SOCIOLOGICAL INQUIRY

Million Dollar Maybe? The Effect of Female Presence in Movies on Box Office Returns*

Andrew M. Lindner, *Skidmore College*Melissa Lindquist and Julie Arnold, *Concordia College*

This study examines processes that contribute to the underrepresentation of women in film by linking the depiction of gender in film to its impact on domestic box office returns. Drawing on a sample of widely distributed movies from 2000 to 2009 (n = 974), we test whether the box office under-performance of films with an independent female presence results primarily from "downstream" public rejection or from an gendered "upstream" division of resources that provides greater studio support to movies about men. Using a series of multivariate regression analyses and controlling for genre, critical appraisal, arthouse label, being a sequel, and including a popular star, films with a female presence earn less at the box office. This effect, however, appears to be largely the consequence of movies that feature women having smaller production budgets, suggesting that the underrepresentation of women in film stems from "upstream" routines of film industry gatekeepers, not a lesser interest in stories about women in the minds of the public.

In Hollywood, *Charlie's Angels* and its sequel, *Charlie's Angels: Full Throttle*, are relative anomalies. Although the crime-foiling Angels are highly sexualized and, in many ways, gender–stereotypical characters, their very presence defies film industry norms. That the *Charlie's Angels* films feature multiple women having interactions that do not revolve exclusively around men is unusual. That both movies, which prominently feature multiple women, had big production budgets and proved successful at the box office is even more remarkable.

While films that portray women in prominent roles have become more common since mid-century, women continue to be underrepresented on the big screen. Even when multiple women appear in a film, they are often ancillary characters in stories about men. Yet, despite a crowded literature on gender inequality in film, we know little about how women's exclusion is situated within the wider entertainment industry. Content analysis alone misses upstream effects, in which content is structured through the actions of media gatekeepers (e.g., producers, directors, and executives in movie studios) in gendered organizations who prejudge what the public wants, as well as downstream effects where audiences respond to content and attitudes are formed and reinforced (Acker 1990; Corts 2001; Griswold 2013).

Sociological Inquiry, Vol. 85, No. 3, August 2015, 407–428 © 2015 Alpha Kappa Delta: The International Sociology Honor Society

DOI: 10.1111/soin.12081

In the current study, examining the case of gender representation in film, we ask: Do movies that feature women who lead lives independent of men underperform at the box office? And, if so, does lower revenue reflect consumer demand for male-centric stories or is it part of a gendered division of resources within the industry decision-making process? Drawing on a sample of the most widely distributed films annually from 2000 through 2009, we combine a content analysis using the Bechdel Test (a test of gender representation) with revenue data to examine whether the inclusion of an independent female presence in a film affects box office performance. Then, by controlling for other film characteristics such as the presence of stars, its critical reception, its genre, and its production budget, we go further to consider whether the box office underperformance of films with an independent female presence results primarily from public rejection or from a gendered division of resources that provides greater studio support to movies about men. While other current studies that assess the extent of women's representation in media are certainly important, our research uses women's representation in film as an independent, not a dependent, variable. We examine whether and why movies with an independent female presence suffer a box office penalty. In doing so, we contribute to sociological theory about both the processes through which gender inequality is produced and the relationships among cultural creators, objects, and recipients.

We begin by reviewing background literature regarding gender content analyses in film, box office performance, and the dynamics of the film industry. We continue by discussing the methods used to select and analyze the 974 films in the sample. Then, we review the results of a cross-tabulation and several multivariate regression analyses. We conclude with a discussion of the implications of these findings for our understanding of how gender inequality is structured within the film industry and consider the ways in which this study might be extended in the future.

Background

Content Analysis

Since the early 1970s, numerous studies have used content analysis to examine depictions of gender in television and movies. The expansion of electronic access to media and transcripts via the Internet has eased the process of conducting such studies and lead to exponential growth in content analyses (Lindner 2012; Rudy, Popova, and Linz 2010). The journal *Sex Roles* alone published 114 gender content analyses between 1978 and 2009 (Rudy, Popova, and Linz 2010). Indeed, such studies are so common that one scholar acknowledged "journals seem glutted with gender studies of contents" (Jolliffe 1993:93). Traditionally, these content analyses document stark inequalities in

media depictions of men and women. Even though a few recent studies have shown that women are featured in an increasing number of leading roles on television and in movies, men still appear more frequently than women and gender-stereotypical depictions remain dominant in popular programming.

Although gender-stereotypical depictions abound, the more fundamental fact is that women are remarkably underrepresented in mass media. This finding holds true in fictional television (Eschholz, Bufkin, and Long 2002; Fejes 1992; Fischer 2010), radio (Gill 1993), news (Gallagher 2005), and advertising (Mager and Helgeson 2011). In a well-known study of primetime television between 1969 and 1978, Gerbner and Signorielli (1979) found that men outnumbered women by a ratio of 2.5 to 1. In their 2010 study of 1,291 health storylines on primetime television shows, Hether and Murphy (2010) found that men dominated every type of role: 57 percent of the patients, 62 percent of the bystanders, 80 percent of the characters that caused the illness, and 60 percent of the caregivers. As Collins notes, since the 1970s, although "women have progressed to nearly representative rates of participation in the working world, the 'reel' and print worlds have continued to overlook them" (2011:292).

The pattern of women's underrepresentation in media extends to film. Lauzen and Dozier's (2005) study of the 100 top-grossing films of 2002 revealed that 73 percent of all major characters in the movies were male. Smith and her coauthors found nearly identical results in G-rated films from 1990 to 2005: 72 percent of the characters were male, and of the films with narrators, 83 percent were male (Smith et al. 2010).

This pattern of underrepresentation has garnered attention beyond the halls of academia. The "Bechdel Test," named after graphic artist Alison Bechdel, originated from her comic strip Dykes to Watch Out For and first appeared in a 1985 strip titled The Rule. To pass the Bechdel Test, a film must meet the following criteria: (1) it must have at least two women with names, (2) those named women must talk to each other, and (3) in those conversations, the women must talk about something other than a man (Bechdel Test Movie List 2010). The Bechdel Test has inspired videos, essays, and, in Sweden, even a film rating system (Associated Press 2013). A few popular studies have found that only about half of all films pass the test (Hickey 2014). Taken together, all existing evidence strongly suggests that women are underrepresented in film, and when present, are often secondary characters in male-centric stories.

