



Management Science

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

Estimating Structural Models of Equilibrium and Cognitive Hierarchy Thinking in the Field: The Case of Withheld Movie Critic Reviews

Alexander L. Brown, Colin F. Camerer, Dan Lovallo,

To cite this article:

Alexander L. Brown, Colin F. Camerer, Dan Lovallo, (2013) Estimating Structural Models of Equilibrium and Cognitive Hierarchy Thinking in the Field: The Case of Withheld Movie Critic Reviews. *Management Science* 59(3):733-747. <http://dx.doi.org/10.1287/mnsc.1120.1563>

Full terms and conditions of use: <http://pubsonline.informs.org/page/terms-and-conditions>

This article may be used only for the purposes of research, teaching, and/or private study. Commercial use or systematic downloading (by robots or other automatic processes) is prohibited without explicit Publisher approval. For more information, contact permissions@informs.org.

The Publisher does not warrant or guarantee the article's accuracy, completeness, merchantability, fitness for a particular purpose, or non-infringement. Descriptions of, or references to, products or publications, or inclusion of an advertisement in this article, neither constitutes nor implies a guarantee, endorsement, or support of claims made of that product, publication, or service.

Copyright © 2013, INFORMS

Please scroll down for article—it is on subsequent pages



INFORMS is the largest professional society in the world for professionals in the fields of operations research, management science, and analytics.

For more information on INFORMS, its publications, membership, or meetings visit <http://www.informs.org>

Estimating Structural Models of Equilibrium and Cognitive Hierarchy Thinking in the Field: The Case of Withheld Movie Critic Reviews

Alexander L. Brown

Department of Economics, Texas A&M University, College Station, Texas 77843,
abrown@econmail.tamu.edu

Colin F. Camerer

Division for the Humanities and Social Sciences, California Institute of Technology, Pasadena, California 91125,
camerer@hss.caltech.edu

Dan Lovo

The University of Sydney Business School, The University of Sydney, Sydney NSW 2006, Australia,
dan.lovallo@sydney.edu.au

Film studios occasionally withhold movies from critics before their release. Because the unreviewed movies tend to be below average in quality, this practice provides a useful setting in which to test models of limited strategic thinking: Do moviegoers seem to realize that no review is a sign of low quality? A companion paper showed that in a set of all widely released movies in 2000–2009, cold opening produces a significant 20%–30% increase in domestic box office revenue, which is consistent with moviegoers overestimating quality of unreviewed movies (perhaps due to limited strategic thinking). This paper reviews those findings and provides two models to analyze this data: an equilibrium model and a behavioral cognitive hierarchy model that allows for differing levels of strategic thinking between moviegoers and movie studios. The behavioral model fits the data better, because moviegoer parameters are relatively close to those observed in experimental subjects. These results suggest that limited strategic thinking rather than equilibrium reasoning may be a better explanation for naïve moviegoer behavior.

Key words: decision analysis; game theory; economics; econometrics; marketing; competitive strategy; bounded rationality; psychology

History: Received July 26, 2010; accepted March 6, 2012, by Brad Barber, behavioral economics. Published online in *Articles in Advance* September 4, 2012.

1. Introduction

Game theory has sometimes been criticized as a descriptive model of business practice, or a source of normative advice, on the grounds that most analysis assumes people forecast accurately what others will do, and choose best responses given their accurate (equilibrium) forecasts. Recently, models have been developed that allow plausible limits on strategic thinking. These models are particularly useful because their basic principles can apply to many different games. One class of models that has been applied to many data sets is a “cognitive hierarchy” (CH) model of levels of steps of thinking (and its close relative, level- k). These models have been used to explain normal form games in a wide variety of experimental¹ and field settings,² but the only applications of these

theories to games with private information so far are analyses of auctions.³ This paper explores the generality of these approaches through the first field application of models of limited strategic thinking to games with private information.⁴

The setting we study is Hollywood movies. Movie studios generally show movies to critics well in advance of the release (so that critics’ reviews can be published or posted before the movie is shown and can be quoted in newspaper ads). However, movies are often deliberately made unavailable to film critics

³ See Crawford and Iriberri (2007b) and Wang (2006).

⁴ This setting is one example of a more general class of disclosure games in which a seller who knows something about a product’s quality can choose whether or not to disclose a signal of its quality (for surveys, see Verrecchia 2001, §3; Fishman and Hagerty 2003). These disclosure games have been extensively studied in economics (see Dranove and Jin 2010 for an exhaustive listing), but this paper is the first to tie the process of disclosure to models of limited strategic thinking and estimate them structurally. This paper examines the disclosure process, testing strategic disclosure as a response of producers to the limited strategic thinking of consumers.

¹ See Nagel (1995), Stahl and Wilson (1995), Camerer et al. (2004), Crawford and Iriberri (2007a), and Crawford et al. (2010).

