UNIVERSITÄT PADERBORN
Die Universität der Informationsgesellschaft

FACULTY OF BUSINESS ADMINISTRATION AND ECONOMICS

Working Paper Series

Working Paper No. 2015-03

Combining Qualitative Comparative Analysis and Shapley Value Decomposition: A Novel Approach for Modeling Complex Causal Structures in Dynamic Markets

Daniel Kaimann

August 2014

Combining Qualitative Comparative Analysis and Shapley Value Decomposition: A Novel Approach for Modeling Complex Causal Structures in Dynamic Markets

Daniel Kaimann

University of Paderborn, Department of Business Administration and Economics
Warburger Str. 100, D - 33098 Paderborn, Germany
E-Mail: Daniel.Kaimann@wiwi.upb.de, Phone: +49 5251 60 3373

August 2014

Abstract Depending on which combination of factors is used in empirical analyses, regression results lead to varying levels of significance or even insignificance and, consequently, to inconsistent results. Linear algebra and linear regression models are apparently not able to analyze complex causal structures in dynamic markets. Boolean algebra, qualitative comparative analysis (QCA) and the game theoretical-based model of the Shapley value could be more suitable for covering dynamic market structures and, consequently, for helping us to understand what significantly affects complex cause-effect relationships. Using proprietary data from the volatile motion-picture industry, we show that a segmentation and brand extension strategy are sufficient for achieving high market performance and that certain conditons (e.g., production budget, critic reviews and brand extension products) appear particularly appropriate for gaining a competitive advantage.

Key words Dynamic markets, Qualitative comparative analysis, Shapley value, Motion-Picture economics

JEL classification C18 · C71 · L10 · L82

Acknowledgment The author thanks the participants of the Portsmouth Business School Research Seminar, the YU/Bruce Mallen Workshop for Scholars and Practitioners in Motion Picture Industry Studies, the Suffolk Economics Seminar Series, and the NBER Productivity Lunch Seminar for their constructive comments on this research. A special thanks goes to Bernd Frick, Claus-Jochen Haake, and especially Darlene Chisholm for her useful comments and discussions.
This work was partially supported by the German Research Foundation (DFG) within the Collaborative Research Centre "On-The-Fly Computing" (SFB 901).

Combining Qualitative Comparative Analysis and Shapley Value Decomposition: A Novel Approach for Modeling Complex Causal Structures in Dynamic Markets

1 Introduction

Most studies focus on cause-effect relationships. Causality is measured by statistical tests of significance. Partial effects of independent variables are summarized to a total effect. The addition leads to a unifinal result; an outcome that is reached from a distinct path. Apparently, linear combination models are appropriate to analyze the net effects of individual variables, but they are insufficient to examine the relationship of complex causal structures in dynamic industries. The introduction of interaction terms might be a solution to capturing complex causal structures in linear models, but the number of interaction terms quickly increases exponentially, which may lead to multicollinearity. Additionally, the cause-effect relationship of a three or four variable interaction term is difficult to interpret. A methodology that reproduces different causal paths to an existing outcome could provide a remedy. A model of complex causality also underlines equifinality, the perception that a system of sets can lead to the same outcome from a different combination of initial conditions (Katz and Kahn 1978). Such complex interactions and systems cannot be reproduced with standard linear relationships and combinations. Consequently, configurational comparative analysis is based on set-theoretic relations rather than on correlations. Qualitative comparative analysis (QCA) relies on set-theoretic methods and thus allows for analyzing cases as combinations of conditions and attributes (Fiss 2007, Ragin 2008, Ragin and Fiss 2008). QCA reveals the logical interconnection of different conditions that all lead to a unique outcome condition. Thus, it is based on interactive models that take all possible configurations of conditions into account.

Previous studies have used QCA methodology to analyze multivariate data in the field of social sciences (Fiss 2007), cross-country comparisons (Woodside 2013, Kvist 2007, Pajunen 2008, Skaaning 2007, Vis, Woldendorp and Keman 2007), strategic management (Ganter and Hecker 2014, Stanko and Olleros 2013, Greckhamer, Misangyi, Elms and Lacey 2008, Kogut, MacDuffie and Ragin 2004) or macroeconomic growth (Boyer 2004, Kogut and Ragin 2006). We apply QCA to the dynamic movie market. Current motion-picture economic studies focus their analyses on cause-effect relationships, variable-oriented researches and the use of linear regression models. Unfortunately, the studies of media economics do not yield consistent results. The divergences of findings imply two major issues. First, the econometric specifications themselves are problematic, using endogenous variables like production budgets, critic ratings or star actors (Litman 1983, Smith and Smith 1986, Litman and Kohl 1989, Prag and Casavant 1994, Eliashberg and Shugan 1997, Sawhney and Eliashberg 1996, Ravid 1999, Nelson *et al*. 2001, Basuroy, Chatterjee and Ravid 2003, Liu 2006, McKenzie and Walls 2013, Nelson and Glotfelty 2012). Second, linear regression models are limited in analyzing the volatile and dynamic American feature film market (Walls 2005, McKenzie 2010). Following previous movie business studies, we use the same influential variables and analyze the U.S. cinema market and the correlation between financial effects (production budget), seasonality and time effects (release before or during a federal holiday), reputation effects (movie reviews from professionals and moviegoers, star popularity effects of actors, sequels, prequels or book adaptations and award wins), discrimination effects (Motion Picture Association of America, MPAA ratings), and domestic box office gross. Contrary to present U.S. cinema market analyses, configurational comparative analysis allows us to model the complexity of the motion-picture industry without having the statistical artifacts of multicollinearity or endogeneity.