Box Office Performance

An obvious explanation for the underrepresentation of women in a profitseeking film industry is that movies with an independent female presence earn less at the box office. Such a pattern would make female-driven films less attractive to movie studios. In this study, we test this simple empirical question

by connecting films' gender content with their box office performance. Many researchers and movie industry analysts have attempted to better predict movie box office returns. Revenue forecasting models are invaluable for studio executives and producers as they make decisions about which movies to "greenlight" and how best to promote their movies. Existing research has focused on producing an optimally predictive model and has identified several highly predictive variables. The number of screens showing the film in the first week of its run, the size of the movie's budget, the presence of stars, and its genre all significantly affect box office performance (Elberse 2006; Neelamegham and Chintagunta 1999). Other research has considered the role of Internet "word-of-mouth" (Liu 2006), seasonality (Einav 2007), critics' reviews (Eliashberg and Shugan 1997), and even the characteristics of character development and plot (Eliashberg, Hui, and Zhang 2007). Asur and Huberman (2010) have produced models explaining more than 95 percent of the variation in box office performance based on buzz on the social network Twitter.

In a particularly important study for developing a theory of movie profitability, Hennig-Thurau, Houston, and Walsh (2007) used path analysis on a sample of 331 films to better understand the inter-relationships among a variety of factors that are predictive of box office performance. These factors include both upstream effects, such as production budget, existing "cultural familiarity" (e.g., being a sequel or based on a popular novel), "starpower," and release dates, and downstream effects, including critical reviews and consumer-perceived quality. Their findings indicate that short-term box office revenues are most heavily affected by the film's production costs. In short, the best predictors of box office performance were upstream effects—decisions made by key film industry gatekeepers about which projects to greenlight and how big of a production budget to grant—rather than downstream effects like consumer perception.

Although no existing academic research has documented how gendered content affects box office performance, two analyses by online news and entertainment Web sites offer some insight. Two writers for the site Vocativ conducted an analysis of "50 of the biggest box office movies of 2013" (Sharma and Sender 2014). They found that only 36 percent of the films passed the Bechdel Test, but that the movies that did earned more at the box office than those that did not pass the test (a total of \$4.22 billion versus \$2.66 billion). Their conclusion was that "Hollywood movies with strong female roles make money" (Sharma and Sender 2014).

Their primary finding may be unexpected, but Sharma and Sender's (2014) analysis is riddled with significant methodological problems. Even holding aside their incredibly small sample, by selecting only "the biggest box office movies," they essentially sample on the dependent variable, reducing the

range of variation. Moreover, their substantive conclusion that "strong female roles make money" is a misrepresentation of what the Bechdel Test can demonstrate (i.e., it offers no way of testing whether there are "strong female roles").

Hickey (2014), writing for data journalism Web site FiveThirtyEight, produced a more serious effort. Using data from BechdelTest.com and The-Numbers.com, he was able to construct a sample of 1,615 films released from 1990 to 2013. Although the size of this sample is a big improvement over the Vocativ analysis, the selection and coding of films in Hickey's sample was non-systematic. The sample included only movies that visitors to the Bechdel-Test.com decided to upload, leading to potential underrepresentation of less widely distributed films. Additionally, neither the BechdelTest.com nor Hickey has a way of ensuring the reliability of the user-submitted codings of whether a given film passed the Bechdel Test. Unlike Sharma and Sender's (2014) findings, Hickey found that a slim majority (53%) of films passed the Bechdel Test. In terms of the box office performance of Bechdel films, once again contrary to Sharma and Sender's conclusions, Hickey reported that passing the Bechdel Test had no effect on box office performance at either the bivariate level (comparing the median box office revenue) or in a regression analysis controlling for budget. However, he did note that the median production budget for movies that passed the Bechdel Test was significantly lower than for those that did not.

The findings of these two analyses offer us very different understandings of gender representation and the box office. In one account, a majority of films pass the Bechdel Test and there is no box office penalty for representing women (Hickey 2014). The other argues that women are underrepresented, but that the movies that do feature women do better at the box office (Sharma and Sender 2014). Neither of these analyses support the view widely held in Hollywood that films featuring women earn less at the box office. This study aims to resolve the apparent dispute using a large and systematic sample, a reliable coding method, and a regression analysis with more control measures. In doing so, we can establish the extent of gender inequality in film and better understand the processes that perpetuate it.

Mechanisms of Gender Inequality in Media

The "cultural diamond," first developed by Griswold (2013), is a "model of culture" that envisions a set of links among (1) "cultural objects" (e.g., movies), (2) creators (e.g., movie producers, directors, screenwriters, actors), (3) recipients (e.g., moviegoers), and (4) the wider social world. The cultural diamond model does not attribute causal directions, but posits that both creators and recipients play a role in constructing the meaning of a cultural object. At the same time, creators and recipients communicate directly, with creators hyping new cultural objects and recipients providing feedback. Alongside cultural intermediaries, like movie critics who provide formal feedback in the form of reviews, recipients also provide feedback through sales of cultural objects (e.g., paying for movie tickets) and related merchandise (e.g., movie-related toys). In turn, "Producing organizations interpret...feedback to assess the popularity of an artist, the effectiveness of promotional activities, and implications for future productions" (Griswold, 2013:76).

If we presume, momentarily, that films with women who lead lives independent of men¹ do, in fact, earn less at the box office, there are two possible explanations—one driven by the movie-going audiences (a downstream effect) and the other shaped by the film industry creators (an upstream effect). Upstream and downstream effects offer two different accounts of where the production of gender inequality is located within the cultural diamond. According to the upstream account, any box office underperformance of Bechdel films would be due to constitutive decisions that take place in the link between the creators (movie industry gatekeepers) and the cultural object (the movie). A downstream effect would be produced in the relationship between the cultural object (the movie) and the recipients (the potential audience). We consider each of these two alternatives in turn.

As Risman (2009) writes, "The gender structure [of every society] has implications at the level of individual analysis...and at the institutional level in the organization and policing of social groups" (p. 83). One possibility is that lesser box office earnings of films with an independent female presence is a downstream effect, operating at the "level of individual analysis." In a larger social system of stratification that "ranks men above women of the same race and class" (Lorber 1994:32), audiences may prefer films that prominently feature men. Or using the cultural diamond model, "creators produce an excess supply of all cultural objects" (Griswold 2013: 81) and, posed with choice, recipients—socialized to value men over women—tend to select male-centric cultural objects. If the downstream explanation were correct, we would expect to find that films with an independent female presence earn less at the box office² (i.e., are less popular with the film-going public), net of other factors such as the film's genre, budget, and star power.