² Goldfarb and Yang (2009) apply these models to firm adoption of 56 K modems, Goldfarb and Xiao (2011) study strategic entry of phone companies into new markets, and Östling et al. (2011) use Swedish lottery choices and experimental analogues.

in advance of their initial release, a practice sometimes called “cold opening.” If moviegoers believe that studios know their movie’s quality (and if some other simplifying assumptions hold; see Brown et al. 2012), then rational moviegoers should infer that cold-opened movies are below average in quality.

Anticipating this accurate negative inference by moviegoers, studios should only cold open the very worst movies. However, this conclusion requires many steps of iterated reasoning (and many simplifying assumptions). So it is an empirical question whether the equilibrium prediction fits behavior. If it does not fit well, it is also an empirical question whether neoclassical explanations can explain the data or whether models of limited strategic thinking, initially designed to explain experimental data as well or better than neoclassical models, fit the studios’ cold-opening decisions.

A fully rational analysis of simple disclosure games, due originally to Grossman (1981) and Milgrom (1981), implies that cold opening should not be profitable if some simple assumptions are met. The argument can be illustrated numerically with a simple example. Suppose movie quality is uniformly distributed from 0 to 100, moviegoers and studios agree on quality, and firm profits increase in quality. If studios cold open all movies with quality below a cutoff of 50, moviegoers with rational expectations will infer that the expected quality of a cold-opened movie is 25. But then it would pay to screen all movies with qualities between 26 and 100, and only cold open movies with qualities of 25 or below. More generally, if the studios do not screen movies with qualities below q^* , the consumers’ conditional expectation if a movie is unscreened is $q^*/2$, so it pays to screen movies with qualities $q \in (q^*/2, 100]$ rather than just $q \in (q^*, 100]$. The logical conclusion of iterating this reasoning is that only the worst movies (quality 0) are unscreened. This conclusion is sometimes called “unravelling.”⁵ We proceed with the maintained hypothesis that complete unravelling should occur in theory, if studios and consumers are perfectly rational.

The CH models also proceed through the steps of strategic thinking in the rational unravelling argument, except that they assume that some fraction of moviegoers end their inference process after a small number of steps. For example, a 0-level moviegoer thinks that cold-opening decisions are random (they convey no information about quality) and hence infers that the quality of a cold-opened movie is

average. A 1-level studio anticipates that moviegoers think this way and therefore opens all below-average movies cold, and shows all above-average movies to critics. Higher-level thinkers iterate more steps in this process. Observed behavior will then be an average of the predicted behaviors at each of these levels weighted by the fraction of moviegoers and studios who do various numbers of steps of thinking. (More details of this model are given in §3.)

The data generally do not agree with the standard full disclosure model. Roughly 10% of the movies in our sample are opened cold (though that fraction has increased sharply in recent years). Regressions show that cold opening appears to generate a box office premium (compared to similar-quality movies that are prereviewed, and including many other controls). Because box office returns are strongly correlated with subjective quality (measured by either critic or fan ratings), the cold-opening premium suggests that fans think the movie is better than it actually is. We also conclude that this explanation is consistent with four of five stylized facts in this environment, none of which can be explained by a neoclassical model.⁶

We then fit a baseline Nash equilibrium model, similar to Seim (2006), in which studios cold open movies when they receive a private idiosyncratic error, and a model using CH, that can be augmented to allow disequilibrium using two separate CH parameters for moviegoers and studios.⁷ Both baseline and CH models have roughly similar estimates for studio choices, because cold openings are quite rare as predicted by equilibrium and CH models with high levels of thinking (there is less data with actual cold-opening box office to fit the model). However, the baseline model cannot predict cold-opening premiums from moviegoer choice whereas the CH model can.

The estimates for moviegoers thinking in the CH model, especially in the period 2000–2005, are roughly consistent with experimentally observed data. Studios in the later period (2006–2009) also have lower estimates of perceived steps of thinking, suggesting they may be learning to best respond to moviegoers’ limited rationality.

The paper is organized as follows: Section 2 discusses data on quality ratings, box office returns, and

⁶ For example, an important observation is that fan ratings of quality are correlated with critic ratings but are systematically lower for cold-opened movies. This is consistent with the hypothesis that fans choose movies based on expected quality and are disappointed more often in cold-opened movies (presumably because their expectations were too high).

⁷ The mismatch between the degree of strategic thinking of moviegoers and studios is not typically observed in experimental data. However, keep in mind that experiments rarely use mixtures of populations that are more or less strategically sophisticated, so it is perhaps not surprising that the estimate of studio strategic thinking is very high and is much higher than the moviegoer estimates.