We are able to reveal complex causal structures with QCA methodology. Unfortunately, we are not able to state any propositions about the net effects of the conditions to the outcome. A way to solve the problem can be found in the Shapley value solution. The Shapley value represents the expected marginal power contribution of a player i to a coalition j. By default, it evaluates the *a priori* power distribution among members of a legislature or committee system, for instance, a council or a parliament (Shapley and Shubik 1954). We adopt the same formal language of coalitional games for modeling the interaction, but instead of studying the marginal power contribution of players or parties to a coalition, we examine conditions and their marginal contribution to the outcome in question. Previous studies, mostly in the field of biomedical research, applied the Shapley value to find significant statistical influences of genes to disease onset (Moretti, Patrone and Bonassi 2006, Esteban and Wall 2009).

Our findings have several managerial and strategic implications. First, the use of QCA for analyzing dynamic markets provides both academics and practitioners a novel approach to structure and understand complex causal structures of volatile markets, such as the motion-picture industry. Second, the Shapley value power allocation appears particularly appropriate for the study of net effects, thus giving managers the opportunity to identify key drivers of product sales and optimizing managerial decisions.

2 Variable-Oriented versus Configurational Comparative Analysis

The ultimate goal of variable-oriented analysis is the deductive testing of previously developed theories. The deductive testing is driven by observing and analyzing cross-case relationships between variables in preferred large N populations. The cases and populations are predefined. The number of cases has a crucial impact on the estimation quality. Consequently, the better the sample size, the better the estimations' precision. The estimation quality is also influenced by the properness of the model specification and formulated hypotheses, which is an essential requirement for the significant testing of cause-effect relationships. Variable-oriented analysis focuses on the empirical testing of the influence of independent variables on one or more dependent variables. Both types of variables have to show a decent grade of variation. The higher the variation quality, the greater the model specification quality. Consequently, the key focus of variable-oriented research is the analysis of the relative importance of individual variables, but not the analysis of complex combinations and interactions of all different variables across cases.

In contrast, configurational comparative analysis defines and analyzes well-defined sets of cases through a simultaneous formation process of theoretical construction and empirical estimations. As the analysis proceeds, the sets of cases become more specific through concept formation and empirical development. Cases must show case homogeneity as well as case heterogeneity. Case homogeneity is needed because cases should reflect conditions that are sufficient for achieving the outcome of interest. Simultaneously, cases should reflect a certain diversity. The main theoretical objective of configurational comparative analysis is the evaluation, assessment and development of theoretical models. One important distinction from variable-oriented research is the focus on multiple causal paths instead of single paths. Each causal path represents a configuration of relevant causal conditions that lead to the outcome in question.

In conclusion, variable-oriented analysis seeks to explain why and which relevant independent variables have a statistically significant influence on the dependent variable; an outcome that varies across cases (Ragin 2000). The primary objective of comparative research is to explain how causal conditions and the combinations of conditions are linked to a specific outcome. In contrast to correlation analysis, the outcome in comparative studies does not vary significantly across cases. Consequently, comparative research approaches the concept of diversity, a concept that combines generality and complexity in data analysis and simultaneously treats cases as a fixed, homogenous population and focuses on the specificity and distinctiveness of individual cases (Ragin 2000).

3 Modeling Complex Causal Structures and Paths

3.1 Data

Following previous studies in motion-picture economics, we use a sample of the 500 all-time box office champions. The information was mainly collected from the Internet databases TheNumbers.com, Box Office Mojo and IMDb. In accordance with previous studies in motion-picture economics and for the purpose of study comparison, we concentrate our analysis on the influence of financial effects (production budget), seasonality and time effects (movie release before a federal holiday), reputation effects (movie reviews from critics and moviegoers, star popularity appeal of actors, sequels, prequels or book adaptations and award wins) and discrimination effects (restricted MPAA ratings). Movie business studies have shown a strong causal connection of these eight elements with box office gross. Consequently, we specify these seven variables as our core causal conditions. Unlike linear regression analysis, QCA is not based on the same data probability assumptions. In addition, causes from a lack of production data and, consequently, missing values are less of an issue, as QCA was developed for small-to-intermediate samples. We use a QCA that rests on a Boolean or binary matrix. Consequently, we have to transform all not genuine dichotomous variables into binary variables. In order to not completely lose the continuous nature of variables, a transformation threshold, well-founded in theory and preceding studies, had to be chosen.

3.2 Independent Conditions

The production budget reflects all the film's negative production costs, excluding compensations and marketing expenditures. Big budget blockbuster movies with high-cost star actors produced by major studios account for U.S. $100 million production costs on average. Art movies, specialty genre movies and other niche-market movies produced by conglomerates' indie subsidiaries or independent producer-distributors have an average U.S. $40 million production budget per release (McDonald and Wasko 2008). Most motion-picture industry studies show a significantly positive influence of production costs on sales revenues (Litman 1983, Ravid 1999, Basuroy, Chatterjee and Ravid 2003, Walls 2005).

