives during product-harm crises. *J. Marketing* 73(6):214–226.

Chernozhukov V, Hansen C (2005) An IV model of quantile treatment effects. *Econometrica* 73(1):245–261.

Chernozhukov V, Hansen C (2006) Instrumental quantile regression inference for structural and treatment effect models. *J. Econometrics* 132(2):491–525.

Chintagunta PK, Gopinath S, Venkataraman S (2010) The effects of online user reviews on movie box office performance: Accounting for sequential rollout and aggregation across local markets. *Marketing Sci.* 29(5):944–957.

De Vany AS, Walls WD (1999) Uncertainty in the movie industry: Does star power reduce the terror of the box office? *J. Cultural Econom.* 23(4):285–318.

De Vany AS, Walls WD (2004) Motion picture profit, the stable paretian hypothesis, and the curse of the superstar. *J. Econom. Dynam. Control* 28(6):1035–1057.

Einav L (2007) Seasonality in the U.S. motion picture industry. *RAND J. Econom.* 38(1):127–145.

Elberse A (2007) The power of stars: Do star actors drive the success of movies? *J. Marketing* 71(4):102–120.

Elberse A, Eliashberg J (2003) Demand and supply dynamics for sequentially released products in international markets: The case of motion pictures. *Marketing Sci.* 22(3):329–354.

Eliashberg J, Elberse A, Leenders MAAM (2006) The motion picture industry: Critical issues in practice, current research, and new research directions. *Marketing Sci.* 25(6):638–661.

Gyourko J, Tracy J (1988) An analysis of public-and private-sector wages allowing for endogenous choices of both government and union status. *J. Labor Econom.* 6(2):229–253.

Heckman JJ (1979) Sample selection bias as a specification error. *Econometrica* 47(1):153–161.

Heitmueller A (2006) Public-private sector pay differentials in a devolved Scotland. *J. Appl. Econom.* 9(2):295–323.

Hennig-Thurau T, Houston M, Sridhar S (2006) Can good marketing carry a bad product? Evidence from the motion picture industry. *Marketing Lett.* 17(3):205–219.

Hennig-Thurau T, Houston MB, Heitjans T (2009) Conceptualizing and measuring the monetary value of brand extensions: The case of motion pictures. *J. Marketing* 73(6):167–183.

Joshi AM, Hanssens DM (2009) Movie advertising and the stock market valuation of studios: A case of great expectations? *Marketing Sci.* 28(2):239–250.

Kamakura W, Basuroy S, Boatwright P (2006) Is silence golden? An inquiry into the meaning of silence in professional product evaluations. *Quant. Marketing Econom.* 4(2):119–141.

Karniouchina EV (2011) Impact of star and movie buzz on motion picture distribution and box office revenue. *Internat. J. Res. Marketing* 28(1):62–74.

Klein RW, Spady RH (1993) An efficient semiparametric estimator for binary response models. *Econometrica* 61(2):387–421.

Koenker RW, Bassett G Jr (1978) Regression quantiles. *Econometrica* 46(1):33–50.

Lee L-F (1978) Unionism and wage rates: A simultaneous equations model with qualitative and limited dependent variables. *Internat. Econom. Rev.* 19(2):415–433.

Lee S (2007) Endogeneity in quantile regression models: A control function approach. *J. Econometrics* 141(2):1131–1158.

Lehmann DR, Weinberg CB (2000) Sales through sequential distribution channels: An application to movies and videos. *J. Marketing* 64(3):18–33.

Liu Y (2006) Word of mouth for movies: Its dynamics and impact on box office revenue. *J. Marketing* 70(3):74–89.

Luo L, Chen X, Han J, Park CW (2010) Dilution and enhancement of celebrity brands through sequential movie releases. *J. Marketing Res.* 47(6):1114–1128.

MacDonald GM (1988) The economics of rising stars. *Amer. Econom. Rev.* 78(1):155–166.

Machado JAF, Mata J (2005) Counterfactual decomposition of changes in wage distributions using quantile regression. *J. Appl. Econometrics* 20(4):445–465.

Melly B (2005) Decomposition of differences in distribution using quantile regression. *Labour Econom.* 12(4):577–590.

MovieScope (2011) Acting up—Why distributors can no longer rely on old-school star power to sell a movie. (September 7), http://www.moviescopemag.com/24-fps/distribution/acting-up-why-distributors-can-no-longer-rely-on-old-school-star-power-to-sell-a-movie/.

Muth JF (1961) Rational expectations and the theory of price movements. *Econometrica* 29(3):315–335.

Neelamegham R, Chintagunta P (1999) A Bayesian model to forecast new product performance in domestic and international markets. *Marketing Sci.* 18(2):115–136.

Oaxaca R (1973) Male–female wage differentials in urban labor markets. *Internat. Econom. Rev.* 14(3):693–709.

Ravid SA (1999) Information, blockbusters, and stars: A study of the film industry. *J. Bus.* 72(4):463–492.

Rosen S (1981) The economics of superstars. *Amer. Econom. Rev.* 71(5):845–858.

Shugan SM (2004) Endogeneity in marketing decision models. *Marketing Sci.* 23(1):1–3.

Sochay S (1994) Predicting the performance of motion pictures. *J. Media Econom.* 7(4):1–20.

Sorenson O, Waguespack DM (2006) Social structure and exchange: Self-confirming dynamics in hollywood. *Admin. Sci. Quart.* 51(4):560–589.

Vogel HL (2007) *Entertainment Industry Economics: A Guide for Financial Analysis* (Cambridge University Press, New York).

Wallace WT, Seigerman A, Holbrook MB (1993) The role of actors and actresses in the success of films: How much is a movie star worth? *J. Cultural Econom.* 17(1):1–27.

Walls WD (2005) Modelling heavy tails and skewness in film returns. *Appl. Financial Econom.* 15(17):1181–1188.

Walls WD (2009) Robust analysis of movie earnings. *J. Media Econom.* 22(1):20–35.