⁵ In the most similar theoretical work, Fishman and Hagerty (2003) do provide a model of a disclosure process with both informed and uninformed consumers, although it does not specifically address limited strategic thinking.

control variables, and presents some regression results on the existence of a box office premium for movies that are cold opened. Section 3 describes the Bayesian-Nash and CH models. Section 4 estimates parameters of those models based on studios' decisions and the box office revenue. Section 5 concludes and discusses future extensions to management-related research.

2. Data

Much of the data and details of regression conclusions are reported in a companion paper (Brown et al. 2012); we will summarize those results that are relevant to the analysis of this paper. In much of the analysis we will make a distinction between movies from 2000–2005 and 2006–2009, a distinction that was not made in the companion paper.

The data set is all 1,414 movies widely released in more than 600 theaters in the United States in their first weekend, over the decade from January 1, 2000, to December 31, 2009. Critic and moviegoer ratings are both used to measure quality. Metacritic.com normalizes and averages ratings from over 30 movie critics from newspapers, magazines, and websites. These have a roughly normal distribution between 0 and 100. For estimation purposes, we use the percentile score of those ratings so that they are standardized to a uniform distribution between 0 and 100. The metacritic rating is available for all non-cold-opened movies on the day they are released and is available on Monday for cold-opened movies. Because these ratings occur so early in a film's release, we assume the ratings help determine box office revenue, and not vice versa (i.e., critics are not influenced by box office). Other variables (such as cold opening, box office revenues, movie genres and ratings, production budgets, and star power ratings) are collected from various data sources (see the Web appendix, available at http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2126774).

Table 1 provides summary statistics for all variables. All these variables were used in a regression model to test if movies that are cold opened have significantly greater opening weekend and total U.S. box office revenues. The table also shows separate variable means for the cold-opened movies. The cold-opened movies are somewhat statistically different in a few dimensions—they tend to be smaller in budget and theater coverage, have less well-known stars, and overrepresent some genres (e.g., suspense/horror).

Each movie, j , has a standardized metacritic.com rating, q_j , a dummy variable for whether a movie was cold opened, c_j ($=1$ if cold), and a vector X_j of other variables. The model is

$$\log y_j = aX_j + bq_j + dc_j + \epsilon_j, \quad (1)$$

where y_j is opening weekend or total U.S. box office for movie j in 2005 dollars, standardized using the consumer price index (<http://www.bls.gov>). Table 2 shows regression results on logged total box office revenue and logged opening weekend revenue, respectively.

The point of these initial regressions is not to estimate a full model with endogenous studio decisions (we will estimate such a model in §4). Instead, the regression is simply a way of determining whether there is a difference in the revenue between cold-opened and reviewed movies. Under the standard equilibrium assumption that all quality information of cold-opened movies is inferred by logical inference of moviegoers, we should see no difference in revenues, and the cold-opening coefficient should be zero. If this is the case, it is good evidence for the Nash unravelling argument, and there is no interesting pattern for the behavioral theories to explain. The cold-opening coefficients in the first row of Table 2 show that cold opening a movie is positively correlated with the logarithm of opening weekend and total U.S. box office revenues.⁸

In this paper the analysis will often be separated over the years 2000–2005 and 2006–2009. This is because the frequency of cold opening doubled in 2006 and persisted at the same level through 2009. (In financial economics it is common to break a long period into subperiods to test for robustness of effects, and we adapt that method here as well.) There is no structural change in the movie industry in 2006 that justifies separating the entire sample into these two periods, but there is a statistical jump in the percentage of cold openings in that year. In our first regressions, it is apparent that the cold-opening premium is greater and more significant over the later years 2006–2009 than the earlier years 2000–2005. However, this result could be due to the small number of cold openings in the earlier period (43 of 778 movies, about 6%) compared to the later period (93 of 525 movies, about 18%).

If the regression model is taken literally, these coefficients suggest that cold opening a movie increases its revenue from 6% to 35%.⁹ However, we caution the reader in such an immediate interpretation of these results because there is no evidence that the relationship is causal. For instance, a critically acclaimed movie with a high metacritic score would probably not make more revenue if it were cold opened. Selection of particular types of movies that benefit

⁸ Note that this relationship is also found between cold opening and opening weekend and total U.S. box office (no logarithm). So this relationship is not just a result of the functional form of the regression.

⁹ For the average gross of a cold-opened movie, \$25 million, this is roughly \$1.5–\$8.75 million of box office revenue.