Consumers consider the production budget a "proxy variable" for the overall quality of a film (Litman 1983). Consequently, the overall costs represent one of our main causal conditions. Following MacDonald and Wasko (2008), we choose a transformation threshold of U.S. $100 million to divide feature films into two categories – big budget blockbusters and independent movies.

Additionally, studies find a significant correlation between annual quarters and financial success (Litman 1983, Sochay 1994, Nelson *et al*. 2001, Einav 2007). We concentrate our analysis on federal holidays because on these days, all non-essential federal government and non-federal government offices are closed, including schools and stores, giving people the opportunity for recreation. Since the "holiday" condition already represents dichotomous characteristics, no further transformation is needed.

Consumers have a need for signals that help them to get over their uncertainties in the buying decision process. Reviews from critics and moviegoers may be one of these signals. The influence of movie critics on ticket sales has been shown in various studies (Litman 1983, Litman and Kohl 1989, Wallace *et al*. 1993, Sochay 1994, Eliashberg and Shugan 1997). We analyze the metascores from RottenTomatoes.com; a movie review website that publishes an aggregated percentage rating index that is created through a weighted summarization of top critics. Generally, Rotten Tomatoes divides its rated movies into two categories. If the positive reviews account for more than 60%, the movie is considered "fresh". If positive reviews make up less than 60%, the film is evaluated as "rotten". We transform our critic conditions according to the classification of Rotten Tomatoes. Consequently, if a film is "fresh", it is coded 1, otherwise 0.

Movies with star actors are shown on 20% more screens than movies without star actors (De Vany 2004). Thus, the star popularity appeal of actors cannot only alleviate viewers' insecurities, but also the uncertainties of distributors and cinema operators. Film studies confirm the influence of star actors on box office gross (Litman and Kohl 1989, Wallace *et al*. 1993, Ravid 1999, De Vany and Walls 1999, Elliot and Simmons 2008). Apparently, an actor's reputation appears to be a good signal for a movie's box office appeal. We use Quigley's Annual List of Box-Office Champions[1] to identify an actor as a star actor. Afterwards, we transform the continuous star actor variable into a binary condition that

[1] The Top Ten Money Making Stars Poll is a list of the previous year's top 10 moneymaking actors. The list is based on a survey of movie theater owners and published in Quigley's Annually International Motion Picture Almanac.

adopts 1 if the number of star actors appearing in a film is greater than or equal to 1.

Due to high information search costs, it can be rational for consumers to rely on past decisions. Sequels, prequels and book adaptations can base themselves on the reputation of their predecessors' success and positively influence consumer choices and, consequently, box office returns (Prag and Casavant 1994). We account for the explicit reputation effect of brand extension products by integrating a dichotomous condition in our analysis.

Apart from brand extension products, award ceremonies also reach the attention of consumers. Due to their product evaluation independency, awards can serve consumers as an additional signal of quality and thus positively influence ticket sales (Nelson *et al*. 2001). The Academy Awards of the Motion Picture Arts and Scienes, also known as the Oscars, present the prevalent award in the entertainment industry. Accordingly, we concentrate our analysis on Oscar wins and conduct the same variable transformation as for star actors.

R-rated motion pictures exclude consumers from consumption through their restricted age classification. Consequently, movie business studies show that movies with a MPAA rating G, PG and PG-13 are more profitable than R-rated movies (De Vany and Walls 2002).[2] Following the studies by Ravid (1999), we analyze the economic influence of MPAA ratings and differentiate movies into two categories – films with an R-rating and with films with no R-rating.

3.3 Outcome Condition

Following previous studies in motion-picture economics, the primary outcome condition of interest is movie performance, measured as box office gross (Eliashberg 1996, Ravid 1999, De Vany 2004, Walls 2005). Box office gross represents a continuous variable and must be transformed into a dichotomous condition before analysis. Consequently, we calculate the return on investment (ROI) and classify each movie into two categories. If the ROI exceeds 1 and more, the outcome condition adopts 1, indicating a successful movie investment. If the ROI is less than 1, a movie is considered unsuccessful and the outcome condition is coded 0.

A consolidated overview of the descriptive statistics and binary transformations of the

[2] The Motion Picture Association of America's film-rating system is used in the U.S. and its territories to rate a film's thematic and content suitability for certain audiences. Since 1990, the MPAA movie ratings are as follows: G - General Audiences (all ages admitted), PG - Parental Guidance Suggested (some material may not be suitable for children), PG-13 - Parents Strongly Cautioned (some material may be inappropriate for children under 13), R - Restricted (under 17 requires an accompanying parent or adult guardian), NC-17 - No One 17 and Under Admitted (for adult viewers only).

causal conditions and outcome measure is displayed in Table 1.

