The antecedents and consequences of restrictive age-based ratings in the global motion picture industry

Mark A.A.M. Leenders [a,1], Jehoshua Eliashberg [b,*]

[a] *Amsterdam Business School, Dept. of Marketing, University of Amsterdam, Plantage Muidergracht 12, 1018 TV, Amsterdam, The Netherlands*
[b] *Marketing Department, The Wharton School, University of Pennsylvania, Philadelphia, PA 19104-6371, United States*

ARTICLE INFO

ABSTRACT

This article analyzes one key characteristic shared by a growing number of industries. Specifically, their products and services are continuously monitored and evaluated by local third-party ratings systems. In this study, we focus on understanding the local drivers of restrictive age-based ratings in the motion picture industry and the effect of local ratings on a movie's performance at the box office. The results show that there is a significant negative relationship between restrictive ratings and opening weekend box-office performance. However, we find no significant effect with respect to cumulative box-office performance. In the second part of the study, we focus on the local regulatory system's role as a key *driver* of restrictive age-based ratings in the motion picture industry. Interestingly, the results suggest that the composition of the board that rates the movie plays a key role. Including pediatrics, psychology, or sociology experts in the evaluation board instead of only parents or laypeople has a strong effect and tends to lead to more lenient rating behavior. In addition, we find that larger ratings boards tend to be more restrictive than smaller ones and that industry representation is not necessarily associated with less restrictive ratings. Countries with cultures characterized as uncertainty avoidant, collective, and feminine also seem to be most lenient in their ratings. The implications of the results are discussed from both international marketing and public policy perspectives.

© 2011 Elsevier B.V. All rights reserved.

1. Introduction

One of the key emerging challenges for managers of global products and services lies in the fact that national third-party ratings agencies evaluate their product before public consumption. Industries in which products are normally evaluated for public safety and protection (e.g., industries in which warning labels are issued) include the following: ethical drugs, food, financial services, motion pictures, music lyrics, computer games, automobiles, electrical appliances, toys, and construction materials.

Third-party ratings systems have unique features and serve different purposes across industries. Often the ratings are designed to provide information to the public and to enable informed decision making by individual consumers. For example, in the toy industry, ratings may be designed to help parents prevent children from swallowing toys. In the healthcare industry, ratings may provide information about the risks of certain medical prescriptions to various types of patients, such as pregnant women. In the motion picture industry, rating systems can be traced to the beginning of the 20th century and are often the focal point of discussions about censorship, free speech,

bad taste, and public health. These issues led to the creation of the motion picture industry's age-based ratings system ('rating system' hereafter), which is one of the most visible, well-known and used systems in the world (Leone & Houle, 2006). Understanding the roles of rating systems in industry-specific contexts will significantly contribute to knowledge development, as marketers in the motion picture industry face unique challenges and unanswered questions regarding the role of third-party movie ratings in the context of movie development and marketing decisions (Eliashberg, Elberse, & Leenders, 2006).

To date, one research stream on entertainment-related ratings has mainly focused on the effects of specific content on vulnerable groups in areas such as pediatrics and communication (e.g., Walsh & Gentile, 2001). Other researchers, such as De Vany and Walls (2002), indicated that R-rated movies underperform on all metrics, including box-office receipts. They argued that Hollywood makes too many R-rated movies (see Medved, 1992). Most research to date has focused on the U.S. market and employed the U.S. rating as a covariate to explain and predict the box-office performances of movies. However, to our knowledge, a gap still exists in our understanding of how movie ratings are generated and how they affect movie-going behavior and commercial performance on a global basis.

In this partially exploratory study, we start with the premise that the effects of a movie's ratings on its box-office performance are not straightforward because different forces are at play. On the one

* Corresponding author. Tel.: +1 215 898 5246; fax: +1 215 898 2534.
 E-mail addresses: m.a.a.m.leenders@uva.nl (M.A.A.M. Leenders),
eliashberg@wharton.upenn.edu (J. Eliashberg).
 [1] Tel.: +31 20 5256078; fax: +31 20 5253681.

hand, restrictive ratings may reduce the movie's market potential, as the movie is suitable for fewer consumers. This effect is called the *tainted fruit* effect. On the other hand, restrictive ratings provide additional information about the movie and evoke reactance behavior against restrictions that can lead to higher demand for these movies. This effect is called the *forbidden fruit* effect. Hence, the effects of movie ratings can be both positive and negative. Moreover, the drivers of restrictive ratings are not straightforward. Although a movie's characteristics (e.g., violence, nudity and sex) may seem likely to affect its rating, we still do not understand their relative impact on movie ratings. Ratings are established locally, and different countries have different regulatory systems that generate these ratings. Thus, in addition to questions on the effects of third-party movie evaluations on box-office performances, questions on the methods by which these evaluations are established have not been addressed in the extant literature. In this study we examine the following questions. What are the key drivers of a movie's rating in a local market? To what extent is the same movie evaluated differently in various local markets around the world? Can these differences be explained? What roles do the characteristics of a movie, the regulatory context, and the national culture play in generating the movie's rating?

The main contributions and conclusions of our study are as follows. First, we find that in a cross-national dataset of nine countries, the aggregate effect of restrictive movie ratings on demand is negative only for the opening weekend. Second, we observe differences in the restrictiveness of the ratings across countries. Specifically, the U.S., the U.K. and Australia are the most restrictive, and France, Spain and Italy are the most lenient. With respect to the drivers of these ratings, we find that violence, gore, and sex tend to be strong catalysts of restrictive ratings in movie industries around the world. Moreover, the local regulatory system and the local culture also play an important role in explaining the differences between movie ratings in different countries. Board composition is also associated with more restrictive ratings, especially if laypeople are involved. The size of the evaluation board matters as well. Larger boards tend to rate movies more restrictively. Hence, in the context of international marketing research, our study addresses the paucity of research on regulatory systems in general. Finally, we demonstrate how local cultural characteristics, such as uncertainty avoidance, masculinity, and individualism, combined with other drivers affect the restrictiveness of the ratings system.

In the U.S., the movie industry's ratings system is an example of a voluntary system with an independent board of parents that has evolved since 1922 and employs laypeople to enact the board's evaluations.[2] From the public's standpoint, the rating has informational value in that it indicates what the public can expect from the movie experience and which audience demographics best fit the movie's content. From a marketing standpoint, the system may pose a challenge because it may reduce the potential audience. For example, in the U.S., the NC-17 rating is called "box-office poison" (Guardian Unlimited, July 9th 1999). In addition, a restrictive rating may cause trade-offs between artistic quality and commercial profitability (Medved, 1992).

As can be seen from Table 1, there are indications that R-rated movies generate less receipts at the U.S. box office. For instance, of the top 20 grossing films in 2008, only three movies were R-rated,

Table 1
Highest grossing movies in the U.S. in 2008.

	2008 box office (US$M)	Rating
The Dark Knight	$531.0	PG-13
Iron Man	318.4	PG-13
Indiana Jones and The Kingdom of The Crystal Skull	317.1	PG-13
Hancock	227.9	PG-13
Wall-E	223.8	G
Kung-Fu Panda	215.4	PG
Madagascar: Escape 2 Africa	177.0	PG
Twilight	176.9	PG-13
Quantum of Solace	166.8	PG-13
Dr. Seuss' Horton Hears A Who	154.5	G
Sex and the City	152.6	R
Mamma Mia!	144.1	PG-13
Chronicles of Narnia: Prince Caspian	141.6	PG
The Incredible Hulk	134.8	PG-13
Wanted	134.5	R
Get Smart	130.3	PG-13
Four Christmases	118.2	PG-13
Juno	112.0	PG-13
Tropic Thunder	110.5	R
Bolt	109.9	PG

and none of them were in the top 10. In addition, there are variations in the movie rating systems utilized by different countries. In the U.S., the Motion Picture Association of America (MPAA) is responsible for rating movies. In the U.K., movie ratings are provided by the British Board of Film Classification (BBFC), and in Germany, these ratings are enacted by the Voluntary Self-Regulation of the Film Industry (FSK). Almost all countries have their own ratings agencies that are involved in many different types of media. Table 2 illustrates the differences in the ratings for the movies *Godzilla* and *My Best Friend's Wedding* in three countries (i.e., U.S., Australia, and Sweden). Appendix 1 shows the minimum ages recommended by the rating systems in different countries.