In the second possibility, films with an independent female presence may perform worse at the box office due to upstream decision-making by industry gatekeepers. That is to say, that gender inequalities operate at the "institutional level" (Risman 2009) or the link between creators and cultural objects within the cultural diamond model (Griswold 2013). Producers and studio executives, trying to keep their budgets in the black, commonly believe that there is no clear formula for success or that "all hits are flukes" (Bielby and Bielby 1994).

Given the pressing need to be profitable and the perceived unpredictability of film performance, cultural producers within the entertainment industry attempt to minimize risk (Peterson and Anand 2004). As Lampel and Shamsie (2003) write, describing the routines of Hollywood executives,

The search for effective routines and practices in the preproduction phase is motivated by the risks involved. Highly paid and experienced managers are limited in the amount of time they can devote to projects. Projects that go through the development and casting process and yet fail to go into production represent a significant loss in terms of opportunity costs to the managers and firms involved ... Poorly conceived projects with resource bundles that are ill suited for production or the market place increase the risk of box office failure... (p. 2195).

The ironic upshot of the routinization of gatekeeping practices is that, despite believing there is no formula to success, studio executives attempt to "predict the unpredictable" by reproducing what they see as successful formulas (Gitlin 1983:32). In Bielby and Bielby's classic study of decisions made by television programming directors in selecting new primetime TV series, they found that network programmers act in an extremely conservative fashion, selecting only from reputable creators, concepts, and genres: "As the level of risk... increases... reliance on established writer-producers is likely to intensify" (1994:1309). Likewise, in the film industry, each additional past credit increases a filmmaker's subsequent chances of being hired for other projects (Faulkner and Anderson 1987). In this way, executives reproduce past models of success in the hopes of avoiding a box office bomb.

Along these lines, it may be that the types of movies that tend to have an independent female presence are not viewed as favorably by studio executives and, therefore, are less likely to be supported with the kind of resources necessary to be profitable as Hickey (2014) found in his analysis of Bechdel films. Anecdotal evidence for studio executives' doubt in women's earning potential is exemplified by a 2007 report by the reputable film industry blog Deadline Hollywood. The blog reported that Warner Brothers' President of Production Jeff Robinov had bluntly informed his staff, "We are no longer doing movies with women in the lead." Warner Brothers, according to the story, immediately issued a denial (Finke 2007). Still, even if studios are not directly discriminatory, executives may provide larger budgets to male-centric genres, such as fantasy and adventure films.

These kinds of gendered assumptions and decisions are typical across a range of organizations and labor markets. Acker's (1990) theory of "gender organizations" argues that gender inequality is built into the very nature of work organizations. One process through which gendered organizations produce inequality is a division of labor. Lorber (1994) describes this division of labor as gendered "segmentation" in which industries divide occupations into

"sectors with different...credential requirements for hiring, different promotion ladders, different work assignments, and different pay scales" (p. 198). In the film industry, there are significant forms of segmentation. For example, big budget action flicks (which tend to come with higher pay for actors) have more roles for men, while smaller budget movies like romantic comedies and indie films have more parts for women.

At an even more fundamental level, Acker (1990) theorizes that work organizations have gendered "organizational logics" that privilege the type of work tasks associated with men and masculinity. If women act in gender stereotypical ways, they will fail to perform the work tasks "correctly." On the other hand, if women adopt masculine characteristics to "properly" perform the task, they face stigma for violating gender norms. For example, Williams, Muller, and Kilanski (2012) studied geoscientists and found that "to excel at teamwork, individuals must be able to engage in self-promotion, which can be difficult for women in male-dominated environments..." (p. 570). Within Hollywood, there might be any number of organizational logics occurring upstream that privilege men over women. Just as Williams, Muller, and Kilanski (2012) found that gendered organizational logics created barriers for women in successfully networking within the oil and gas industry, it may be that women in Hollywood are disadvantaged in networking, budget negotiations, and/or casting decisions. These interactions are all crucial as studio executives decide what resources to allocate to a given film project.

In this upstream explanation, box office underperformance results not from the audience's rejection of films with women, but from dynamics of "gendered organizations" (including segmentation and organizational logics) which lead studios to allocate fewer resources to films featuring women. That is, the labor of female performers is relegated to projects without as much potential to generate as great box office returns. In the cultural diamond model, the origin of the effect would be the link between creator and cultural object. If this is the case, we expect that any box office penalty for passing the Bechdel Test should vanish when controlling for factors like genre and the size of production budget.

Our study offers an opportunity to understand whether downstream factors (at the individual level of consumer demand) or upstream factors (an unequal division of industry resources at the institutional level) better explain the financial underperformance of films with an independent female presence at the box office, locating the source of gender inequality within the cultural diamond model. Strong arguments can be made for both explanations. However, we suspect the ranking of men over women is so deeply socialized at the individual level that lower box office returns for Bechdel movies (if observed) are likely to be the product of consumer demand. Thus, we hypothesize a downstream effect: All things being equal, movies with an independent female presence (as

measured by the Bechdel Test) will have lower domestic box office totals than movies without one.

Methods

Sample and Analytical Strategy

This study links the gendered content of films with their domestic box office performance. According to the Internet Movie Database (IMDB; www.imdb.com), over 5,000 feature films were released worldwide each year in the first decade of the twenty-first century. Many of these films had no theatrical opening and were released "straight-to-DVD." Others were foreign films that were never released in the U.S. market. Because of our interest in audience response to movies, we limited our sample to the films that were most readily available to the American viewing public. Using IMDB, we selected as our sample the 100 movies with the widest distribution (as measured by total number of screens) in each year between 2000 and 2009 (for a total of 1,000 films).

Although distribution is correlated with the dependent measure of gross domestic box office revenue, the hundred most widely distributed films annually incorporates far more than just the big box office "winners" (including arthouse movies like 2007's Waitress and big budget flops like 2004's The Alamo). Moreover, according to the Arbitron Cinema Advertising Study (2007), the average American goes to the movies fewer than eight times a year. Even among the key movie-going demographic, ages 12-24, the average person visits the movie theater only 16 times per year. By selecting the 100 most widely distributed films for each year, we are likely to capture most of the movies the average American would have seen in the cinema as well as the majority of the most highly visible filmic representations of women in that year.