Table 1: Descriptive Statistics and Binary Transformation of Causal and Outcome Conditions

Causal Conditions	Obs	Mean	Std. Dev.	Transformation
Financial Effects				
Production Budget	428	89.490.452	59.492.367	$\begin{cases}1, if\ Budget_i \geq US\$100\ million \\ 0, if\ Budget_i < US\$100\ million\end{cases}$
Seasonality and Time Effects				
Federal Holiday Release	428	.46	.50	Already represents dichotomous variable.
Reputation Effects				
Critic Reviews	428	.65	.24	$\begin{cases}1, if\ Review\ Score_i \geq 60\% \\ 0, if\ Review\ Score_i < 60\%\end{cases}$
Moviegoers Reviews	428	.71	.14	$\begin{cases}1, if\ Review\ Score_i \geq 60\% \\ 0, if\ Review\ Score_i < 60\%\end{cases}$
Star Actors' Appeal	428	.35	.48	$\begin{cases}1,\ if\ No.\ of\ Stars_i \geq 1 \\ 0,\ if\ No.\ of\ Stars_i < 1\end{cases}$
Sequel / Prequel / Book Adaptation	428	.56	.50	Already represents dichotomous variable.
Award Wins	428	.50	1.38	$\begin{cases}1,\ if\ No.\ of\ Award\ Wins_i \geq 1 \\ 0,\ if\ No.\ of\ Award\ Wins_i < 1\end{cases}$
Discrimination Effects				
Restricted MPAA Rating	428	.21	.40	Already represents dichotomous variable.
Outcome Condition	**Obs**	**Mean**	**Std. Dev.**	**Transformation**
Box Office Gross	428	183.587.517	88.297.162	$\begin{cases}1, ROI_i > 1 \\ 0, ROI_i \leq 0\end{cases}$

3.4 Methodology

We use the QCA methodology, a set-theoretic approach that reveals the complex structure of how causal conditions contribute to a given outcome. This approach is uniquely suited for analyzing cause-effect relationships as it considers individual cases as configurational sets of attributes (Fiss 2007, Ragin 2008, Ragin and Fiss 2008). QCA rests on the assumption that causal configurations are based on a set-subset relationship. Thus, QCA allows us to delete attributes that are unrelated to the outcome measure. For example, if we are interested in analyzing what configurations lead to high-selling products, QCA considers only the "high-selling-set" and examines the combination of cases' attributes related with the associated outcome. By applying logical reduction algorithms, QCA minimizes the number of causal conditions into a reduced "subset" of configurations that lead to the outcome measure.

To empirically identify minimized causal configurational paths, QCA is hierarchically structured. After the transformation of causal and outcome measures into a Boolean matrix, we are able to build a "truth table". Truth tables form the central element in QCA. Each row of the truth table represents one of the 2^n possible configurations of n variables that models sufficient conditions and various configurational sets of attributes correlated with the outcome in question. A condition is defined as sufficient if it can lead to the outcome of interest by itself. Consequently, as we include eight independent measures, our truth table consists of 256 logical combinations of attributes. Nonetheless, some rows contain many cases, some rows several and some rows even include no cases, meaning that there is no empirical evidence for the existence of the particular combinations of attributes related with the row.

For further analysis, only rows are considered that achieve a minimum level of consistency of a solution. Consistency describes what percentage of cases are consistent with the solution formula, i.e., what percentage of cases are correctly described by that particular solution formula. The consistency is calculated from the proportion of all causal condition X with the outcome Y in relation to all cases X. For any outcome Y and N number of cases, the consistency of a causal condition or of a combination of conditions X is defined as

$$Consistency = \sum_{i=1}^{N} \min(X_i, Y_i) \Big/ \sum_{i=1}^{N} X_i \tag{1}$$

Consequently, consistency tells us *how well we explain*. Ragin (2008) recommends a minimum consistency threshold of 0.75.

In order to simplify the information comprised in the truth table, we implement a logical minimization based on the truth table algorithm. The results of the process of logical minimization or summarizing are causal combinations that are sufficient for the outcome in question and are expressed in the form of a solution formula. In QCA, a solution formula reflects the relationship between an outcome in question with the causal conditions and the help of Boolean operators. The three basic Boolean operators are logical OR (+), logical AND (*) and NOT (the negation of causally relevant conditions is normally presented with small initial letters as opposed to large initial letters for positive causally relevant conditions). These operators are able to express any possible conjunction of dichotomous conditions with the binary outcome. For instance, given an outcome set Y and predictor sets A and B, QCA examines which combinations of A and B (i.e., A · B, A · b, a · B, a · b) are most likely to produce Y.

3.5 Estimation Results

Table 2 illustrates the results of the QCA and a parsimonious solution of core causal conditions that lead to the outcome in question. The parsimonious solution identifies only conditions that provide strong empirical evidence relative to the outcome. Any peripheral conditions are neglected because they only provide marginal explanatory power and, consequently, give little insight into relevant causal conditions.

Following the solution table notation of Ragin and Fiss (2008), we use black circles to indicate the presence of a condition, and white circles to indicate its absence. Blank spaces indicate either the absence or the presence of a causal condition and thus are referred to as "don't care" situations. We set the minimum consistency threshold to 0.80, which is above the minimum consistency level of 0.75 recommended by Ragin (2008) and, consequently, even more conservative and restrictive. The minimum frequency cutoff is set at four, meaning that the minimum number of cases required for a solution is four. In total, 313 causal configurations exceed the minimum solution frequency. In addition, 99 cases also pass the minimum consistency threshold of 0.80.