At first glance, local movie evaluation/rating systems clearly appear to have different reactions to movies that contain certain characteristics, such as violence, sex, and gore. In our study, we do not explicitly focus on the question of whether certain countries are more lenient to attributes such as violence and nudity. Instead, we adopt a broader perspective and focus on the drivers (i.e., movie ingredients, institutional board characteristics, and cultural characteristics) and effects of the *restrictiveness* of ratings on the demand for and the performances of movies. Thus, by studying movies that go through different ratings processes, we can learn about the effects of these processes and the methods by which these processes can be potentially managed.

The calls for both more and less restrictive ratings have been recurring topics in public policy circles. For example, the British parliament urged the BBFC to toughen its stand on violence in movies and video games (Variety, 2008). Our study offers important insights into how one can exercise some indirect control over movie ratings without directly interfering with the ratings board's decisions. Because

---

[2] In the U.S., the ratings board consists of women and men representing a range of racial and employment backgrounds. Their main goal is to evaluate movies and provide parents and moviegoers with advance information on films. In doing so, the ratings board enables moviegoers to determine which movies they will allow their children to watch. The MPAA ratings system originated from the filmmakers' community to preempt government regulation (Peacock, 2001). Its ratings are as follows: G (for general audiences), PG (parental guidance suggested), PG-13 (parents strongly cautioned), R (restricted), and NC-17 (no one 17 and under is admitted). Not every movie has to be submitted to the ratings board. However, movies that are not rated (NR) do not run in mainstream cinemas.

Table 2
Example of the differences in age ratings in different countries for the same movie.

Movie	Country	Rating	Age
Godzilla	US	PG13	13
Godzilla	Aus	PG	10
Godzilla	Swe	11	11
My Best Friend's wedding	US	PG13	13
My Best Friend's wedding	Aus	M	15
My Best Friend's wedding	Swe	Btl	All

Note: 'Age' is the minimum recommended age for watching the movie without supervision.

direct interference tends to provoke questions about censorship, we show how one may change movie rating system outcomes by restructuring the ratings boards.

2. Relevant literature, conceptual framework, and research hypotheses

Given the increasing reach and influence of the media, its direct link to democracy and freedom of speech, and the growing amount of harmful and offensive material published by the media, the media has a long tradition of balancing commerce and public responsibility. According to some media representatives and scholars, the rating system achieves its purpose. However, others claim that the system is not working (Garry & Spurlin, 2007).

Although the potentially negative effects of media on children have received attention in areas such as pediatrics and communication research (e.g., Bushman & Stack, 1996; Christenson, 1992; Walsh & Gentile, 2001), only a few studies have addressed the role and effects of *age-based ratings* on movie-going behavior from an international marketing perspective. Several descriptive and predictive studies have employed MPAA ratings as one possible determinant of the success of movies (e.g., Sawhney & Eliashberg 1996; Neelamaghan & Chintagunta, 1999; Ravid, 1999). These studies typically find that the G, PG, and, to a lesser extent, PG13 ratings are associated with better performance metrics, such as domestic box-office ticket sales, video revenues, and return on investment. These types of movies may exhibit superior performances because more people can watch movies with lenient ratings. In the same spirit, De Vany and Walls (2002) analyzed whether Hollywood studios produce too many R-rated movies. They found that slates emphasizing movies with restrictive ratings may be costly to the studios in terms of revenues, return on production costs, and profits. In another study, Bagella and Becchetti (1999) found that the most restrictive Italian rating, which was measured as a binary variable, had no significant effect on theater admissions (VM18, see Appendix 1).

2.1. The consequences of movie ratings

Third-party product evaluations and movie ratings are publicly communicated to the product's consumers, such as target audiences and parents. Christenson (1992) tested the effect of the U.S. recording industry's ratings on the attitudes of children. A group of middle school students gave lower evaluations to music albums whose covers had advisory labels indicating explicit lyrics, which implies that the album should only be bought by mature audiences. The children also reported less interest in buying these types of albums. This phenomenon has been called the 'tainted fruit' effect (Bushman & Stack, 1996; Christenson, 1992; Grier, 2001). In general, the tainted fruit theory posits that "warning labels should decrease the attractiveness of a given product because the product might harm the consumer" (Bushman, 1998). Earlier studies on warning labels on products such as alcohol and tobacco have shown that the labels' effectiveness varies but that warning labels can effectively stimulate safe behaviors, especially in familiar product categories in which the cost of compliance is low (Argo & Main, 2004). The tainted fruit effect may be especially strong in the motion picture industry because the largest demographic of moviegoers is families with children under 12 years old (MPAA audience study 2004). Hence, if a media product receives a more restrictive rating and if a tainted fruit effect exists, then the product will be consumed by less people (i.e., only those who consider it appropriate) and its business performance will be lower compared with a media product with broad appeal. Thus, we state our null hypothesis as follows:

H1(0). The relationship between the restrictiveness of a movie's rating and the commercial performance of the movie is negative.

Having stated the null hypothesis, we must analyze the alternative hypothesis. In contrast to the tainted fruit effect, a forbidden fruit effect may also exist. According to reactance theory (Brehm, 1966), if a person's freedom to behave in a particular setting is threatened, then the person will experience an unpleasant motivational state in which he or she is pressured to re-establish the lost freedom. Hence, consumers may respond to a restrictively rated product as if it was a forbidden fruit (Bushman & Cantor, 2003; Bushman & Stack, 1996; Klein, 1993; Pechmann & Shih, 1999). For example, a statement such as "This Product is for Adult Users Only" may stimulate young people to consume it because they seek to simulate adult behavior. Scholars have tested the forbidden fruit theory in different settings, including restroom graffiti, anti-smoking warnings, fat food warnings, drinking legislation, and media products (e.g., Bushman & Cantor, 2003; Pechmann & Shih, 1999). For media products such as movies, an adolescent below the recommended age may resist the restriction and actually be tempted to watch it instead. Morkes, Chen, and Roberts (1997) studied the effects of different types of audience restrictions, including MPAA movie ratings, on adolescents. The researchers asked their respondents to read brief descriptions of movies, which were randomly labeled with an MPAA classification, and then rate the attractiveness of the movies. The researchers found that the adolescents' desires to see the film increased as the rating's restrictiveness increased. In particular, boys seem to experience high levels of reactive behavior (Brehm & Weinraub, 1977; Bushman & Cantor, 2003).

The forbidden fruit and tainted fruit effects are likely to be complemented by a range of other forces that ultimately result in a net ratings effect within a local market. Ratings are likely to provide a cue to moviegoers about the movie's content and the attractiveness of the movie. Half of the tickets in the U.S. are sold to audiences between 12 and 29 years old (MPAA audience study 2004). Some portion of the tickets may have been purchased by audience members for whom the movie is not recommended. Garry and Spurlin (2007) showed that many young moviegoers can and do see inappropriate films in the cinema, even in countries with considerable legislation and enforcement mechanisms to prevent such behavior. The percentage of moviegoers who ignore ratings is much higher for children in public schools (71%) than for children in private schools (29%). Hence, we formulate the alternative hypothesis as follows:

H1(A). The relationship between the restrictiveness of a movie's rating and the commercial performance of the movie is positive.

In the next section, we discuss the antecedents of the local ratings.

2.2. The antecedents of ratings

In addition to investigating the *consequences* of age-based ratings on the commercial performances of movies in different countries, marketing researchers and managers must understand the *drivers* of these ratings. From a public policy standpoint, for example, third-party rating systems are important for consumer protection, and their effectiveness can be enhanced if one understands their antecedents.

2.2.1. The characteristics and ingredients of movies

The process of rating movies entails the evaluation of their ingredients. Movie ingredients are used for creating entertainment experiences as well as for artistic and esthetic expression, but they may be hazardous to the public or to certain vulnerable groups. Risky ingredients are classified in categories such as 'violence', 'sex', nudity, and 'language' by ratings agencies. These categories constitute creative elements that may be beneficial to a movie's storyline and desirable to certain audiences, but these elements may also cause potential harm or disruption to adolescents who may imitate the "real-world" behavior that they see in the movie. In particular, violence and its effects on society have received the most attention and publicity. Previous

studies have estimated that children watch 21–23 h of television per week on average, and from cartoons alone, children may watch 200 violent incidents per week. In addition, children may view 16,000 simulated murders and 200,000 violent acts by the time that they complete elementary school (Caron, 2001; Federal Communications Committee, 2003). Controversial elements, such as violence, are often augmented with potentially less harmful elements, such as rough humor and sensuality. Sex also plays an important role as a controversial ingredient because it may relate to real-life issues, such as teen pregnancies. Hence, we propose the following hypothesis:

H2. The presence of risky ingredients in movies is associated with more restrictive ratings for movies.