After excluding documentaries, the sampling frame included 997 films. Of these, we were able to secure full information for 974 movies. The unit of analysis for this study is at the level of the movie. Our analytic strategy is to examine the relationship between the Bechdel Test and box office performance on the bivariate level by performing cross-tabulations followed by multiple iterations of multivariate linear regression analysis to isolate the independent effect of each of the independent variables on the dependent variable.

Dependent Variable

The dependent variable for this study was gross domestic box office revenue (as reported by the Internet Movie Database). Film studios also pay close attention to revenue from international box office returns and DVD rentals and sales, but due to inconsistencies in reporting across international markets and the delay in revenue in DVDs sales, these metrics are less reliable and less widely reported for films outside the top-grossing blockbusters of the year. For this reason, gross domestic box office revenue offered the most consistent and readily available measure of economic performance. Gross domestic box office revenue was measured in millions; earnings ranged from \$17.2 million to \$754 million. Complete descriptive statistics are in Table 1.

Independent Variables

This study uses the Bechdel Test (whether a movie features two or more named women who speak to each about something other than a man) as its central independent measure of female representation. The Bechdel Test is frequently cited in critical studies of literature, gender studies (Anthropy 2012; Power 2009; Thompson and Armato 2012), and in popular analyses (Hickey 2014; Sharma and Sender 2014), but has not been used as a measurement tool in other social scientific research to date. The Bechdel Test is a fairly blunt measure: It is not a test of how feminist a movie is and cannot measure the extent of stereotypical depictions. It merely captures the extent to which

Table 1Descriptive Statistics of Variables

Variable	Minimum	Maximum	Mean	SD	Median
MetaCritic Score	9	98	52.33	17.092	51.00
Box Office (in millions)	17	754	79.96	70.945	55.00
Bechdel	0	1	.43	.495	.00
Star	0	1	.17	.373	.00
Sequel	0	1	.12	.328	.00
Arthouse	0	1	.15	.360	.00
Production Budget (in millions)	.02	300.00	56.42	44.06	45.00
Adventure	0	1	.25	.432	.00
Animation	0	1	.08	.266	.00
Comedy	0	1	.43	.495	.00
Drama	0	1	.41	.493	.00
Fantasy	0	1	.14	.345	.00

n = 974.

women exist in the movie and have experiences independent of men. For studies intending to describe the nature of gender representation, the Bechdel Test would prove inadequate. However, given our need for a simple metric of the extent of independent female representation in films, the Bechdel Test is ideal as it is a composite measure of three metrics that have been used widely in existing research. A large number of studies have explored the sheer number of women represented in movies (Lauzen and Dozier 2005), the number of speaking lines woman have (Smith et al. 2010), and how male centric the content is (Smith and Choueiti 2011). The Bechdel Test essentially offers a low threshold for each of these, but requires that a film pass all three. By laying out three clear, dichotomous measures, as opposed to ratings on a scale, the Bechdel Test also requires fewer subjective judgments, increasing intercoder reliability. (By contrast, a measure of how "feminist" a movie is would be highly subjective even if intercoder reliability was high.)

Even with the Bechdel Test's minimal requirements, past evidence suggests that many movies do not pass. With any more rigorous measure of gender equality, passing would be so rare as to limit variation to the point where meaningful analysis would not be possible. Indeed, the low bar of merely including women with lives independent of men is precisely what makes the Bechdel Test a useful measure.

The Bechdel Test examines not only if multiple women are present, but also whether their lives have focal points other than men. We operationally defined "two or more named women" as "two or more named characters who are specifically identified as girls or women, portrayed by an actress, or who otherwise displayed stereotypically feminine characteristics to indicate gender."³ We defined "talk to each other" as "an exchange of words in which each of the women says something and speaks at least two times." Finally, the conversation was deemed "not about a man" if the women discussed "any topic unrelated to men at any time in the conversation."

The Bechdel Test variable was coded by the researchers and undergraduate students who volunteered to code movies. Coders were shown an online video explaining the Bechdel Test and were given oral and written instructions including the operational definitions of each of the three levels of the Bechdel Test. Then, they were given a paper form for each film and told to circle "Yes" or "No" for each of the three levels of the Bechdel Test, indicating which, if any, of the parts the film passed. Coders were also asked to flag any borderline cases for the researchers to review. The researchers manually entered the information from the paper forms into an electronic spreadsheet. All movies were obtained from the library, NetFlix, and personal collections.

While volunteer coders may be less experienced in content analysis coding, the inherent simplicity of the test and the precise definitions appear to have yielded a high degree of intercoder reliability. The researchers double-coded 10 percent of the sample (100 films) and found identical codings in all but four cases. We measured reliability using Krippendorf's alpha, one of the most conservative measures of reliability, which compares observed vs. expected intercoder disagreement. Krippendorf's alpha varies from 0 to 1 with higher numbers indicating greater agreement and $\alpha \ge .800$ generally regarded as an acceptable level of reliability (Krippendorf 2013). The percent agreement and Krippendorf's alpha for each level of the Bechdel Test were as follows: two or more women (100%, $\alpha = 1$), talk to each other (98%, $\alpha = .92$), about something other than a man (96%, $\alpha = .83$). We report each level of the Bechdel Test separately at the univariate level. However, for all other analyses, we used a dummy variable designating that the movie "passed" all three levels of the Bechdel Test (Bechdel movie = 1, others = 0).

The study controls for the size of the films' production budgets. While there are, of course, famous examples of big budget flops and microbudget success stories, movie studios tend to expect big returns on big investments. Movies with large production budgets typically have the advantage of celebrity casts and large advertising campaigns. For these reasons, we anticipate that movies with larger budgets will tend to earn more money at the box office. Data on production budgets were drawn from the Web site Box-OfficeMojo.com, which is the leading source of movie box office information on the Internet and is regularly cited in major publications like USA Today, The New York Times, and the Wall Street Journal. When production budget information was not available from BoxOfficeMojo.com, we attempted to locate the relevant information from reputable industry periodicals (e.g., Hollywood Reporter), the entertainment sections of major newspapers, and the movie's official press documents. The production budget in our sample was measured in millions of dollars, ranging from \$.02 m (that is, \$20,000 for Paranormal Activity) to \$300 m (Pirates of the Caribbean: At World's End).