The QCA solution table shows six core causal configurations. Two solutions do not exceed the unique consistency threshold of 0.80. Additionally, the solution table illustrates an overall solution coverage and raw configuration coverages. The coverage describes how many individual cases X that stand in relation to the total number of cases Y are explained by a solution formula. The coverage thus examines what ratio of cases is actually declared by the solution formula. For any outcome Y and N number of cases, the coverage of a sufficient causal condition or of a combination of conditions X is defined as

$$Coverage = \sum_{i=1}^{N} \min(X_i, Y_i) \Big/ \sum_{i=1}^{N} Y_i \tag{2}$$

Consequently, the coverage tells us *how much of a phenomenon we explain*. Our comparative analysis of core causal conditions shows an acceptable overall solution coverage of 0.55. Solutions 1a and 1b have the highest raw coverage with 0.23 and 0.15, respectively.

Solutions 2a, 2b, 3a and 3b all present a raw coverage lower than 0.10. Consequently, the solution coverage of 1a and 1b is almost twice as large as the solution coverage of the other four solutions.

Both solutions 1a and 1b indicate that releasing medium and low budget movies during federal holidays is essential for high performance. The interaction of financial and time effects is particularly strong if they are combined with professional critic reviews and non-restricted movie ratings (solution 1a) or the non-appearance of star actors and brand extension products (solution 1b).

Solution 1a indicates that movie studios should select families with young children and teenagers as their target group. As parents have to protect their children against sexual and violent content in movies, it is likely that they will inform themselves about the content before they go to the cinema. Additionally, as R-rated motion pictures indicate adult material, it is unlikely that families will go to R-rated movies. Comedy films use humor as a driving force and thus have no need for expensive special effects or star actors. The aim of a comedy movie is to tell an entertaining story with vivid characters. Consequently, solution 1a gives evidence that the production and distribution of comedy movies during holiday seasons is sufficient for achieving high box office performance. This finding gives support to Prag and Casavant (1994), who show that the comedy genre has a significantly higher marketability compared to other movie genres.

Solution 1b indicates a second important path to high performance, combining medium and low cost productions with federal holiday releases, the appearance of unknown actors and a sequel, prequel or book adaptation, which is consistent with a brand extension strategy. Consequently, configuration 1b confirms the importance of a brand extension marketing strategy for the introduction of new products suggesting that consumers rely heavily on brand extension products (Tauber 1988). Following the brand extension literature, movie managers should consider that a successful brand extension mainly consists of two factors: a high product feature similarity and brand concept consistency (Park *et al*. 1991, Broniarczyk and Alba 1994, Völckner and Sattler 2006). If movie studios manage to consider both factors in new movie projects and releases, brand extension motion pictures provide a competitive advantage in their existing market but also have the potential to conquer new markets.

Solutions 2a, 2b, 3a and 3b indicate a substitution effect of signals. Solutions 2a and 2b have low and medium budget movies and non-federal holiday releases in common. While configuration 2a signals a reputation effect of star actor appeal and brand extension products, configuration 2b indicates a strong signaling effect of award wins. Solutions 3a and 3b provide a similar picture as solutions 2a and 2b. Although production budget is not part of both configurations, solution 3a indicates that a non-federal holiday release combining actors without star appeal and R-rated movies is sufficient for achieving high performance. Consequently, this solution suggests that it is sufficient to produce motion pictures that include hard language, intense or persistent violence, and sexually-oriented nudity to attract a significant number of attendants. Solution 3b indicates that the concurrent presence of a non-federal holiday release, an original film screenplay and the win of one or more Academy Awards are sufficient for achieving high rental movies. These indicators refer to the art house film genre, motion pictures that are often experimental and not designed for commercial profit. Art house movies must draw on less-known actors and smaller special effects, due to their tight production budget and missing support of major studios. Mainly because of their artistic and experimental film style, art movies are often considered for Academy Awards. This finding is in line with Nelson *et al*. (2001), who show the positive economic effects of Oscar wins on a movie's cinematic performance. Nevertheless, it should be noted once more that solutions 2a, 2b, 3a and 3b all possess a raw coverage lower 10%. Consequently, their empirical predication should be further verified.

Table 2: Solution Table of Core Causal Configurations for Achieving High Movie Performance

	Solution					
Configuration	**1 a**	**1 b**	**2 a**	**2 b**	**3 a**	**3 b**
Financial Effects						
Production Budget	○	○	○	○		
Seasonality and Time Effects						
Federal Holiday Release	●	●	○	○	○	○
Reputation Effects						
Critic Reviews	●					
Moviegoers Reviews						
Star Actors' Appeal		○	●		○	
Sequel / Prequel / Book Adaptation		●	●			○
Award Wins				●		●
Discrimination Effects						
Restricted MPAA Rating	○				●	
Unique Consistency	0.91	0.89	0.83	0.91	0.66	0.78
Raw Coverage	0.23	0.15	0.08	0.09	0.09	0.06
Overall Solution Consistency			0.83			
Overall Solution Coverage			0.55			

● circles indicate the presence of a condition, and ○ circles indicate its absence. Blank spaces indicate "don't care".