2.2.2. The characteristics of local evaluation boards

During a typical evaluation process, the members of the national ratings board view a movie, discuss it, and vote on the film's rating. If the movie's distributor is not satisfied with this rating, he or she can re-edit the film and re-submit it. However, distributors who do so face strong resistance from the creative community (i.e., director, actors, and actresses) and often face a costly delay in the movie's premiere. The ratings boards generally aim to reflect the norms of society, and the board members often collect local feedback (e.g., complaints and surveys) from the public to fine-tune their judgments.

We focus on three key characteristics of the ratings board: (1) its composition in terms of experts versus laypeople (i.e., novices), (2) the extent of industry representation within the board, and (3) its size.

Prior researchers have extensively studied the differences in risk assessment between experts and laypeople (e.g., Fischhoff, Lichtenstein, Slovic, Derby, & Keeney, 1981). The research has shown that experts do not necessarily assess risks better than laypeople, even if both are presented with the same available data. One reason is that many controversial issues force experts to move beyond the limits of the available data and convert their incomplete knowledge into judgments by falling back on intuitive processes and rules of thumb similar to those employed by laypeople.

Previous research has also shown that laypeople tend to be more cognitively involved in issues related to their self-interests and the interests of people with whom they empathize. As a layperson's level of personal interest in a topic or event increases, he or she thinks more about the issue as well. People work hard to form opinions on topics signifying positive or negative outcomes that are relevant to their personal interests (Petty & Cacioppo, 1981, 1986). As a result, cognitively involved people are likely to perceive more types of risks and frame these risks as losses (Kahneman & Tversky, 1973). We expect to see similar behavior in the case of movie ratings if parents make up parts of the ratings boards. In addition, people who are highly involved in an issue tend to see *all* topic-related arguments as important (Heath & Douglas, 1990). In Karger and Wiedemann's (1997) environmental risk assessment study, they stated the following: "Laypeople have a 'worst-case' scenario in mind. Furthermore, laypeople are quite sure of their assessment." The evidence also suggests that during criminal trials, laypeople who are acting as jurors tend to respond emotionally to various facts and, consequently, impose stiffer punishments than professionals. Moreover, during criminal trials, professional judges tend to be more lenient in the court than the laypeople who are acting as jurors (Institute for Security Studies, 2000). Based on the above information, we hypothesize the following:

H3. The presence of laypeople in an evaluation board is associated with more restrictive ratings.

Another characteristic of evaluation boards that may drive differences in ratings is the extent to which industry representatives are involved in the rating process. Interestingly, the degree of industry representation varies in motion picture industries around the world. In some countries (e.g., Australia, Hong Kong, and Italy), the industry is not involved with any ratings boards. In the U.S. and the U.K., there is some degree of industry representation because the ratings are provided by industry-backed organizations, such as the MPAA and BBFC, respectively. In Germany, some members of the ratings boards are appointed by the industry. Because economic considerations drive most industries, including the mainstream movie industry, industry representation is likely to lead to more lenient evaluations that do not 'taint the fruit' and, thus, exclude too many moviegoers.

H4. Industry representation in ratings boards is associated with more lenient ratings.

The third characteristic of evaluation boards over which countries often differ is the size of the board. Prior researchers have extensively studied how decision makers shift toward risk seeking or risk avoidance (Crott & Zuber, 1983; Myers & Lamm, 1976; Stoner, 1961; Sunstein, 2000). Both shifts increase in amplitude with group size (Vidmar & Burdeny, 1971). However, whether the decision makers shift toward risky or safe depends on the situation. The study by Reynolds, Joseph, and Sherwood (2009) is particularly relevant to our context. They found that decision makers are less risk-averse in making decisions for themselves and more risk-averse in making decisions that affect others. Hence, we expect evaluation board members to be generally cautious, and we predict that this tendency is stronger for larger boards. In larger ratings boards, negative information is also likely to receive more weight than positive information (Rozin & Royzman, 2001), and the board members may identify more controversial aspects in a movie. Thus, we hypothesize the following:

H5. Larger evaluation boards are associated with more restrictive ratings.

2.2.3. The characteristics of the local culture

Research on international marketing has always been an active area of academic study. In the media space, global consumer segments are beginning to emerge. Nowadays, advances in communication technology are shrinking distances and exposing consumers to a myriad of different influences beyond their national borders. As a result, many studies have attempted to identify global consumer segments based on demographic characteristics, such as age, gender, and buying power (Douglas & Craig, 1997). For example, global teens are being increasingly regarded as similar in terms of clothing, music and media consumption (Kjeldgaard & Askegaard, 2006). However, prior studies also suggest that distinct cultural traits underlie systematic differences in consumer behavior, especially in consumer responses to new products (Stremersch & Tellis, 2004).

This section explores the relationship between a country's culture and its local ratings boards. Hofstede (2001) defines national culture as the collective programming of the mind, and he shows that culture comprises the following components: masculinity, individualism, uncertainty avoidance, power distance, and time orientation. These dimensions have been widely employed in the international marketing literature, and they have been expanded and validated in numerous studies (Hofstede & Hofstede, 2005). We did not employ cultural differences to study the effects of the ratings across countries because we conceptualized culture as a driver of the rating rather than a moderating variable (i.e., a construct that affects the relationship between the rating and a movie's performance). There are two reasons for utilizing such a conceptualization. First, culture and ratings are not independent of each other. The raters who comprise the evaluation board are part of the same culture as the moviegoers, and they constantly receive feedback from society through surveys and complaints. Second, even if

one assumes independence, we do not have sufficient knowledge at this point to convincingly formulate hypotheses concerning the moderating impact of general cultural characteristics on the relationship between ratings and demand.[3]

In theory, the decision to restrictively rate a movie relates to the probability that adolescents will exhibit risky behaviors and the consequences of these behaviors on society (Grier, 2001; Walsh & Gentile, 2001). In practice, these decisions may relate to social norms regarding 'normality' and 'good taste' as well as the factors endangering the youth and society. In accordance with this goal, many ratings boards host periodic discussions on censorship. Here, we include those Hofstede's (2001) cultural dimensions that are most closely related to risk perceptions and risk attitudes.

One cultural dimension that fits our criteria is 'uncertainty avoidance'. Uncertainty avoidance refers to the extent to which the members of a culture feel threatened by uncertain or unknown situations (Hofstede, 2001). In uncertainty-avoidant cultures, risk-taking behavior is limited to known risks. The impact of violent movies is hard to assess and since the consequences of risky behavior can be severe (e.g., potential school shootings), we expect countries with higher levels of uncertainty avoidance to contain a more conservative population and thus the ratings to be less restrictive because these communities are less at risk.

Another relevant cultural dimension is 'individualism'. Individualism pertains to societies in which the ties between individuals are loose, and people are generally expected to look after themselves and their families (Hofstede, 2001). Because individualism is associated with entrepreneurship and risk-taking in general and because individualistic communities are more prone to risky behavior, we conjecture that individualism is associated with more restrictive ratings.

Finally, the degree to which a country's culture can be characterized as masculine may drive a board's rating decisions as well. Masculinity has received some attention in the areas of violence and sex in motion pictures but not in the rating process per se (Powrie, Davies, & Babington, 2004). Masculinity refers to the roles of men and women in social groups, especially in the middle class. In masculine countries, men are expected to be assertive, competitive and tough (Hofstede, 2001, p. 280, Hoppe, 1998). More masculine societies (e.g., Spain and Italy) tend to place greater value on wealth, success, ambition, material things, and achievement, whereas more feminine societies (e.g., Sweden) tend to place greater emphasis on people, helping others, preserving the environment, and equality (Hofstede, 2001; Lynn, Zinkhan, & Harris, 1993; Steenkamp, Hofstede, & Wedel, 1999). Hence, in masculine countries, the public at large may be more prone to controversial behavior. Accordingly, we expect more masculine countries to be more restrictive in their rating decisions, as shown by the following hypothesis:

H6(0). Restrictive ratings are more prominent in masculine countries and individualistic countries. Additionally, lenient ratings are more prominent in countries marked by high levels of uncertainty avoidance.