The study also controlled for several genres of the films in the sample. Past research has shown that films that adhere to genre conventions are better received by audiences because they make the movie more intelligible (Hsu, Hannan, and Koçak 2009; Neale 2000). Using information from MetaCritic.com, a movie rating Web site owned by the Columbia Broadcasting System (CBS), we recorded all genres listed for each film. Then, we conducted a preliminary analysis to determine which genres were most closely linked with the primary variables of interest. The genre dummy variables used as controls in our analyses were Adventure, Drama, and Fantasy (genre = 1, other = 0).

Because a famous star can increase a film's box office performance on the strength of their personal charisma and reputation (Hennig-Thurau, Houston, and Walsh 2007), we included a variable for actor popularity. The star variable

was constructed using the Forbes Celebrity 100 list, which measures the fame of celebrities every year. The Forbes Celebrity 100 list is determined by an algorithm based on earnings, wealth, television/radio exposure, print exposure, popularity in web searches, and popularity with the public. In our study, movies were dummy-coded for the presence of a star who appeared on Celebrity 100 list in the year prior to the film's release (movie with a star = 1, other = 0). In our sample, 16.7 percent of the movies featured a star from the list.

We also controlled for the nature of the movie's distribution label. Some distribution labels are known for releasing blockbusters (Universal, Warner Bros., etc.), and others tend to distribute "art-house" films. These so-called arthouse films tend to be serious, complex, and aimed at a niche market rather than a mass audience. Films distributed on an art-house label may be more likely to feature an independent female presence and may affect box office performance. In the past, many artsy movies were made primarily by independent production companies, and this gave rise to the term "indie movies." Today, major studios frequently produce and distribute art-house movies, too (Baumann 2007). However, rather than distributing the film under the company's primary distribution label (e.g., Universal), the art-house movies are usually released on a subsidiary label dedicated to "indie films" (e.g., Focus Features, a division of NBC Universal). For these reasons, we developed a dummy variable for distribution by an art-house label.⁵ Using BoxOfficeMojo.com, we recorded the distribution studio of each movie. Then, we examined the distributors' Web sites for any mention of distributing "art-house," "art," "independent," "indie," and/or "eclectic" films. These art-house distribution labels were coded as 1, and all others were coded as 0. In our sample, 15.3 percent of the films were distributed on an art-house label.

Some past literature has focused on the role of film critics in influencing the public's interest in films. Hirsch, for example, has argued that film critics act as "surrogate consumers" who work to anticipate and represent their audiences' tastes (1972:649). At the same time, movie studios attempt to co-opt movie critics with targeted promotional campaigns. In turn, studios use critics' quotes to legitimate their films as part of their marketing to the public (Baumann 2007). Some research has shown that critics' review can, in fact, influence the public's decisions about movie attendance (King 2007). For this reason, we controlled for critical consensus using a composite score of movie critics' reviews taken from Metacritic (www.metacritic.com).

Metacritic converts movie reviews into scores on a 100 point scale (e.g., a movie that receives 3 of 4 stars would receive a Metacritic score of 75). Quantifying each review not only allows users to easily make comparisons across various reviewers, it also allows Metacritic to produce a composite rating for each film. The score is produced using a weighted average that gives more influence to some critics and publications due to their perceived quality and prestige within the field. Unlike some other movie compiler sites which use only a bimodal rating of "fresh" or "rotten," Metacritic's composite scores range from 0 to 100, capturing more subtle distinctions between films. According to Metacritic, scores from 81 to 100 are "Universally Acclaimed," 61-80 received "Generally Favorable Reviews," 40-60 are "Mixed or Average Reviews," 20-39 are "Generally Unfavorable Reviews," and 0-19 indicates "Overwhelming Dislike." The scores in our sample range from 9 to 98. While movie ratings themselves are fundamentally subjective, Metacritic scores offer one of the best measures of critical consensus in film and have been used in several academic inquiries (Eastment 2009; King 2007; Mahesh et al. 2010).

The final independent variable is a sequel dummy. If a movie was highly successful at the box office, we would expect that its sequel would also perform well, almost regardless of its actual quality. In all cases of ambiguity, we used the Internet Movie Database to determine whether a movie was a sequel (sequel = 1, others = 0). In our sample, 12.2 percent of the movies were sequels.

Results

A minority (44 percent) of the movies in our sample passed the Bechdel Test. This finding is a middle ground between Sharma and Sender's (2014) 36 percent and Hickey's (2014) 53 percent passing. While only 13 percent of the movies failed to have two named females, the women in the other 87 percent of the movies still clearly inhabit a world based around men. Twenty-eight percent of the movies never showed the women speaking to each other and, in another 15 percent, the women spoke exclusively about men. The majority of popular movies lack an independent female presence (Table 2).

Given the well-documented tendency of films to have fewer female characters in male-centric stories, we speculated that movies with an independent female presence might earn less at the box office. That, in turn, would make future female-driven movies less attractive to studio executives. As can be seen in the cross-tabulation of the Bechdel Test and box office performance in Table 2, the results support this premise (contrary to the popular analyses of Sharma and Sender [2014] and Hickey [2014]). Here, box office returns are broken into quartiles with the lowest quarter ranging from \$17.2 to 35.7 million and the highest quarter ranging from \$95.1–754 million. Of the films that fell within the first quartile of earnings, 55.1 percent did not pass the Bechdel Test. The second and third quartiles are fairly similar. By contrast, of the films in the fourth quartile (the films that made the most money), 64.4 percent did not pass

	Box Office (in millions)						
	\$17.2–35.7	\$35.8–54.6	\$54.7–95	\$95.1–754			
Bechdel test							
Fail (%)	55.1	53.6	55.0	64.4			
Pass (%)	44.9	46.4	45.0	35.6			

Table 2 Cross-Tabs of Bechdel Test and Box Office Earnings

n = 9/4; χ^- test, p < .05.

the Bechdel Test. These findings suggest that films that do not feature an independent female presence earn significantly more money and much of the effect appears to be concentrated in the upper quartile (p < .05) (Table 3).