Table 1 Summary Statistics for Variables

Variable	Mean, (standard error), all movies	Minimum and maximum	Difference between means, cold-opened and reviewed movies, (standard error) ^a
Total box office revenue (in millions)	58.501 (1.806)	0.117 677.796	-38.321*** (2.642)
First weekend box office (in millions)	17.961 (0.507)	0.086 141.918	-6.826*** (1.948)
Metacritic rating	49.601 (0.780)	0.743 99.845	-31.746*** (1.741)
IMDb user rating	5.862 (0.034)	1.100 8.900	-1.472*** (0.109)
Theaters opened	2,498.633 (20.899)	601.000 4,366.000	-363.006*** (61.008)
Production budget (in millions)	45.871 (1.094)	0.446 281.740	-24.693*** (2.161)
Advertising expenditures (in millions)	19.007 (0.259)	0.470 57.114	-10.413*** (0.480)
Average competitor budget	42.642 (0.870)	0.000 281.740	-10.432*** (0.000)
Average competitor advertising expenditures	18.269 (0.233)	0.000 59.076	-3.469*** (0.676)
Average star ranking of lead roles ^b	2.125E+06 (2.122E+06)	1.500 3.000E+09	-1.472*** (0.109)
Summer open (1 = June, July, August)	0.250 (0.012)	0.000 1.000	-0.061* (0.034)
Adaptation or sequel (1 = yes)	0.612 (0.013)	0.000 1.000	0.092** (0.039)
Days released before Friday ^c (1 = Thursday, etc.)	0.201 (0.018)	-4.000 4.000	-0.116** (0.053)
Opening weekend continues after Sunday (1 = Monday, etc.)	0.111 (0.009)	0.000 2.000	0.020 (0.029)
Months released earlier in foreign country (months)	0.417 (0.076)	0.000 81.610	0.002 (0.220)
Action or adventure (1)	0.149 (0.009)	0.000 1.000	-0.051* (0.026)
Animated (1)	0.065 (0.007)	0.000 1.000	-0.060*** (0.011)
Comedy (1)	0.361 (0.013)	0.000 1.000	-0.068* (0.039)
Documentary (1)	0.007 (0.002)	0.000 1.000	0.013 (0.011)
Fantasy or science fiction (1)	0.069 (0.007)	0.000 1.000	0.005 (0.022)
Supense or horror (1)	0.179 (0.010)	0.000 1.000	0.304 (0.040)
Year of release (2003 = 0)	1.670 (0.076)	-3.000 6.000	1.531*** (0.220)
PG (1)	0.173 (0.010)	0.000 1.000	-0.127*** (0.022)
PG-13 (1)	0.463 (0.013)	0.000 1.000	0.059 (0.042)
R(1) (r)	0.325 (0.012)	0.000 0.000	0.083** (0.041)
Observations ^d	1,414	1,414	163/1,251

^aT-tests assume unequal variance. Standard error is the square root of the weighted average of sample variances.

^bFor movies that do not have a second actor (e.g., a nature documentary with a narrator). The second star value is chosen arbitrarily high at six billion to represent the effect of no second star.

^cThis value is calculated in regard to the Friday of a movie's "opening weekend." We follow the industry's classification on opening weekend, and make no decisions ourselves.

^dThere are 1,414 observations for all variables except 1,413 (1,251 screened, 162 cold) for metacritic and 1,303 (1,155 screened, 136 cold) for production budget.

*Significant at the 10% level; **significant at the 5% level; ***significant at the 1% level.

Table 2 Regressions on Box Office Revenues (in Millions)

Dependent variable:	Log opening weekend box office revenue			Log total box office revenue		
	2000–2009 (1)	2000–2005 (2)	2006–2009 (3)	2000–2009 (4)	2000–2005 (5)	2006–2009 (6)
<i>Cold opening</i>	0.204*** (0.057)	0.059 (0.084)	0.301*** (0.086)	0.292*** (0.064)	0.123 (0.095)	0.386*** (0.097)
<i>Metacritic rating</i>	0.005*** (0.001)	0.004*** (0.001)	0.007*** (0.001)	0.005*** (0.001)	0.005*** (0.001)	0.006*** (0.002)
<i>IMDb user rating</i>	0.010 (0.020)	0.040 (0.025)	−0.040 (0.033)	0.100*** (0.022)	0.122*** (0.028)	0.069* (0.037)
Log theaters opened	1.169*** (0.067)	0.999*** (0.083)	1.273*** (0.107)	0.989*** (0.075)	0.823*** (0.095)	1.092*** (0.121)
Log production budget	0.011 (0.026)	−0.072** (0.031)	0.103** (0.042)	0.008 (0.029)	−0.098*** (0.036)	0.127*** (0.047)
Log advertising expenditures	0.512*** (0.044)	0.630*** (0.057)	0.412*** (0.067)	0.742*** (0.049)	0.921*** (0.065)	0.595*** (0.075)
Average log competitor budget	−0.053** (0.023)	−0.025 (0.032)	−0.061* (0.034)	−0.021 (0.026)	−0.009 (0.036)	−0.017 (0.038)
Average log competitor advertising expenditures	−0.029 (0.030)	−0.076* (0.040)	0.016 (0.046)	−0.033 (0.034)	−0.077* (0.046)	0.011 (0.051)
Average log star ranking	−0.010* (0.006)	−0.015** (0.007)	−0.026* (0.014)	−0.015** (0.006)	−0.015* (0.008)	−0.026* (0.016)
Adaptation or sequel	0.060* (0.036)	0.157*** (0.038)	−0.313*** (0.095)	0.106*** (0.041)	0.170*** (0.044)	−0.196* (0.107)
Genre dummy variables included	Yes	Yes	Yes	Yes	Yes	Yes
MPPAA rating dummy variables included	Yes	Yes	Yes	Yes	Yes	Yes
Release date timing variables included ^a	Yes	Yes	Yes	Yes	Yes	Yes
Observations	1,303	778	525	1,303	778	525
R ²	0.684	0.712	0.702	0.700	0.727	0.707