4 Marginal Contribution of Conditions

4.1 Theoretical Background of the Shapley Value

We identify configurational sets of causal conditions that lead to the outcome of interest by using QCA. Nevertheless, we are so far not able to determine the net effects or particular marginal contribution of each causal condition to the outcome measure. The Shapley value power distribution remedies the shortcomings of QCA net effect calculation. Shapley and Shubik (1954) introduced the concept of power indices for the first time. They were able to establish a method for the *a priori* evaluation of the distribution of power among parties and members of a committee system.

The calculation of the Shapley value is based on the structure of simple games. A game v is called a simple game if it allocates every coalition a value of either 0 or 1, consequential $v: 2^N \rightarrow \{0,1\}$. Thus, the Shapley value and QCA are based on the same binary data structure.

Simple games have been widely applied to the analysis of the distribution power allocation in coalitions, for example in councils or parliaments. We adopt the same formal language of coalitional games for modeling the interaction among causal conditions. A coalition that is correlated with 1 is called a winning coalition. Otherwise, it is defined as a losing coalition. The set of winning coalitions in the simple game v is identified as $W(v) \coloneqq \{S \in 2^N | v(S) = 1\}$. To describe a simple game, it is consequently sufficient to list the total set of the winning coalitions. The minimal winning coalitions are those winning coalitions that cannot be reduced further without losing their status as "winners", i.e., $MW(v) \coloneqq \{S \in W(v) | v(S') = 0 \ \ (S' \subsetneq S)\}$.

The Shapley value $\phi_i(v)$ assigns each simple game v a vector $\phi(v) = (\phi_1(v), \dots, \phi_n(v)) \epsilon \mathbb{R}_+^n$. The Shapley value $\phi_i(v)$ of player i in game v is a weighted sum of terms of the form $[(v(S) - v(S\backslash\{i\})]$ and is defined as

$$\phi_i(v) = SSI_i(v) \coloneqq \sum_{S \epsilon W(v): i \epsilon S} \frac{(n-s)!\,(s-1)!}{n!} (v(S) - v(S\backslash\{i\}) \qquad (i \epsilon N) \tag{3}$$

Thus, the Shapley value $\phi_i(v)$ for player i depends on the fulfilling "role" of a player i, i.e., how much he contributes to a certain coalition. The Shapley value $\phi_i(v)$ is the expected marginal contribution of player i to a randomly selected coalition. In determining the marginal power contribution of player i, the player's ranking order of a coalition is not important. The result of the Shapley value $\phi_i(v)$ is a solution concept that allocates each game a unique allocation. If two players are symmetric, i.e., if they fulfill the same "role", they should be assigned the same Shapley value $\phi_i(v)$. Additionally, the sum of individual Shapley values of v and w equals the Shapley value of the sum game. If a player does not contribute anything to a coalition he should receive a zero share.

4.2 Estimation of the Conditions' Shapley Value

In addition to the analysis of interactions and complex cause-effect relationships, we identify the net effects of the causal conditions to the outcome in question. Table 3 shows the results of the Shapley value calculation and the marginal contribution of the eight causal conditions: production budget, movie release before a federal holiday, movie reviews from critics and moviegoers, star popularity appeal of actors, sequels, prequels or book adaptations, award wins and restricted MPAA ratings.

The Shapley value indicates that the production budget has the largest marginal contribution to high performance, with 80%. Reviews from professional critics and reputation effects from previous products in the form of sequels, prequels or book adaptations have the second and third biggest net effect on the outcome of interest. An approximately total marginal contribution of 12% is allocated to these two conditions. These two reputation conditions are closely followed by user reviews, which generate a marginal contribution of 4.15%. The condition “Star Actors’ Appeal” have a net effect of 2.47%. Thus, the Shapley value indicates that star actors contribute sparsely towards achieving high performance. The three conditions Restricted-rated movies, Oscar wins, and a movie’s release during a federal holiday have a combined net effect below 1.5% on the outcome. Consequently, their causal influence on a movie’s performance can be neglected.

The estimation of the Shapley value lends support to the QCA solutions 1a and 1b, which also focus on the conditions of production budget, critic reviews and reputation effects from previous products. Consequently, the Shapley value power allocation supports a segmentation and brand extension strategy for achieving high performance at the box office. Additionally, critic reviews significantly contribute to high performance. This finding reinforces Eliashberg and Shugan (1997), who argue that critics adopt the role of an influencer who possesses a credible expertise and has the reputation of being an opinion leader. Consequently, reviews of professional critics have the ability to influence box office ticket sales. Our empirical findings also suggest that critics have an influencing effect which outweighs the peer group effect and reputation of user reviews. Consequently, this result stands in contrast to previous studies stating that critic reviews influence consumer choices *ex ante*, whereas user reviews or word-of-mouth influence buying decisions *ex post* (Dellarocas *et al*. 2007, Chintagunta *et al*. 2010, Archak *et al*. 2011).