In addition to the risk perspective that motivates H6(0), we also argue that each ratings board is influenced by the underlying culture such that the board reflects the same values and beliefs as the citizens in the country. Therefore, in a more risk-averse country, the board may be more conservative as well, and the ratings handed out by the board will be restrictive. We can apply the same reasoning to the other cultural

variables (i.e., ratings board characteristics) to derive different conjectures. Based on this reasoning, we propose an alternative hypothesis:

H6(A). Restrictive ratings are more prominent in feminine countries and collective countries. Additionally, restrictive ratings are more prominent in countries marked by high levels of uncertainty avoidance.

3. Data, variables, and analytical approach

3.1. Data

We collected data on movie ratings, evaluation systems, national cultures, and box-office performances from April 1996 to the end of 2000 in the following nine countries: Australia, France, Germany, Hong Kong, Italy, Spain, Sweden, the U.K., and the U.S. The countries were selected based on the availability of the local box-office data and ratings data. Additionally, the countries needed to maintain the same evaluation system during the data collection period. For example, we did not include the Netherlands because of the changes to its ratings system during the period of the study.

The movies included in the database represent mainstream movies that were primarily produced by Hollywood studios and released by major global distributors and their labels. The movies in our data are distributed as follows: Columbia 13%, Disney 10%, DreamWorks 6%, Fox 13%, MGM 3%, Miramax 8%, New Line 8%, Paramount 11%, Polygram 5%, Universal 9%, and Warner Bros 12%. The movies' theatrical runs were substantially longer in U.S. theaters than elsewhere (See Table 3a).

Excluding marketing costs, the average budget of the movies in the sample was $55 million. The most prominent genres were drama (38%), action (36%), and comedy (33%), and the least popular genres were romance (21%), sci-fi (10%), and horror (7%). A complete list of the movies can be obtained from the authors.

3.2. The dependent variables

3.2.1. Opening weekend and cumulative box-office performances

The box-office performance of each movie in each country was measured in U.S. dollars. We measured the opening weekend and the cumulative box-office performances of all of the movies in the countries. These performance measures were compiled from the *Showbizdata.com* and *VNU/ACNielsen* databases. To adequately compare the movie's performance across the different countries, we normalized the dependent variables. For the opening weekend box-office receipts, the normalization is done by calculating the total receipts generated by the top ten movies in each weekend, and identifying the maximum ticket sales across all weekends in our 1996–2000 time period. This can be viewed as the best performing national weekend. By this approach, we derive for any given movie a relative demand share measure (DSH), which is comparable across different countries. This approach also reduces the need for incorporating a

Table 3a
Release patterns of movies in the sample.

Country	1–4 weeks in theaters	4–8 weeks in theaters	8–12 weeks in theaters	>12 weeks in theaters	Average number of weeks in theaters
Australia	52.4%	30.5%	8.9%	8.2%	4.86 (3.6)
France	63.1%	27.9%	6.5%	2.5%	4.28 (3.2)
Germany	54.7%	31.6%	10.8%	2.9%	4.94 (3.9)
Hong Kong	75.6%	22.8%	.8%	.8%	3.59 (2.4)
Italy	52.7%	37.3%	5.5%	4.5%	4.74 (3.6)
Spain	46.3%	35.0%	14.6%	4.1%	5.61 (4.2)
Sweden	43.0%	27.1%	18.7%	11.2%	6.84 (5.6)
U.K.	48.9%	34.8%	13.3%	3.0%	5.40 (3.9)
U.S.	5.2%	18.9%	21.7%	54.2%	14.64 (7.9)

Note: Data is from Showbizdata.com. The standard deviations are in parentheses.

---

[3] For example, one might expect that in 'masculine' countries, restrictive ratings lead to higher box-office ticket sales because 'masculine' audiences prefer this type of movie. However, when we ran an exploratory analysis on our dataset, we did not find a significant interaction between masculinity and the rating on either the opening weekend box-office sales or the cumulative box office sales. Similar exploratory analyses with respect to the other cultural characteristics indicated that no clear pattern in the relationship exists between ratings and movie attendance across cultures.

broad range of country-related controls in the analyses (such as cinema density). The demand share variable is given by:

$$DSHO_{ij} = \frac{\text{(Opening weekend) box office receipts of movie i in country j}}{\text{Maximum box office weekend receipts (top 10) in country j during data collection}}.$$

(1)

The same procedure was used to develop a normalized cumulative box office measure ($DSHC_{ij}$). The cumulative box office receipts of movie (i) were divided by the maximum cumulative box office of any movie in that country in the period of study.

3.3. The independent variables

3.3.1. Screens

Researchers have shown that the number of screens on which a movie plays in its opening weekend indicates the amount of money spent on marketing and explains the movie's performance to a significant degree (Elberse & Eliashberg, 2003; Eliashberg et al., 2006; Gemser, Leenders, & Wijnberg, 2008; Karniouchina, 2011; Litman, 1983). We obtained our data on the number of opening screens in the different countries from the Showbizdata.com and VNU/ACNielsen databases. Based on the local capacity, we normalized the number of screens during the opening weekend in a fashion similar to the demand share variables. We performed the normalization by identifying the weekend in which the ten best-performing movies utilized the highest number of screens. We used the weekend with the most screens allocated to the top ten as the base. Called Screen Share (SSH), this variable is used to measure the push to distribute and market the movie in the local market.

3.3.2. Age-based ratings

We obtained the age-based ratings systems of the movies released in the different international markets from the local ratings boards. Because most ratings systems utilize an age-based cut-off point, we coded the movie ratings corresponding to the recommended minimum age at which point adolescents can watch the movie without being accompanied by their parents. Our coding scheme (see Appendix 1) allowed us to perform comparisons across the countries. In the scheme itself, a lower score indicates a less restrictive rating by the evaluation board. For validity testing, we also coded the rating as a dummy (0 = 'all ages' and 1 = 'not for all ages').

3.3.3. Movie characteristics

We compiled a list of movie ingredients from the MPAA database (see www.MPAA.org.) and used this information as a rough proxy for the content of the movie. A score of '1' on an ingredient variable denotes the presence of the ingredient. Each movie can contain multiple ingredients. The MPAA considers a wide range of ingredients, but not all of them are stated in each movie and in our study. We included an ingredient if it appeared in at least 5 movies. These ingredients and the number of movies containing them (in parentheses) were the following: language (150), violence (93), sex (66), thematic elements (25), nudity (20), sensuality (18), drugs (16), (rough) humor (16), and gore (10).

3.3.4. Ratings board characteristics

We obtained the compositions of the ratings board from publications available at the local ratings boards' offices and from local guidelines regarding the formation of the board in each country. In particular, we directly measured the average size (labeled 'Size') of the board by the average number of people who had been serving on the board over time. We used a dummy variable to code for whether experts are included in the board (Expert = '0' if not included; Expert = '1' otherwise). We also captured the degree of industry representation in the board ('Indrep') such that Indrep = '0' if no representatives were present and Indrep = '1' otherwise (i.e., the

representatives are either directly or indirectly on the board or are represented only in the appeal processes).

3.3.5. Cultural characteristics

The national index scores for Hofstede's cultural characteristics, especially masculinity, individualism, and uncertainty avoidance, are available for all countries whose movies are investigated by this study (see Hofstede, 2001; Hofstede & Hofstede, 2005). For validation purposes, we also collected the national index scores from other studies, such as Hoppe (1998).

3.3.6. Control variables

We included several control variables for this study. First, we measured star power (e.g., Karniouchina, 2011) by log-transforming the ranking of the highest ranked actor on *Premiere*'s list of global star power, which ranks the 100 most powerful actors in film. For example, the actor with the highest star power had a star power variable with a score of '0' (log(1)). The actors and actresses not on the list were assigned a score of log (100). We compiled the movies' budgets and genres from IMdB. The budget was again log-transformed. We opted to utilize non-exclusive genre dummies, as a movie can have multiple genres. Table 3b provides some key descriptive statistics.

The table shows that the U.S. had the largest number of movies in the dataset for two possible reasons. First, because no box office data were available for some movies outside the U.S., we could not include them. Second, other missing data points may have caused U.S. movies to be overrepresented in our dataset. The descriptive statistics show that the U.S. had the most restrictive ratings (i.e., a mean age cut-off at 14.14 years old) and that France and Italy were the most lenient in their movie ratings (i.e., mean age cut-offs at 1.55 and 2.69, respectively). However, the ratings averages are not based on equal samples of movies, and we only have limited data for certain countries, such as Italy (n = 26). The average size of the board was the highest in the U.S. (11) and relatively small in Sweden (2) and the U.K. (2). Some countries relied only on laypeople as raters (e.g., Germany), whereas other countries included experts (e.g., France and Italy). In four countries (e.g., France), the industry exerted some type of influence on the rating process. The average budget of the movies in the U.S. sample was slightly lower than the budgets of the movies in the other countries, and the U.S. actors' star power ranks were somewhat higher than those of the other countries.