Notes. All regressions include a constant term. Tables displaying coefficients for all regressors are available in the Web appendix.

^aRelease date timing variables are “year,” “days released before friday,” “opening weekend continues after sunday,” “months released earlier in a foreign country,” and “summer open.”

*Significant at the 10% level; **significant at the 5% level; ***significant at the 1% level.

from cold opening is likely to be contributing to the regression results. Cold-opened movies all have metacritic scores under 67 and have a mean of 30. We do not have data for high-quality movies that are cold opened, because studios never make this choice. We account for the effect of this selection for cold opening using propensity matching in this section and later in our structural model (§§3 and 4).

Propensity score matching techniques involve running a logistic regression to determine which other variables are the most associated with a cold opening (see the Web appendix). These predicted cold probabilities can be used to estimate which movies were the most likely to be cold opened. Running weighted regressions with these values, we can ignore movies that are very unlikely to be cold opened and match movies that were and were not cold opened but had similar propensities. Table 3 shows the results of three types of propensity score matching for weekend and cumulative U.S. box office data.

The propensity score matching results find that cold opening is correlated with a 35%–55% positive increase in revenue for U.S. opening weekend

and cumulative box office. This result suggests a poor-quality movie could increase its revenue by one-third to one-half by cold opening. Nearest neighbor matching—a technique that matches each cold-opened movie (*j*) with the regular released movie that has the closest propensity (to *j*) to have been cold opened—finds the highest positive correlation of 50%. Other matching techniques that use more movies (596 versus 72), but weigh each film differently, predict a lower value for the cold-opening premium (30%–40%). Taken together, these results suggest that the positive cold-opening premium is not a result of comparing cold-opened movies to larger, blockbuster movies that would never be cold opened, because the propensity of these large movies is low and they are ignored and receive low weight. Instead, the better differential performance of cold-opened movies compared to their equally poor quality screened-for-critics, counterparts is associated with the cold-opening premium.

However, there is further evidence to suggest that the cold-opening premium was not as pronounced during 2000–2005 as it was in 2006–2009. When these

Downloaded from informs.org by [131.215.23.31] on 28 February 2014, at 16:09. For personal use only, all rights reserved.

Table 3 Propensity Score Matching Results for Logged U.S. Cumulative and Weekend Box Office, 2000–2009 Movies

Dependent variable	Specification	Number in treatment group	Number in control group	Average treatment effect on the treated	Standard error	t-statistic
Log <i>opening weekend box office revenue</i>	Nearest neighbor	138	76	0.482	0.178	2.715***
Log <i>opening weekend box office revenue</i>	Stratification	138	587	0.404	0.193	2.093**
Log <i>opening weekend box office revenue</i>	Kernel matching	138	587	0.358	0.151	2.369**
Log <i>total box office revenue</i>	Nearest neighbor	138	76	0.530	0.201	2.645***
Log <i>total box office revenue</i>	Stratification	138	587	0.455	0.190	2.401**
Log <i>total box office revenue</i>	Kernel matching	138	587	0.400	0.191	2.097**

Notes. Cold opening is the treatment variable. A logit regression including all 23 variables and a constant (see the Web appendix for the regression) was used to generate propensity scores. All specifications are over the area of common support.

*Significant at the 10% level; **significant at the 5% level; ***significant at the 1% level.

propensity score matching techniques are used on the two specific periods, 2000–2005 and 2006–2009, separately, the coefficient for the cold-opening premium in the first period is insignificant (and negative in sign; see the Web appendix). This provides more evidence of a regime shift after 2005 in the profitability of cold-opened movies, or it may be due to other factors. Nonetheless, the overall result on the profitability of cold-opened movies is strong: there is a pronounced “cold-opening” premium in the data in the entire sample, using both regression and propensity matching.