Table 3: Shapley Value Decomposition of Causal Conditions

Causal Conditions	Shapley Value	Percentage
Financial Effects		
Production Budget	34.003	80.07
Seasonality and Time Effects		
Federal Holiday Release	.027962	0.07
Reputation Effects		
Critic Reviews	2.7275	6.42
Moviegoers Reviews	1.7636	4.15
Star Actors' Appeal	1.0501	2.47
Sequel / Prequel / Book Adaptation	2.349	5.53
Award Wins	.079267	0.19
Discrimination Effects		
Restricted MPAA Rating	.46844	1.10
Total	42.469	100

5 Discussion and Implications

The introduction of the conditional stable-distribution regression analysis by Walls (2005) and the study of the Australian DVD industry using Pareto distribution models by McKenzie (2010) mark a fundamental step in the analysis of volatile industries. We present a complementary study and introduce a new methodology of analyzing cause-effect relationships and outcome performance in dynamic markets. The set-theoretic approach allows controlling for elements of a configuration that are relevant for an outcome of interest but, moreover, identifying sufficient combinations of elements and their relationship with the outcome. The novel approach and their findings connote direct implications for the management and marketing literature, but also for movie business practitioners. Most current research implies a linear cause-effect relationship, leading to a possible mismatch between the theoretical construct and the management reality.

QCA is appropriate to analyze complex cause-effect relationships and multiple interactions. Consequently, QCA appears to be useful to identify complementary and substitutionary key success factors. Accordingly, configurational comparative analysis represents a complementary analysis method to existing standard approaches, such as cluster analysis or interactions in multiple regressions. Nevertheless, QCA also has limitations. Due to the fact that QCA is based on complex interactions that consider all possible configurations of n variables, its Boolean matrix increases exponentially with the number of causal conditions. Consequently, the appropriate ratio of causal conditions to the number of cases is limited and must be considered by researchers. Additionally, QCA is inappropriate for the study of net effects of causal conditions.

We introduce a novel methodology to specify marginal contributions of causal conditions to the outcome of interest. The Shapley value evaluates the distribution of power among conditions and appears particularly useful for determining the marginal power contribution of individual conditions.

The novel approach can be adapted for research in other industry contexts in which researchers are confronted with complexity and multiple interactions of key success factors. Consequently, QCA and the Shapley value appear to be generalizable to industries with dynamic market structures. Additionally, our findings not only support the existing literature on brand extension, segmentation strategy and signaling effects of critics but also enables managers to estimate efficiently the market value of different marketing strategies (Tauber 1988, Broniarczyk and Alba 1994, Völckner and Sattler 2006, Eliashberg and Shugan 1997, Dellarocas *et al*. 2007, Archak *et al*. 2011). For stakeholders, the novel methodologies provide transparency and complementary evidence to the prevalent anecdotal evidence that informs many industries.

In summary, we introduce two novel approaches to analyze complex causal relations in dynamic industries. Using example data from the motion-picture industry, we show that a segmentation and a brand extension strategy are sufficient for achieving high market performance. In addition, the conditions of production budget, critic reviews and brand extension products particularly appear to drive high performance in volatile markets and, consequently, represent a greater source for competitive advantage than any other influencing factors.

References

Archak, N.; Ghose, A.; Ipeirotis, P.G. 2011: “Deriving the pricing power of product features by mining consumer reviews”. *Management Science*, 57(8), 1485-1509.

Basuroy, S.; Chatterjee, S.; Ravid, S. A. 2003: “How critical are critical reviews? The box office effects of film critics, star power and budgets”. *Journal of Marketing*, 67(4), 103-117.

Broniarczyk, S. M.; Alba, J. W. 1994: "The Importance of the Brand in Brand Extension". *Journal of Marketing Research*, 31(2), Special Issue on Brand Management, 214-228.

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

Dellarocas, C.; Zhang, X.; Awad, N.F. 2007: “Exploring the value of online product reviews in forecasting sales: The case of motion pictures”. *Journal of Interactive Marketing*, 21(4), 23-45.

De Vany, A.; Walls, W. D. 1999: “Uncertainty in the Movies. Can Star Power reduce the Terror of the Box-Office?”. *Journal of Cultural Economics*, 23(4), 285-318.

De Vany, A.; Walls, W. 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.

De Vany, A. 2004: “Hollywood Economics. How extreme uncertainty shapes the film industry”. New York.

Einav, L. 2007: “Seasonality in the U.S. Motion Picture Industry”. *The RAND Journal of Economics*, 38(1), 127-145.

Eliashberg, J.; Shugan, S. M. 1997: “Film Critics. Influencers or Predictors?”. *Journal of Marketing*, 61(2), 68-78.

Esteban, F. J.; Wall, D. P. 2009: “Using game theory to detect genes involved in Autism Spectrum Disorder”. *TOP*, 19(1), 121-129.

Fiss, P. C. 2007: “A set-theoretic approach to organizational configurations”. *Academy of Management Review*, 32(4), 1180-1198.

Ganter, A.; Hecker, A. 2013: "Configurational paths to organizational innovation: qualitative comparative analyses of antecedents and contingencies." *Journal of Business Research*, In Press.

Greckhamer, T., Misangyi, V. F., Elms, H., & Lacey, R. 2008: “Using qualitative comparative analysis in strategic management research: An examination of combinations of industry, corporate, and business-unit effects". *Organizational Research Methods*, 11(4), 695-726.

Katz, D., & Kahn, R. L. 1978: "The social psychology of organizations". 2nd ed. New York: Wiley.