3.4. Analysis

To test the hypotheses concerning the consequences of the ratings systems, we use a log-linear multi level model. Our logistic formulation considers the DSH variables, which lie between 0 and 1, and can accommodate an S-shaped functional form (Dekimpe, Francois, Gopalakrishna, Lilien, & Van den Bulte, 1997). The model was linearized and specified for both DSHO (i.e., opening weekend box-office sales) and DSHC (i.e., the cumulative box-office sales). For a movie (i) in a particular country (j), we calculate the following model:

$$\ln\left(DSH_{ij}/1 - DSH_{ij}\right) = \alpha_i + \beta1\, SSH_{ij} + \beta2\, Rating_{ij} + \beta3\, Star\, power_i + \beta4\, Budget_i$$
$$+ \beta5\, Action_i + \beta6\, Comedy_i + \beta7\, Drama_i + \beta8\, Horror_i$$
$$+ \beta9\, Romance_i + \beta10\, Sci-fi_i + e_{2ij}$$

(2)

where:

$$\alpha_i = \alpha + u_{2i}.$$

To test the *antecedents* of the local ratings, we constructed a linear multilevel model with a movie-specific intercept. The local rating was the dependent variable, and the movie's ingredients, the characteristics of the local board, and the characteristics of the local culture

Table 3b
Key descriptive statistics.

	Aus	Fra	Ger	Hon	It	Spa	Swe	UK	US
Variable									
Mean ratings classification	12.99	1.55	10.76	10.38	2.69	7.70	10.61	12.81	14.14
Board variables									
Presence of experts	No	Yes	No	No	Yes	No	No	No	No
Industry representation	0	1	1	0	0	0	0	1	1
Average size of the board	5	7	7	9	7	7	2	2	11
Cultural variables									
Masculinity	61	43	66	57	70	42	5	66	62
Individualism	90	71	67	25	76	65	71	89	91
Uncertainty avoidance	51	86	65	29	75	86	29	35	46
Other variables									
Screen share (SSH)	.07	.08	.08	.10	.07	.09	.06	.09	.07
Starpower (rank)	69	64	67	65	61	67	61	66	76
Budget (Million $)	52	56	55	57	56	55	56	53	46
Action	.34	.42	.34	.41	.38	.34	.37	.35	.26
Comedy	.34	.33	.36	.29	.28	.33	.35	.33	.40
Drama	.38	.36	.37	.33	.42	.41	.37	.38	.40
Horror	.07	.08	.08	.10	.08	.07	.07	.07	.05
Romance	.23	.17	.21	.19	.19	.23	.22	.21	.21
Sci-fi	.10	.11	.11	.11	.11	.11	.09	.11	.07
Number of movies in the sample	157	93	136	73	26	122	107	135	227

were the independent variables. For a movie (i) in a country (j), we construct the following model:

$$
\begin{aligned}
\text{Rating}_{ij} = {} & \gamma_i + \beta 11\,\text{Violence}_i + \beta 12\,\text{Sex}_i + \beta 13\,\text{Drugs}_i + \beta 14\,\text{Nudity}_i + \beta 15\,\text{Humor}_i \\
& + \beta 16\,\text{Language}_i + \beta 17\,\text{Gore}_i + \beta 18\,\text{Thematic}_i + \beta 19\,\text{Sensuality}_i \\
& + \beta 20\,\text{Expert}_j + \beta 21\,\text{Indrep}_j + \beta 22\,\text{Size}_j \\
& + \beta 23\,\text{Mas}_j + \beta 24\,\text{Indiv}_j + \beta 25\,\text{Uncer}_j + e_{3ij}
\end{aligned}
\tag{3}
$$

where:

$$\gamma_i = \gamma + u_{3i}.$$

In both Eqs. (2) and (3), $e_{ij} \sim N(0, \sigma_e^2)$ and $u_i \sim N(0, \sigma_u^2)$. As is typical with hierarchical linear models, u_i is assumed to be independent of the residuals e_{ij}. We present the results in the following section.

4. Results

4.1. The consequences of the movies' ratings

Table 4 presents the results of the estimated parameters of model (2). We employed a restricted (or residual) maximum likelihood (REML) procedure to estimate the model. REML is commonly used to estimate multilevel models (Hox, 2002; Wuyts, Stremersch, Van den Bulte, & Franses, 2004). The metric −2RLL captures the model fit.

The restrictiveness of the rating indicates that a negative and significant coefficient exists for the opening weekend's demand share model ($\beta 2 = -.049$, $t = -10.76$, $p < .001$) but that a non-significant coefficient exists in the cumulative demand share model ($\beta 2 = .005$, $t = .77$ $p = .44$). To test the sensitivity of the estimates, we conducted an analysis by removing the U.S. from the dataset. In the smaller sample of the remaining eight countries, the coefficient of the opening weekend's demand share was similar ($\beta 2 = -.051$, $t = -10.20$, $p < .001$), but the coefficient

Table 4
The relationship between the age rating and the movie's performance (i.e., opening weekend's and cumulative demand share).

	Opening weekend B	t-value	Cumulative box-office B	t-value
Intercept	−4.652***	−12.05	−6.122***	−10.26
Rating	−.049***	−10.76	.005	.77
Screens (SSH)	15.596***	19.98	17.594***	14.91
Starpower (rank)	−.097*	−2.26	−.091	−1.36
Budget	.205*	2.44	.422***	3.29
Action	.303*	2.35	−.139	−.70
Comedy	.226*	1.94	.405*	2.25
Drama	.202	1.83	.016	.09
Horror	.812***	4.03	.288	.92
Romance	.275*	2.19	.491**	2.54
Sci-fi	.367*	2.31	.139	.56
−2RLL = 1765 (opening weekend)[a]			−2RLL = 2415 (cumulative box-office)	
Variance parameter estimate (S.E.):			Variance parameter estimate (S.E.):	
$\sigma u_{2i}^2 = .21$ (.03); Wald Z = 5.2***			$\sigma u_{2i}^2 = .51$ (.09); Wald Z = 5.3***	
$\sigma e_{2ij}^2 = .39$ (.02); Wald Z = 18.2***			$\sigma e_{2ij}^2 = .89$ (.05); Wald Z = 18.0***	

p*** < .001 (two-tailed), p** < .01 (two-tailed), p* < .05 (two-tailed).
[a] −2RLL = 1815 (opening weekend) if the rating variable was not included, which indicates a better fit for the full model.

of the cumulative demand share was again insignificant ($\beta2 = -.011$, $t = -1.52$, $p = .13$)). We also used the binary rating variable (i.e., $0 = $ 'all ages' and $1 = $ 'not all ages') instead of the continuous rating variable. The coefficient of the dummy variable was still negative and highly significant for the opening weekend share ($\beta2$ (dummy) $= -.61$, $p < .001$) but not for the cumulative box office share ($\beta2$ (dummy) $= -.09$, $p = .19$). Adding country dummies did not alter any of these results. In a single level regression, the Variance Inflation Factors (VIF) for all of the variables ranged from 1 to 1.5, except for budget, which had a VIF of 1.8. All of the VIFs are acceptable, and we found no signs of multicollinearity by using Belsley, Kuh, and Welsch (1980) condition index.

The budget and screen share variables generally show a positive association with performance, and the lower ranked actors are negatively related to performance. Based on these findings, we do not reject H1(0), where the opening weekend's demand share is used as a performance measure, in favor of H1(A). That is, the data provide support for the tainted fruit theory when a movie's box-office performance is measured by the opening weekend's box office receipts.

4.2. Different ratings systems for the same movie in different countries

We are also interested in the extent to which the ratings of the *same* movie vary across *different* countries. To explore this question more closely, we match the sample of movies shown in the U.S. and those shown in other countries while using the U.S. ratings as the benchmark in each pair of countries. Table 5 reveals some interesting insights.