2.1. Five Stylized Facts About Cold Openings

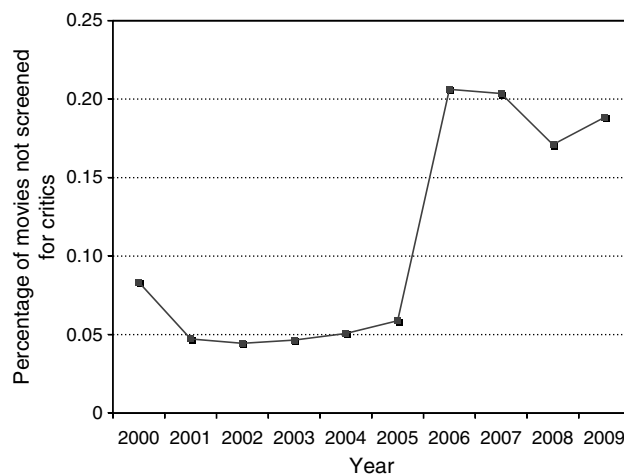
Our companion paper, Brown et al. (2012), notes five main stylized facts about cold openings that any explanation for cold opening must explain. That paper argues that other explanations such as moviegoers not learning about reviews, angry critics, and consumer-critic differences are unlikely to explain all five facts. The box office premium could be due to an omitted variable that is correlated with the decision to cold open, a possibility that is difficult to rule out. Interviews with industry executives did not suggest any such variable. A promising candidate variable is an unusually good print ad or movie trailer that makes an awful movie look great. Studios should spend extra on marketing to promote such movies if they cold open them. However, the interaction between (demeaned) marketing budget and cold-opening dummy has a negative and highly significant effect on box office in both time periods (the full-sample coefficient is -0.276 , $t = -3.46$). Thus, unusually expensive marketing is associated with *lower* cold-opening box office, which is inconsistent with a “great trailer” type of omitted variable explanation.

We do not go into details of other explanations here (see Brown et al. 2012). Instead we note that the standard Nash model cannot explain any of these facts,

but an extended version of the CH model has an explanation for all five:

1. There is an apparent correlation between cold opening and U.S. box office revenue.
2. The correlation is very similar whether quality ratings are derived from critics (metacritic) or from fans who saw the movie (Internet Movie Database (IMDb)).
3. The correlation is less pronounced in non-U.S. markets, especially the foreign language market Mexico, where releases are typically later, after U.S. reviews are available.
4. IMDb fan ratings are about 0.5 points lower (on a 10-point scale) for cold-opened movies than for comparable-quality movies that were not cold opened, suggesting fans are disappointed.
5. Cold openings are rare overall, but are increasingly frequent over the years in the sample (as shown in Figure 1).

The standard neoclassical model cannot easily explain any one of these five facts. It does not predict

Figure 1 Percentage of Widely Released Movies Cold Opened by Year, 2000–2009

a cold-opening premium (facts 1 and 2). It does not predict differential performance of cold openings in other markets, because quality information should be correctly inferred in all markets (fact 3). It does not explain why cold-opened movies have lower IMDb fan ratings (fact 4). Because it predicts cold openings should not happen, or should only happen by studio idiosyncratic error, it does not predict why they should be more frequent in the second part of the data set (fact 5).

A CH model can provide plausible explanations for the first four facts, and the conclusions of experiments that inspired the CH model can provide an explanation for fact 5. Cold openings generate a box office premium due to the limited strategic thinking of moviegoers in nondisclosure, because they believe cold-opened movies have higher quality than in actuality (facts 1 and 2). In foreign markets where quality information is already known, cold openings do not have this premium because moviegoers infer quality correctly (fact 3). More moviegoers go to cold openings than if the movie had been screened for critics, because they infer quality incorrectly, therefore the average fan rating of cold-opened movies will be lower than screened movies (fact 4).¹⁰

Fact 5 is that the rate of cold opening goes up over time. The baseline model explanation of this fact is that idiosyncratic error is going up over time, which is unlikely. However, the CH model does not have an immediate explanation of the rate of increase either. Keep in mind that CH models were initially developed to explain “preequilibrated” behavior in one-shot games (and to supply initial conditions for learning models). They may or may not have much explanatory power in settings like this, in which studios and moviegoers make a few decisions a year over ten years. Our explanation is that early in the sample, studios underestimated how naïve moviegoers could be (evidenced by the very high estimates of strategic sophistication in 2000–2005; see §4 for these estimations). Whereas moviegoers learned slowly, studios learned more quickly after noticing that cold-opened movies often did fine at the box office despite no

reviews (or, in the CH approach, *because* there were no reviews). This asymmetry in learning is consistent with the rise in cold openings over time, but there may well be other explanations that are not incorporated in either equilibrium or CH models.

3. The General Model

The initial regressions in §2 were not designed to understand the endogenous choice of studios to cold open and the likely reactions of moviegoers. Instead, we create a structural model of movie viewing and studio choice where moviegoers choose whether to see a movie and studios choose whether to screen the movie for critics. Our aim is to create a model that can be analyzed with box office data and studio choice, in which each side simultaneously maximizes utility and profit, respectively, but also a model that we may augment to allow estimation of parameters of limited strategic thinking concerning beliefs. To ensure that we can calculate equilibrium strategies for moviegoers and consumers, this model is static. However, we will examine the model over two different time periods to account for the sudden and sustained increase in cold opening at the end of our data set (i.e., 2000–2005 and 2006–2009; see Figure 1).