Using a series of multivariate regression analyses (Table 3), we examined whether the negative effect of passing the Bechdel Test on box office performance would hold up when controlling for several other variables, particularly production budget. Model 1 in Table 3, which examines only the effect of the Bechdel Test, reaffirms the findings of our cross-tabulation: Movies that pass the Bechdel Test earn significantly less at the box office than those that do not pass the Bechdel Test (p < .05). Based on Model 1, Bechdel movies earn \$11.3 million dollars less than an otherwise similar film that does not feature an independent female presence.

Each subsequent model controls for the effect of additional independent variables. Model 2 incorporates three variables with statistically significant effects on box office performance: whether the movie is a sequel (p < .001), its critical appraisal as measured by MetaCritic Score (p < .001), and whether it features a major star (p < .001). These three variables improve the model greatly, explaining over 19 percent of the variation in box office performance $(R^2 \text{ for Model 2 minus } R^2 \text{ for Model 1})$. The effect of MetaCritic Score may suggest that positive reviews drive audiences to the theater or that higher quality films appeal to both critics and audiences alike. Either way, it suggests that better reviewed movies perform better at the box office. The effects of having a star and being a sequel indicate that using established actors and stories reduce the risk of a box office flop. Model 3 adds three genres—Adventure (p < .001), Fantasy (p < .001), and Drama (p < .001)—to the analysis. Simi-

	Model 1	Model 2	Model 3	Model 4	Model 5
Constant	84.843	8.118	4.876	3.796	-22.162***
Bechdel	-11.356*	-10.658**	-7.756*	-6.598*	2.582
MetaCritic Score		1.217***	1.254***	1.338***	1.067***
Sequel		63.575***	49.298***	51.370***	28.383***
Star		29.742***	31.564***	29.858***	13.507**
Adventure			22.910***	20.459***	-1.040
Fantasy			32.497***	31.092***	17.659**
Drama			-20.943***	-17.685***	-13.865***
Arthouse				-28.396***	-7.228
Production Budget (in millions)					.793***
R^2	.006	.199	.290	.309	.467

 Table 3

 Linear Regression of Box Office Earnings

larly, Model 4 adds a measure of whether the movie was produced by an "arthouse" production label (p < .001).

While all of these variables have significant effects on box office performance, it is the introduction of a measure of production budget in Model 5 that transforms our understanding of the relationships among the variables. First, it must be noted that production budget alone accounts for almost 16 percent of the variation in box office performance (increasing the model's R^2 to .467). Holding all else constant, for every additional \$10 million spent on production budget, a film earns \$7.93 million more at the box office (p < .001). After controlling for production budget, the effect of being released by an art-house label and being in the Adventure genre become non-significant. In other words, arthouse movies make less money largely because they have smaller production budgets and just the opposite is true of Adventure films.

More importantly for our purposes, production budget appears to explain much of the relationship between Bechdel and box office earnings, leading us to reject our hypothesis. While movies that pass the Bechdel Test do have significantly lower domestic box office totals than other films, the effect appears to stem from the fact that movies with large production budgets (also the biggest box office earners) tend not to pass the Bechdel Test. Put another way, movies with an independent female presence tend to have smaller production budgets and consequently earn less money. However, given two otherwise equivalent movies with the same budget, audiences do not appear to penalize those with a female presence, providing strong evidence for an upstream effect.

Discussion

Our results indicate that less than half of the most popular movies from 2000 to 2009 featured an independent female presence (i.e., passed the Bechdel Test). Although Bechdel movies did earn less money at the box office, it does not appear to be the case that the public is simply disinterested in movies featuring women (i.e., a downstream effect). Rather, there appears to be an upstream effect: the origin of gender inequality in films' box office performance occurs in the link between creators and the cultural object within the cultural diamond model (Griswold 2013). Reaffirming the findings of past literature (e.g., Hennig-Thurau, Houston, and Walsh 2007), the best predictor of a movie's box office performance is its production budget. Given the significantly smaller budgets of Bechdel movies, it is unsurprising that they earn less at the box office. In this section, we consider the implications of the results for existing theory, discuss some limitations of our study and propose several directions for future research.

Four decades of gender content analyses of television and movies have shown that women appear less frequently than men and often in male-centric stories. Our findings support this conclusion. The Bechdel Test offers an incredibly low threshold for female inclusion in a film. As we have noted, it also cannot measure whether a movie presents crass stereotypes of women or whether it is feminist in nature. Indeed, the popular movie Twilight (2008) often criticized for its stereotypical depiction of a dependent teen girl—passes the Bechdel Test. Yet, even with such a minimal test of the representation of women, the majority of movies in our sample did not pass. Of the upper quartile of films in box office earnings, only a third passed the test. These findings suggest that the overwhelming majority of the movies Americans watch are largely about men.

One of the innovations of the current study was linking a gender content analysis with the economic outcome of box office performance. In making this connection, we found Bechdel movies earned less at the box office than other films. Although we anticipated this pattern, we also hypothesized that Bechdel movies would suffer a penalty at the box office even when controlling for other factors like production budget. We argued that lower box office returns for Bechdel movies could be the product of downstream effects with gender operating in the relationship between the cultural object and the recipients in the cultural diamond with audiences preferring male-centric movies. If this were the case, given two otherwise equal movies, the film without an independent female presence would earn more at the box office. This interpretation was not supported by the results.

Alternatively, we proposed there might be upstream effects with gender operating at the institutional level with film industry gatekeepers allocating smaller production budgets to the type of movies that tend to have an independent female presence (e.g., romantic comedies) than the type of big budget movies that feature almost exclusively men (e.g., adventure and fantasy). Our findings reveal that, even taking genre into account, films that portray (even briefly) an independent female presence tend to have smaller production budgets than those that do not pass the Bechdel Test. However, holding all else constant, when film industry gatekeepers allocate comparable production budgets to movies with an independent female presence, those movies perform as well as others. Therefore, the small but observable relationship between gender content and box office performance appears to be the consequence of institutional upstream effects rather than individual-level downstream effects.

Past literature has shown that industry gatekeepers adopt routines and adhere to established formulas to reduce the risk of a box office flop. It is possible, as the controversial *Deadline Hollywood* report suggests, that studio executives and film producers consciously avoid allocating large production budgets to female-centric movies because they doubt the earning potential of such projects. On the other hand, it is possible that industry gatekeepers unknowingly allocate smaller budgets to the type of films that pass the Bechdel Test as they pursue a variety of other risk-averse routines and gendered organizational logics. Or that the types of film projects that feature an independent female presence tend to have more limited production needs. All of these possibilities suggest that Hollywood studios are "gendered organizations." The findings of this study demonstrate that, whether consciously or not, gatekeeping practices upstream lead to systematically smaller production budgets for Bechdel movies and, in turn, smaller box office returns.