As stated previously, a lower rating implies a more lenient evaluation. The table shows that the differences in the ratings of the same movie between the U.S. and the other countries are significant and that *all* non-U.S. countries tend to be *more* lenient relative to the U.S., with the largest differences occurring between the U.S. and France (14.04 vs. 1.55), the U.S. and Italy (14.54 vs. 2.69), and the U.S. and Spain (13.87 vs. 7.70). We further reduce the sample of matched movies to include only the movies with runtimes (i.e., length of the movie in minutes) in the non-U.S. market that are identical to those in the U.S. and thereby better control for their content. Our results show that similar differences exist for these movies as well.

The French system has a history of providing lenient ratings, especially for local movies that are often subsidized for cultural

Table 5
Comparing the age rating across countries (pairwise with the U.S.).

Samples matched on movie title[a]		F-value
US	Aus	
13. 94 (N = 156)	13.07 (N = 156)	F = 3.67 (p = .056)
US	Fra	
14.04 (N = 93)	1.55 (N = 93)	F = 466.24 (p = .000)
US	Ger	
13.87 (N = 136)	10.76 (N = 136)	F = 31.14 (p = .000)
US	Hon	
14.58 (N = 73)	10.38 (N = 73)	F = 44.16 (p = .000)
US	It	
14.54 (N = 26)	2.69 (N = 26)	F = 93.65 (p = .000)
US	Spa	
13.87 (N = 122)	7.70 (N = 122)	F = 74.74 (p = .000)
US	Swe	
14.23 (N = 107)	10.61 (N = 107)	F = 39.49 (p = .000)
US	UK	
13.91 (N = 135)	12.81 (N = 135)	F = 4.17 (p = .042)

[a] We checked for the possible effects of language dubbing. We deleted all movies that have 'language' as an ingredient that may have been dubbed in the foreign country. This additional restriction on the sample did not result in substantial differences in the results for any of the comparisons.

purposes. Interestingly, in early 2000, a lenient rating on the movie *Baise-moi*, which had obtained a '16' rating, triggered a broad discussion in France. Many argued that the movie should be X-rated because of its pornographic content and violence. The newspaper *Le Monde* called the movie 'sick', whereas others argued that the low-budget movie, which was shot without artificial lighting, was highly artistic.

The lenient ratings system in Spain may be related to the adverse effects of the dictatorship and heavy censorship that existed in the country until 1975. Most Spanish citizens are against censorship of any kind and prefer personal responsibility and liberalism. For example, cinemas in Spain rarely ask for age identification, which may explain the lenient ratings system in Spain. We found lenient ratings systems in certain continental European counties in accordance with the findings of a study conducted on behalf of the European Commission (Olsberg/SPI & KEA, 2003). In general, Anglo-American countries seem to be more restrictive. For example, in Australia, the restrictive ratings board censored the movie Hannibal. However, censorship cases in the U.K. have been rare (Variety, 2008).

To obtain additional insight with regard to a country's level of restrictiveness, we compared the proportion of movies that is rated "for all ages" across the countries in our study. In a pairwise comparison with the US, this analysis also shows that the U.S. is restrictive relative to other countries, as only 4% of American movies are rated "for all ages". In comparison, Australia (7%), the U.K. (7%), Hong Kong (9%), Sweden (9%), Germany (10%), Italy (19%), Spain (40%), and France (66%) all exhibited higher percentages of movies rated "for all ages". Except for the differences between the U.S. and Australia as well as the U.S. and the U.K., all of the other differences are significant ($p < .10$). When we explore the use of the more restrictive categories (e.g., "for ages 14 and higher"), we learn that the Anglo-American countries are again the most restrictive, as more than 50% of the movies are rated "for ages 14 and higher". In comparison, Spain (14%), France (24%), and Germany (30%) all exhibited lower percentages of movies rated "for ages 14 and higher". The pair-wise comparison between Spain and France is significant ($p < .10$), which shows that the Spanish system is especially averse to restrictions for young adults. This finding may again be related to the country's history of censorship. In sum, we note that the U.S., the U.K., and Australia have the most restrictive ratings systems and that France, Spain, and Italy have the most lenient ratings systems in our sample. In the next section, we will systematically study certain characteristics of ratings boards that relate to the ratings systems in the various countries.

4.3. The antecedents of ratings systems

Table 6 presents the coefficients estimated for the model described in Eq. (3).

Among the parameters of movie ingredients, violence shows the most positive association with restrictive ratings ($\beta11 = 3.03$, $p < .001$). The presence of sex and gore are also significantly associated with restrictive ratings. Interestingly, we find that (rough) 'humor' ($\beta15 = -1.20$, $p = .07$) and 'sensuality' ($\beta19 = -1.75$, $p < .05$) have negative associations with restrictive ratings. This finding suggests that certain types of rough humor and sensuality may be viewed by some ratings boards as harmless and perhaps even beneficial for younger viewers. This result raises the question of why 'beneficial' ingredients (e.g., educational and artistic elements) are not included in the ratings criteria. In other industries (e.g., the pharmaceutical industry), the benefit (i.e., efficacy) of the product is always weighed against its cost (i.e., side effects).

Table 6
Movie-, evaluation board-, and cultural drivers of the age rating, as specified in Eq. 3.

Independent variable	Hypoth	Coeff. (B)	t-value
Intercept		2.95^{***}	4.65
Movie's ingredients			
Violence	H2	3.03^{***}	6.83
Sex	H2	1.67^{***}	3.26
Drugs	H2	1.47	1.66
Nudity	H2	1.32	1.85
Humor (rough)	H2	-1.20	-2.05
Language	H2	.83	1.81
Gore	H2	2.58^{**}	2.65
Thematic elements	H2	$-.32$	$-.46$
Sensuality	H2	-1.75^{*}	-2.19
Board characteristics			
Experts	H3	-8.31^{***}	-21.89
Industry representation	H4	.38	1.17
Board size	H5	$.21^{***}$	5.29
Cultural characteristics			
Masculinity	H6	$.02^{***}$	3.75
Individualism	H6	$.08^{***}$	11.60
Uncertainty avoidance	H6	$-.05^{***}$	-7.35
$-2\text{RLL} = 5806$			

Variance parameter estimate (S.E.): $\sigma\, u^2_{3i} = 7.09\ (.92)$; Wald $Z = 7.71^{***}$
Variance parameter estimate (S.E.): $\sigma\, e^2_{3ij} = 9.76\ (.47)$; Wald $Z = 20.8^{***}$

$p^{***} < .001$ (two-tailed), $p^{**} < .01$ (two-tailed), $p^{*} < .05$ (two-tailed).

The other results in Table 6 provide general support for the hypotheses. The inclusion of experts in the board is significantly associated with more lenient ratings ($\beta20 = -8.31$, $t = -21.89$, $p < .001$). Larger boards are also associated with more restrictive ratings ($\beta22 = .21$, $t = 5.29$, $p < .001$). Additional analyses with different models showed that there is no support for an inverted U-shaped relationship between board size and restrictive ratings.

We found that the degree of industry representation is not associated with the ratings systems ($\beta21 = .38$, $t = 1.17$, $p > .24$).[4] All of the VIFs in the estimated single level regression are less than 1.9, with condition indexes less than 21 (Belsley et al., 1980). These results indicate that multicollinearity is not a concern. In sum, we find empirical support for H2, especially for violence, sex, and gore. Additionally, we find support for H3 and H5 but not for H4.

Table 6 shows that uncertainty avoidance and individualism have the strongest relationship with the restrictiveness of the ratings system. Cultures that have high levels of uncertainty avoidance tend to be less restrictive in their movie evaluations ($\beta25 = -.05$, $t = -7.35$, $p < .001$). Hofstede (2001, p.174 and p.199) discusses several examples in other areas that are in accordance with this finding. For example, Hofstede's analysis across 14 Western European Countries showed that speed limits on the highway are more lenient in countries with high levels of uncertainty avoidance. In our study, individualism and, to a lesser extent, masculinity are positively associated with more restrictive ratings (($\beta24 = .08$, $t = 11.6$, $p < .001$) and ($\beta23 = .02$, $t = 3.75$, $p < .001$), respectively). These findings show that the public may be more at risk in these communities, though we cannot observe the underlying mechanism. Thus, H6(0) is not rejected in favor of H6(A).