Formally, let movie j have characteristics, X_j , that are known to the studio and moviegoers. We assume that studios know the quality of their movie, q_j , and then choose whether to open cold ($c_j = 1$) or to screen for critics in advance ($c_j = 0$). Moviegoers do not know q_j and form a belief $E_m(q_j | c_j, X_j)$ that depends on a movie’s characteristics, X_j , and whether it was cold opened, c_j .¹¹

To model moviegoer utility functions and studio profit functions, we use an approach similar to Seim (2006), who examined the equilibrium entry decision in the video rental market of multiple firms. Whereas that paper examined an equilibrium of homogeneous firms, our paper examines the equilibrium between moviegoers and studios who have different objective functions.¹² Moviegoers form utility estimates of a given movie based upon its characteristics and expected quality, subtracting the ticket price, γ :

$$U(X_j, E_m(q_j | c_j, X_j)) = \alpha E_m(q_j | c_j, X_j) + \beta X_j - \gamma + \epsilon_j. \quad (2)$$

¹⁰ For an explanatory example, consider the moviegoer who overestimates a movie’s quality based on strategic naïveté. Suppose that if he knew the movie’s true quality, or knew that it was cold opened and inferred correctly its lower quality, he would not see it in theaters. This type of moviegoer is present in the audience of cold-opened movies but does not go to movies screened for critics. Thus, IMDb ratings for cold-opened movies are lower, all else being equal, because these moviegoers bring down their ratings, but not for movies screened for critics. By analogy, imagine an expensive restaurant that posts a menu online but does not list prices. If the highest-priced restaurants withhold prices, and naïve diners do not infer that relation, they will always be complaining about the surprisingly expensive prices at the restaurants they go to that did not post a menu.

¹¹ It is not crucial that moviegoers literally know whether or not a movie has been cold opened (e.g., surveys are likely to show that many moviegoers do not know). The essential assumption for analysis is that beliefs are approximately accurate for prereviewed movies and formed based on some different behavioral assumption for cold-opened movies.

¹² This general approach has also been used in previous studies of limited rationality (see Goldfarb and Yang 2009, Goldfarb and Xiao 2011), although all only studied producer behavior.

The term ϵ_j represents moviegoers' idiosyncratic preferences over movie j . Similar to Seim (2006), we assume this term is private information known to the moviegoer and independently and identically distributed from a logisitic distribution (e.g., McFadden 1974). We define the opportunity utility of not going to the movies as zero.¹³ The probability that the moviegoer will go to movie j with characteristics X_j and expected quality $E_m(q_j | c_j, X_j)$ at ticket price γ ¹⁴ is¹⁵

$$\begin{aligned} p(X_j, E_m(q_j | c_j, X_j)) &= P(U(X_j, E_m(q_j | c_j, X_j)) > 0) \\ &= P(\epsilon_j > \gamma - \alpha E_m(q_j | c_j, X_j) - \beta X_j) \\ &= \frac{\exp(\alpha E_m(q_j | c_j, X_j) + \beta X_j - \gamma)}{\exp(\alpha E_m(q_j | c_j, X_j) + \beta X_j - \gamma) + \exp(0)} \\ &= \frac{1}{1 + \exp(\gamma - \alpha E_m(q_j | c_j, X_j) - \beta X_j)}. \end{aligned} \quad (3)$$

We use a representative-agent approach to model moviegoers. We assume $p(X_j, E_m(q_j | c_j, X_j))$ is the total share of moviegoers that go to movie j .¹⁶ We define the constant M as the maximum amount of box office revenue that could be earned in the period if every moviegoer went to a movie. Then we can define expected revenue as

$$\begin{aligned} \bar{R}(c_j, X_j, q_j) &= Mp(X_j, E_m(q_j | c_j, X_j)) \\ &= \frac{M}{1 + \exp(\gamma - \alpha E_m(q_j | c_j, X_j) - \beta X_j)}. \end{aligned} \quad (4)$$

Movie studios make decisions whether to screen movies for critics based on their expected revenues

¹³ This is without loss of generality because a constant term is included in the revenue regression, which in this model is equivalent to the estimated utility of not going to the movie.

¹⁴ The term γ is fixed at the average U.S. ticket price in 2005, \$6.71 (recall box office revenues are in 2005 dollars).

¹⁵ Note that this formalization is a single variable logit and not a multinomial one. We do not assume that a moviegoer can only go to one movie in this period. We make this decision because it is not clear whether movies crowd out other movies at the theater or have positive spillover effects. The sign on our initial regression for the competition term (see Table 2) suggests that, if anything, the latter explanation is more likely. Given these facts and that our estimation period is the opening weekend, and the widely released movies usually are released no more than four at a time, diversified over genre, we do not find the crowding out explanation plausible enough to warrant using a multinomial logit.