This study helps us better understand the processes undergirding the underrepresentation of women in film, but there are several shortcomings. First, the Bechdel Test is an imperfect measure of women's representation. Even though a more rigorous standard of gender equality would likely find so few acceptable films as to render quantitative analysis nearly impossible, a dichotomous measure like the Bechdel Test cannot capture the immense variation in women's representation across films.

Second, our study establishes the importance of production budget for the box office performance of Bechdel movies, but we can only guess at industry gatekeepers' decision-making processes. It is possible that executives do not think of films featuring women as box office losers but, instead, select and fund male-centric movies for other reasons. Indeed, studio executives may feel constrained in their decision-making by institutional processes and/or industry norms. It is also not clear whether such decisions are affected more by processes of "segmentation" or "organizational logics" or some combination of the two. Access to film industry elites is sure to be limited (Ortner 2010), but future qualitative research ought to examine the ways in which producers and studio executives think about, assess, and discuss gender when making movie funding decisions.

Finally, the current research also tells us little about how Bechdel movies come to be. Smith and Choueiti (2011) found that family films with women "behind-the-scenes" as writers, directors, and/or producers had significantly more female speaking characters. Future research ought to examine how the demographic characteristics (especially race, gender, and age) of film creators influence the likelihood of producing and funding Bechdel movies. In pursuing this line of inquiry, scholars ought to pay close attention to intersectionality. For example, are female directors of different races and points in the lifecourse more or less likely to direct Bechdel films? Are white or older female directors more successful in budget negotiations with studios? Answering questions like these would allow us to better understand how Bechdel movies are created and why they typically have smaller production budgets.

Although the current study's findings offer only initial answers to questions regarding the effect of gender representation on box office performance, it advances our previous understanding of the processes behind the underrepresentation of women in film by linking a content analysis with box office data. Our results suggest that Bechdel movies, in which women are represented as leading lives independent of men, tend to make less money at the domestic box office. However, our analysis indicates Bechdel movies perform worse not because the public rejects stories about women downstream, but rather because of upstream practices that make it less likely for female characters to have a life independent of men in big budget movies.

ENDNOTES

*Please direct correspondence to Andrew M. Lindner, Skidmore College, Department of Sociology, 815 N. Broadway, Saratoga Springs, NY 12866, USA; e-mail: alindner@skidmore.edu.

¹Throughout this study, we refer variously to "Bechdel movies," movies with "independent female presence," and movies that pass or fail the Bechdel Test. We understand the Bechdel Test as a measure of whether a film has a female presence (two or more named females) and depicts women as having lives independent of men (by speaking with each other and having conversation on topics other than men). We use these terms interchangeably.

²Profitability (i.e., return relative to investment) is quite a different question. Arthouse films are often more profitable than blockbusters. However, due to our interest in consumer demand, we used a measure that better reflected the number of people who saw the film rather than its investment value.

³This complex definition was necessary to code the rare cases of gender ambiguity (for example, "female" robots in animated movies).

⁴We also ran all analyses using an ordinal measure of the Bechdel Test, which ranged from 0 to 3 capturing which of the parts of the Test the film passed. However, the analyses using the dichotomous and ordinal measures produced nearly identical results. For the sake of clarity, we employed the dichotomous measure.

⁵In preliminary analyses, we used a dummy variable measuring whether the studio distributing the film was actually an independent production company. When controlling for production budget, it produced no significant effects.

REFERENCES

- Acker, Joan. 1990. "Hierarchies, Jobs, Bodies: A Theory of Gendered Organizations." Gender & Society 4:139–58.
- Anthropy, Anna. 2012. Rise of the Videogame Zinesters. New York: Seven Stories Press.
- Arbitron. 2007. "The Arbitron Cinema Advertising Study." Retrieved April 10, 2012. http://www.arbitron.com/study/cinema_study.asp.
- Associated Press. 2013. "Swedish Cinemas Take Aim at Gender Bias with Bechdel Test Rating." *The Guardian Online*, November 6. Retrieved November 6, 2013. http://www.theguardian.com/world/2013/nov/06/swedish-cinemas-bechdel-test-films-gender-bias>.
- Asur, Sitaram and Bernardo A. Huberman. 2010. *Predicting the Future with Social Media*. Cornell University Library. Retrieved January 7, 2010 http://arxiv.org/abs/1003.5699>.
- Baumann, Shyon. 2007. Hollywood Highbrow: From Entertainment to Art. Princeton, NJ: Princeton University Press.
- Bechdel Test Movie List. 2010. Retrieved September 20, 2010. http://bechdeltest.com.
- Bielby, William T. and Denise D. Bielby. 1994. "'All Hits are Flukes': Institutionalized Decision Making and the Rhetoric of Network Prime-Time Program Development." *American Journal of Sociology* 99:1287–313.
- Collins, Rebecca. 2011. "Content Analysis of Gender Roles in Media: Where are We Now and where Should We Go?" Sex Roles 64:290-8.
- Corts, Kenneth S. 2001. "The Strategic Effects of Vertical Market Structure: Common Agency and Divisionalization in the US Motion Picture Industry." *Journal of Economics & Management Strategy* 10:509–28.
- Eastment, Diana. 2009. "The Moving Image." ELT Journal 63(2):191-3.
- Einav, Liran. 2007. "Seasonality in the U.S. Motion Picture Industry." The RAND Journal of Economics 38:127–45.
- Elberse, Anita. 2006. "The Power of Stars: Do Star Actors Drive the Success of Movies?" *Journal of Marketing* 71:102–20.
- Eliashberg, Jehoshua, Sam K. Hui, and Z. John Zhang. 2007. "From Story Line to Box Office: A New Approach for Green-Lighting Movie Scripts." *Management Science* 53:881–94.