We conducted several robustness checks. First, we included cinematic characteristics, such as genre, into our analysis. However, we found that the movies' ingredients are more strongly associated with the rating and that the genres did not obtain significant coefficients. Similar results were also found when we used Hoppe's (1998) national

index scores. We also deleted the cultural variables and added each country's economic and demographic variables to the model. These additional models showed the same pattern of coefficients in terms of magnitude, sign, and significance as the model with the movies' ingredients and characteristics of the ratings boards. Finally, we included the U.S. rating as a predictor of the ratings in the other countries. Again, this finding did not substantially alter the results.

5. Summary and discussion

As more attention is given to customer protection in markets around the globe, rating systems are becoming an important factor that need to be considered by marketing managers. In this exploratory study, we developed and tested a framework for examining the antecedents and consequences of local age-based ratings systems in the motion picture industry. Our results suggest that movie ratings can play a significant role in determining a movie's commercial success around the world, especially during the opening weekend. Our explanations are related to the audiences' tainted fruit behaviors and to the notion that movies with more restrictive ratings may be less attractive to certain (i.e., infrequent) moviegoers, such as families.

Interestingly, our study does not show that restrictive ratings have similarly negative effects on a movie's cumulative box office performance. More research is needed to understand how the tainted fruit effect is complemented by the forbidden fruit effect over time and how other factors may eliminate the negative performance effect generated in the first weekend. However, we believe that our findings provide a first step to understanding the marketing implications of these systems with regard to aggregate demand.

Our results also indicate that ratings vary across countries, even when we control for some cutting (possible editing) and local adaptation of movies. In addition, our study provides insights into the local drivers of the rating processes of the highly restrictive boards. The characteristics of the rating organization (e.g., its size and the involvement of experts or laypeople) drive the leniency or restrictiveness of the ratings system is in a given country.

Moreover, our analyses of the effects of the ratings and their drivers yield interesting insights into each country's underlying culture. We see that uncertainty-avoidant cultures tend to be less restrictive, whereas masculine and individualistic countries are more restrictive. These findings are in accordance with the hypothesis that citizens are less at risk in uncertainty avoidant cultures and more at risk in masculine and individualistic countries. However, we cannot rule out other underlying mechanisms that may drive these cultural effects.

5.1. Conclusions and future research

In this study, we conducted several sensitivity analyses and employed different variables, measures, and models to validate our results. However, given the cross-sectional approach undertaken in this study, we cannot conclude that restrictive ratings have a negative *causal* effect on a movie's box-office performance. More experimental and longitudinal research is needed in this regard.

Our results show that certain cultural variables play an important role as *drivers* of a rating, but we cannot rule out other relationships and effects that may result from cultural differences. For example, more research is needed to examine the role of enforcement mechanisms (e.g., closing a theater's doors to underage moviegoers) and the perceptions of the relations between ratings and censorship in certain countries (e.g., Spain). We also observe a negative relationship between the restrictiveness of a movie's rating in a given country and its opening weekend box-office performance in the same country. However, we cannot observe the detailed mechanisms that enhance or diminish a movie's box-office performances in a given

---

[4] In addition, we considered separate dummies that measure different types of industry representation. Examples include whether industry representation exists only in the appeals process (Dumrep1) and whether industry can select the board's members (Dumrep2). The results were again not significant.

country. Although our hypotheses hold to a large degree, alternative explanations (e.g., the effect of 'good taste') cannot be ruled out. Additionally, given our initial approximations of each movie's ingredients, we cannot rule out that the performance-diminishing effect of the restrictive ratings is caused by less demand for certain ingredients instead of the rating per se.

Finally, we need to acknowledge some limitations of our study. First, our sample consists of a limited number of typical Hollywood productions. This sample allowed us to compare the movies while still obtaining different ratings for the movies in different countries. In addition, the U.S.-centric nature of the dataset is in accordance with the dominant position of American movies in the global market. However, this sample also limits our findings to these types of movies. Second, although our overall sample size is quite reasonable and comparable with those of other studies, the data at the country level is not sizeable enough to produce a hierarchical estimation or detailed insights in terms of the country-by-country effects of ratings. Third, we focused on national box-office data. Instead, using segment-level data (e.g., certain age groups) would allow one to answer important questions. For example, does a restrictive rating lead to different effects in different segments? What part of the effect is due to a 'forbidden fruit' effect? These questions are important, and they can be addressed with more appropriate data that may become available in the future. For instance, if more theater owners follow GKC Theatres' experiment of allowing non-accompanied teenagers to watch R-rated movies by showing a parent-approved pass card (Dennis, 2004), then the resulting data will be of great use to future studies.

5.2. Implications for international marketing in the motion picture industry

Movie producers and marketers who work in global industries in which movies are evaluated by local third-party evaluation systems should recognize the tradeoff between the product's appeal and its potential evaluation. The effects of ratings are measurable, and ratings can, to some extent, be managed indirectly. For example, an appropriate international business strategy may require film producers to adapt their movies to local conditions and edit their films by cutting different versions of the product. In the home entertainment sector, different versions of DVDs already exist, with different directors' cuts as well as different scenes and endings.

In addition, our study clarifies whether producers should appeal a local board's rating decision. As far as completed movies are concerned, executives are aware that appeals can be lengthy and costly and that they should be considered only in situations in which a negative rating will strongly affect the commerciality of the product. Our study shows that the relation between restrictive ratings and long-term negative commercial implications is unclear. However, we do find that restrictive ratings have a negative effect on a film's box-office performance on opening weekend. Thus, the more important the opening weekend is, the more important the rating becomes from a commercial perspective. Given that the industry tends to focus on the opening weekend of the movie, this finding is significant. Of course, non-commercial factors, such as the moral and ethical implications of broad consumption, should also be considered seriously. In fact, ratings systems in other media and entertainment contexts can be technologically enforced with innovations such as the V-chip for television sets and restrictive search bots for children surfing the Internet. In addition, there are software packages that delete certain images and words from DVDs while the DVDs are being played. More research is needed on the effects of these technologies and their significance to marketing in the global media.

From a public policy perspective, our results provide interesting insights into a key question linked to rating systems in general. What drives a ratings board to hand out restrictive ratings? Our empirical results imply that the structure of the board can become a policy instrument for establishing both more and less restrictive evaluations. More specifically, if the goal is to have more restrictive ratings, then this study suggests that the board should be large and include more laypeople. Of course, whether the goal should be to have more or less restrictive ratings depends on the actual public health and safety circumstances in the specific country of interest. Hence, the situation should be continually monitored, and the appropriateness of the board should be periodically reviewed.

Acknowledgments

Appendix 1. Local ratings in the sampled countries

Country	Organization	Local ratings	Minimum age
Australia	The Office of Film and Literature Classification	G	0
		PG	10
		M/MA	15
		R	18
France	Ministry of Culture	U	0
		−12	12
		−16	16
		−18	18
Germany	Freiwillige Selbstkontrolle der Filmwirtschaft	U	0
		6	6
		12	12
		16	16
		18	18
Hong Kong	Television and Entertainment Government Authority	I	0
		IIa	10
		IIb	13
		III	18
Italy	Governmental Organization	T	0
		VM14	14
		VM16	16
		VM18	18
Spain	Governmental Organization	T	0
		7	7
		13	13
		18	18
Sweden	Statensbiografbyrå (SBB)	Btl	0
		7	7
		11	11
		15	15
U.K.	The British Board of Film Classification	U	0
		PG	10
		12	12
		15	15
		18	18
U.S.	Motion Picture Association of America	G	0
		PG	10[a]
		PG13	13
		R	17

[a] PG ratings typically do not specify a cut-off age. However, telephone interviews suggest that the local boards employ the age of 10 as a practical benchmark (AUS, UK, US).

References

Argo, J. J., & Main, K. J. (2004). Meta-analyses of the effectiveness of warning labels. *Journal of Public Policy & Marketing*, *23*(2), 193–208.

Bagella, M., & Becchetti, L. (1999). The determinants of motion picture box office performance: evidence from movies produced in Italy. *Journal of Cultural Economics*, *23*(4), 237–256.

Belsley, D. A., Kuh, E., & Welsch, R. E. (1980). *Regression diagnostics.* New Jersey: Wiley.

Brehm, J. W. (1966). *A theory of psychological reactance.* New York: Academic Press.

Brehm, S., & Weinraub, M. (1977). Oppositional behavior in two-year-old children: A reactance theory interpretation. *Journal of Personality and Social Psychology*, *35*(11), 830–836.

Bushman, B. J. (1998). Effect of warning and information labels on consumption of full-fat, reduced-fat, and no-fat products. *Journal of Applied Psychology*, *83*(1), 97–101.