Kogut, B., MacDuffie, J. P, & Ragin, C. C. 2004: "Prototypes and strategy: Assigning causal credit using fuzzy sets". *European Management Review*, 1(2), 114-131.

Kogut, B., & Ragin, C. 2006: "Exploring complexity when diversity is limited: Institutional complementarity in theories of rule of law and national systems revisited". *European Management Review*, 3(1), 44-59.

Kvist, J. 2007: "Fuzzy set ideal type analysis". *Journal of Business Research*, 60(5), 474-481.

Litman, B. R. 1983: "Predicting Success of Theatrical Movies. An Empirical Study". *Journal of Popular Culture*, 16(4), 159-175.

Litman, B. R.; Kohl, L. S. 1989: "Predicting Financial Success of Motion Pictures. The 80's Experience". *Journal of Media Economics*, 2(2), 35-50.

Liu, Y. 2006: "Word-of-Mouth for Movies: Its Dynamics and Impact on Box Office Revenue". *Journal of Marketing*, 70(3), 74-89.

McDonald, P.; Wasko, J. 2008: "The contemporary Hollywood film industry". Malden: Blackwell.

McKenzie, J. 2010: "How do theatrical box office revenues affect DVD retail sales? Australian empirical evidence". *Journal of Cultural Economics*, 34(3), 159-179.

McKenzie, J.; Walls, W. D. 2013: "Australian films at the Australian box office: performance, distribution, and subsidies". *Journal of Cultural Economics*, 37(2), 247-269.

Moretti, S.; Patrone, F.; Bonassi, S. 2006: "The class of microarray games and the relevance index for genes". *TOP*, 15(2), 256-280.

Nelson, R. A.; Donihue, M. R.; Waldman, D. M.; Wheaton, C. 2001: "What´s an Oscar worth?" *Economic Inquiry*, 39(1), 1-16.

Nelson, R. A.; Glotfelty, R. 2012: "Movie stars and box office revenues: an empirical analysis". *Journal of Cultural Economics*, 36(2), 141-166.

Pajunen, K. 2008: "Institutions and inflows of foreign direct investment: A fuzzy-set analysis". *Journal of International Business Studies*, 39, 652–669.

Park, C. W.; Milberg, S.; Lawson, R. 1991: "Evaluation of Brand Extensions: The Role of Product Feature Similarity and Brand Concept Consistency". *Journal of Consumer*

Research, 18(2), 185-193.

Prag, J.; Casavant, J. 1994: "An Empirical Study of the Determinants of Revenues and Marketing Expenditures in the Motion Picture Industry". Journal Cultural Economics, 18(3), 217-235.

Ragin, C. C. 1987: "The Comparative Method. Moving beyond qualitative and quantitative strategies". Berkeley/Los Angeles/London: University of California Press.

Ragin, C. C. 2000: "Fuzzy-Set Social Science". Chicago/London: University of Chicago Press.

Ragin, C.C. 2008: "Redesigning social inquiry: Fuzzy sets and beyond". Chicago: University of Chicago Press.

Ragin, C. C., & Fiss, P.C. 2008: "Net effects analysis versus configurational analysis: An empirical demonstration". C. C. Ragin (Ed.): "Redesigning social inquiry: Fuzzy sets and beyond". 190-212. Chicago: University of Chicago Press.

Ravid, A. S. 1999: "Information, Blockbusters and Stars. A Study of the Film Industry". *Journal of Business*, 72(4), 463-492.

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

Shapley, L. S.; Shubik, M. 1954: "A method for evaluating the distribution of power in a committee system". *The American Political Science Review*, 48(3), 787-792.

Smith, S. P.; Smith, V. K. 1986: "Successful Movies. A Preliminary Empirical Analysis". *Applied Economics*, 18(5), 501-507.

Sochay, S. 1994: "Predicting the Performance of Motion Pictures". *Journal of Media Economics*, 7(4), 1-20.

Stanko, M. A.; Olleros, X. 2013: "Industry growth and the knowledge spillover regime: Does outsourcing harm innovativeness but help profit?" *Journal of Business Research*, 66(10), 2007-2016.

Tauber, E. M. 1988: "Brand Leverage: Strategy for Growth in a Cost Controlled World". *Journal of Advertising Research*, 28(4), 26-30.

Vis, B.; Woldendorp, J.; Keman, H. 2007: "Do miracles exist? Analyzing economic performance comparatively". *Journal of Business Research*, 60(5), 531-538.

Völckner, F.; Sattler, H. 2006: "Drivers of Brand Extension Success". *Journal of Marketing*, 70(2), 18-34.

Wallace, T. W.; Steigerman, A.; Holbrook, M. B. 1993: "The Role of Actors and Actresses in the Success of Films. How much is a movie Star worth?”. *Journal of Cultural Economics*, 17(1), 1-27.

Walls, W. D. 2005: “Modelling Movie Success When “Nobody knows anything”: Conditional Stable-Distribution Analysis of Film Returns”. *Journal of Cultural Economics*, 29(3), 177-190.

Woodside, A. G. 2013: "Moving beyond multiple regression analysis to algorithms: Calling for adoption of a paradigm shift from symmetric to asymmetric thinking in data analysis and crafting theory." *Journal of Business Research*, 66 (4), 463-472.