- Eliashberg, Jehoshua and Steven M. Shugan. 1997. "Film Critics: Influencers or Predictors?" Journal of Marketing 61:68-78.
- Eschholz, Sarah, Jana Bufkin, and Jenny Long. 2002. "Symbolic Reality Bites: Women and Racial/ Ethnic Minorities in Modern Film." Sociological Spectrum 22:299-334.
- Faulkner, Robert R. and Andy B. Anderson. 1987. "Projects and Emergent Careers: Evidence from Hollywood." American Journal of Sociology 92:879–909.
- Fejes, Fred. 1992. "Masculinity as Fact: A Review of Empirical Mass Communication Research on Masculinity." Pp. 9-22 in Men, Masculinity, and the Media, edited by Steve Craig. Newbury Park, CA: Sage.
- Finke, Nikki. 2007. "Warner's Robinov Bitchslaps Film Women; Gloria Allfred calls for Warner's Boycott." Deadline Hollywood. Retrieved August 9, 2011. http://www.deadline.com/2007/10/ warners-robinoff-gets-in-catfight-with-girls/>.
- Fischer, Sabrina. 2010. "Powerful or Pretty: A Content Analysis of Gender Images in Children's Animated Films." Master's thesis. Auburn University.
- Gallagher, Margaret. 2005. Who Makes the News? Global Media Monitoring Project 2005. World Association for Christian Communication.
- Gerbner, George and Nancy Signorielli. 1979. Women and Minorities in Television Drama, 1969-1978. Philadelphia, PA: The Annenberg School of Communications, University of Pennsylvania.
- Gill, Ros. 1993. "Justifying Injustice: Broadcaster's Account of Inequality in Radio." Pp. 75-93 in Discourse Analytic Research: Repertoires and Readings, edited by Erica Burman, Ian Parker. London, UK: Routledge.
- Gitlin, Todd. 1983. Inside Prime Time. London, UK: Routledge.
- Griswold, Wendy. 2013. Cultures and Societies in a Changing World. Thousand Oaks, CA: Sage.
- Hennig-Thurau, Thorsten, Mark B. Houston, and Gianfranco Walsh. 2007. "Determinants of Motion Picture Box Office and Profitability: An Interrelationship Approach." Review of Managerial Science 1:65-92.
- Hether, Heather and Sheila Murphy. 2010. "Sex Roles in Health Storylines on Prime Time Television: A Content Analysis." Sex Roles 62:810-21.
- Hickey, Walter. 2014. "The Dollar-And-Cents Case Against Hollywood's Exclusion of Women." FiveThirtyEight. Retrieved April 1, 2014. .
- Hirsch, Paul M. 1972. "Processing Fads and Fashions: An Organizational-Set Analysis of Cultural Industry Systems." The American Journal of Sociology 77:639–59.
- Hsu, Greta, Michael T. Hannan, and Özgecan Koçak. 2009. "Multiple Category Memberships in Markets: An Integrative Theory and Two Empirical Tests." American Sociological Review 74:150-69.
- Jolliffe, Lee. 1993. "Yes! More Content Analysis." Newspaper Research Journal 14:93-7.
- King, Timothy. 2007. "Does Film Criticism Affect Box Office Earnings? Evidence from Movies Released in the U.S. in 2003." Journal of Cultural Economics 31(3):171-86.
- Krippendorf, Klaus. 2013. Content Analysis: An Introduction to Its Methodology. Thousand Oaks, CA: Sage.
- Lampel, Joseph and Jamal Shamsie. 2003. "Capabilities in Motion: New Organizational Forms and the Reshaping of the Hollywood Movie Industry." Journal of Management Studies 40:2189-210.
- Lauzen, Martha M. and David M. Dozier. 2005. "Maintaining the Double Standard: Portrayals of Age and Gender in Popular Films." Sex Roles 52:437-46.
- Lindner, Andrew M. 2012. "An Old Tool with New Promise." Contexts 11:70-2.

- Liu, Yong. 2006. "Word of Mouth for Movies: Its Dynamics and Impact on Box Office Revenue." Journal of Marketing 70:74-89.
- Lorber, Judith. 1994. Paradoxes of Gender. New Haven, CT: Yale University Press.
- Mager, John and James G. Helgeson. 2011. "Fifty Years of Advertising Images: Some Changing Perspectives on Role Portrayals along with Enduring Consistencies." Sex Roles 64:238–52.
- Mahesh, Joshi, Das Dipanjan, Kevin Gimpel, and Noah A. Smith. 2010. "Movie reviews and revenues: an experiment in text regression." Proceedings of the 2010 Annual Conference of the North American Chapter of the Association for Computational Linguistics 293-6.
- Neale, Stephen. 2000. Genre and Hollywood. London, UK: Routledge.
- Neelamegham, Ramya and Pradeep Chintagunta. 1999. "A Bayesian Model to Forecast New Product Performance in Domestic and International Markets." Marketing Science 18:115–36.
- Ortner, Sherry B. 2010. "Access: Reflections on Studying Up in Hollywood." Ethnography 11:211-
- Peterson, Richard A. and N. Anand. 2004. "The Production of Culture Perspective." Annual Review of Sociology 30:311-4.
- Power, Nina. 2009. One-Dimensional Woman. Hants, UK: O Books.
- Risman, Barbara J. 2009. "From Doing to Undoing: Gender as We Know It." Gender & Society 23:81-4.
- Rudy, Rena M., Lucy Popova, and Daniel G. Linz. 2010. "The Context of Current Content Analysis of Gender Roles: An Introduction to a Special Issue." Sex Roles 62:705-20.
- Sharma, Versha and Hanna Sender. 2014. "Hollywood Movies with Strong Female Roles Make More Money." Vocativ. Retrieved June 6, 2014. http://www.vocativ.com/culture/celebrity/ hollywood-movies-strong-female-roles-make-money/>.
- Smith, Stacy and Marc Choueiti. 2011. Gender Disparity on Screen and Behind the Camera in Family Films. Executive Report for the Geena Davis Institute. Retrieved June 15, 2014. http://seejane.org/wp-content/uploads/full-study-gender-disparity-in-family-films-v2.pdf.
- Smith, Stacy, Katherine Pieper, Amy Granados, and Marc Choueiti. 2010. "Assessing Gender-Related Portrayals in Top-Grossing G-Rated Films." Sex Roles 62:774-86.
- Thompson, Martha E. and Michael Armato. 2012. Investigating Gender. Malden, MA: Polity Press.
- Williams, Christine L., Chandra Muller, and Kristine Kilanski. 2012. "Gendered Organizations in the New Economy." Gender & Society 26:549-73.