Bushman, B. J., & Cantor, J. (2003). Media ratings for violence and sex, implications for policy makers and parents. *American Psychologist*, *58*(2), 130–141.

Bushman, B. J., & Stack, A. D. (1996). Forbidden fruit versus tainted fruit: effects of warning labels for an attraction to television violence. *Journal of Experimental Psychology. Applied*, *2*(3), 207–226.

Caron, E. R. (2001). Blood, guts, and first amendment: regulating violence in the entertainment media. *The Kansas Journal of Law and Public Policy*, *11*(4), 89–112.

Christenson, P. G. (1992). The effects of parental advisory labels on adolescent music preferences. *Journal of Communications*, *42*(1), 106–113.

Crott, H. W., & Zuber, J. A. (1983). Biases in group decision making. In R. W. Scholtz (Ed.), *Decision making under uncertainty*. Amsterdam: Elsevier Science Publishers.

De Vany, A., & Walls, D. (2002). Does Hollywood make too many r-rated movies? Risk, stochastic dominance, and the illusion of expectation. *Journal of Business*, *75*(3), 425–451.

Dekimpe, M. G., Francois, P., Gopalakrishna, S., Lilien, G. L., & Van den Bulte, C. (1997). Generalizing about trade show effectiveness: A cross national comparison. *Journal of Marketing*, *61*(4), 55–64.

Dennis, J. (2004). *Movie chain lets parents give under 17s a pass-to R-rated fare*: Philadelphia Inquirer (June 13).

Douglas, S. P., & Craig, C. S. (1997). The changing dynamic of consumer behavior: Implications for cross-cultural research. *International Journal of Research in Marketing*, *14*(4), 379–395.

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

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

Federal Communications Committee (2003). *Comments on media policy coalition.* (April 21).

Fischhoff, B., Lichtenstein, S., Slovic, P., Derby, S. L., & Keeney, R. L. (1981). *Acceptable risk.* New York: Cambridge University Press.

Garry, P. M., & Spurlin, C. J. (2007). The effectiveness of media rating systems in preventing children's exposure to violent and sexually explicit media content. *Oklahoma City University Law Review*, *32*(2), 215–236.

Gemser, G., Leenders, M. A. A. M., & Wijnberg, N. M. (2008). Why some awards are more effective signals of quality than others: A study of movie awards. *Journal of Management*, *34*(1), 25–54.

Grier, S. A. (2001). The Federal Trade Commission's report on the marketing of violent entertainment to youths: Developing policy tuned research. *Journal of Public Policy and Marketing*, *20*(1), 123–132.

Heath, R. L., & Douglas, W. (1990). Effects of involvement on reactions to sources of messages and to message clusters. In J. E. Grunig, & L. R. Grunig (Eds.), *Public relations research annual* (pp. 179–183). Hillsdale, NJ: Lawrence Erlbaum Publications.

Hofstede, G. H. (2001). *Culture's consequences: Comparing values, behaviors, institutions, and organizations across nations.* Thousand Oaks, California: Sage Publications.

Hofstede, G., & Hofstede, G. J. (2005). *Cultures and organizations: Software of the mind.* NY: McGraw-Hill.

Hoppe, M. H. (1998). Validating the masculinity/femininity dimension on elites from nineteen countries. In G. Hofstede, & Associates (Eds.), *Masculinity and femininity: The taboo dimension of national cultures* (pp. 29–43). Thousand Oaks, CA: Sage Publishers.

Hox, J. J. (2002). *Multilevel analysis. Techniques and applications.* Mahwah, NJ: Lawrence Erlbaum Associates.

Institute for Security Studies (2000). Justice versus Retribution. *Monograph 45*, www.iss.co.za.

Kahneman, D., & Tversky, A. (1973). On the psychology of prediction. *Psychological Review*, *80*(4), 237–251.

Karger, C. R., & Wiedemann, P. M. (1997). Perception and appraisal of environmental problems: Laypeople's worst-case assumptions. *Society for Risk Analysis Annual Meeting.*

Karniouchina, E. V. (2011). Impact of star and movie buzz on motion picture distribution and box office revenue. *International Journal of Research in Marketing*, *28*(1), 62–74.

Kjeldgaard, D., & Askegaard, S. (2006). The globalization of youth culture: The global youth segment as structures of common differences. *Journal of Consumer Research*, *33*(2), 231–247.

Klein, R. (1993). *Cigarettes are sublime.* Durham, NC: Duke University Press.

Leone, R., & Houle, N. (2006). 21st century ratings creep: PG-13 and R. *Communication Research Reports*, *23*(1), 53–61.

Litman, B. R. (1983). Predicting success of theatrical movies: An empirical study. *The Journal of Popular Culture*, *16*(4), 159–175.

Lynn, M., Zinkhan, G. M., & Harris, J. (1993). Consumer tipping: A cross-country study. *Journal of Consumer Research*, *20*(3), 478–488.

Medved, M. (1992). *Hollywood vs. America: Popular culture & the war against traditional values.* New York: HarperCollins Publishers.

Morkes, J., Chen, H. L., & Roberts, D. F. (1997). Adolescents' responses to movie, television and computer game ratings and advisories. *Paper presented at the annual meeting of the International Communication Association, Montreal, Canada.*

Myers, D. G., & Lamm, H. (1976). The group polarization phenomenon. *Psychological Bulletin*, *83*(4), 602–627.

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

Olsberg/SPI and KEA (2003). *Empirical study on practice of the ratings of films distributed in cinemas television DVD and videocassettes in EU and EEA member states.* Prepared on Behalf of the European Commission*Final Report.*

Peacock, R. B. (2001). *The art of movie making: Script to scene.* Upper Saddle River, NJ: Prentice Hall.

Pechmann, C., & Shih, C. -F. (1999). Smoking scenes in movies and antismoking advertisements before movies: Effects on youth. *Journal of Marketing*, *63*(3), 1–13.

Petty, R. E., & Cacioppo, J. T. (1981). *Attitudes and persuasion: Classic and contemporary approaches.* Dubuque, Iowa: Wm. C. Brown.

Petty, R. E., & Cacioppo, J. T. (1986). *Communication and persuasion: Central and peripheral routes to attitude change.* New York: Springer-Verlag.

Powrie, P., Davies, A., & Babington, B. (2004). *The trouble with men: Masculinities in European and Hollywood cinema.* New York: Columbia University Press.

Ravid, S. A. (1999). Information, blockbusters, and stars: A study of the film industry. *Journal of Business*, *72*(4), 463–492.

Reynolds, D. B., Joseph, J., & Sherwood, R. (2009). Risky shift versus cautious shift: Determining differences in risk taking between private and public management decision-making. *Journal of Business & Economics Research*, *7*(1), 63–78.

Rozin, P., & Royzman, E. B. (2001). Negativity bias, negativity dominance, and contagion. *Personality and Social Psychology Review*, *5*(4), 296–320.

Sawhney, M. S., & Eliashberg, J. (1996). A Parsimonious Model for Forecasting Gross Box-Office Revenues of Motion Pictures. *Marketing Science*, *15*(2), 113–131.

Steenkamp, J. -B. E. M., Hofstede, F. T., & Wedel, M. (1999). A cross-national investigation into the individual and national cultural antecedents of consumer innovativeness. *Journal of Marketing*, *63*(2), 55–69.

Stoner, J. A. F. (1961). *A Comparison of Individual and Group Decisions Involving Risk.* Unpublished thesis, MIT, Cambridge, MA.

Stremersch, S., & Tellis, G. (2004). Understanding and managing international growth of new products. *International Journal of Research in Marketing*, *21*(4), 421–438.

Sunstein, C. R. (2000). Deliberative trouble? Why groups go to extremes. *The Yale Law Journal*, *110*(1), 71–120.

Variety (2008). *U.K. sensor bans serial killer film.* (February 28th).

Vidmar, N., & Burdeny, T. C. (1971). Effects of group size and item type in the "group shift" effect. *Canadian Journal of Behavioral Science*, *3*(4), 393–407.

Walsh, D., & Gentile, D. (2001). A validity test of movie, television, and video-game ratings. *Pediatrics*, *107*(6), 1302–1308.

Wuyts, S., Stremersch, S., Van den Bulte, C., & Franses, P. H. (2004). Vertical marketing systems for complex products: A triadic perspective. *Journal of Marketing Research*, *41*(4), 479–487.