¹⁶ We choose this approach rather than aggregating $p(X_j, E_m(q_j | c_j, X_j))$ over some N to avoid arbitrarily large precision in our observations. Because box office numbers and studio decisions will be combined in a maximum-likelihood-estimation process to estimate the parameters of this model jointly, we believed each observation should be counted equally. If we chose to have N consumers to make up the box office, we then have N times more precision on our moviegoer data compared to studio data.

(Equation (4)). Studios also have idiosyncratic error term, v_j , about the additional success of the movie if it is cold opened. As with the moviegoer error term, v_j is private information to studios and independently and identically logistically distributed. Studios will cold open a movie if $\bar{R}(1, X_j, q_j) + v_j > \bar{R}(0, X_j, q_j)$. The probability that a studio will cold open movie j given its characteristics, X_j , and quality, q_j , is

$$\begin{aligned} \pi(X_j, q_j) &= P[v_j < \bar{R}(0, X_j, q_j) - \bar{R}(1, X_j, q_j)] \\ &= \frac{\exp[\bar{R}(1, X_j, q_j)]}{\exp[\bar{R}(1, X_j, q_j)] + \exp[\bar{R}(0, X_j, q_j)]} \\ &= \frac{1}{1 + \exp[\bar{R}(1, X_j, q_j) - \bar{R}(0, X_j, q_j)]} \\ &= \frac{1}{1 + \exp[M[p(X_j, E_m(q_j | 1, X_j)) - p(X_j, E_m(q_j | 0, X_j))]]}. \end{aligned} \quad (5)$$

The term $E_m(q_j | 0, X_j)$, the expected quality of a movie that is released to critics, is determined exogenously. Because critics can write about a movie they screen, as well as reveal their estimates about quality in ways that are relatively costless (i.e., Internet sites, newspapers), we assume that if a movie is screened to critics, its quality is then perfectly known to moviegoers. We also assume studios are aware that critics reveal their quality. Assumption 1 states that for movies screened to critics, moviegoers have accurate perceptions of the quality of the movie, and studios have accurate perceptions about moviegoer perceived quality.¹⁷

ASSUMPTION 1. $E_m[q_j | 0, X_j] = q_j$.

We also make a simplifying assumption about moviegoer perceived quality that allows our structural model to match our motivating example. Recall that in our disclosure example, we went through iterations of quality (e.g., 50, 25, 12.5) without discussing other movie characteristics (X_j). Our models will also make this assumption.

ASSUMPTION 2. $E_m[q_j | 1, X_j]$ does not depend on X_j . That is, $E_m[q_j | 1, X_j] \equiv E_m[q_j | 1]$.

As Milgrom (1981) and Grossman (1981) demonstrate for all disclosure games, $E_m[q_j | 1] = 0$ in Nash equilibrium, as the system completely unravels.¹⁸ Our

¹⁷ Quality could also be known with noise and all results would hold if moviegoers are risk neutral.

¹⁸ Alternatively, one could consider the value $E_m[q_j | 1]$ to be bounded on the interval $[0, 100]$ in a general form of the cursed equilibrium model (Eyster and Rabin 2005). We will estimate the Nash equilibrium model in this paper, and leave the alternative specification in the Web appendix.

estimation techniques will use maximum-likelihood estimation to estimate the parameters α and β that best fit the joint system. This estimation technique is explained in detail in §3.2.

3.1. A Cognitive Hierarchy Model for Moviegoers and Distributors

The alternative structural behavioral model that originally inspired this research, the cognitive hierarchy model (Camerer et al. 2004), makes a different assumption about $E_m(q_j | 1)$ than the equilibrium restrictions in Assumption 2. The behavioral model relaxes the assumption that moviegoers go through all the iterations of strategic thinking necessary to reach the game’s Nash equilibrium and corresponding quality estimate. Similarly, distributors may best respond to moviegoers who have only done a limited number of steps of strategic thinking. The CH model can characterize aggregate strategic behavior with a single parameter, τ .¹⁹

The CH model assumes that there is a population of individuals who do varying numbers of steps of iterative strategic thinking. The parameter τ determines the distribution of steps of thinking by the one-parameter Poisson distribution

$$P(x = n | \tau) = \frac{\tau^n e^{-\tau}}{n!}, \quad (6)$$

where τ is the mean number of steps of strategic thinking. To develop the model similar to our baseline, we restrict moviegoer inference of quality to not include specific movie characteristics, by revising Assumption 2.

ASSUMPTION 2’. For all k , $E_m^k[q_j | 1, X_j, \tau]$ does not depend on X_j . That is, $E_m^k[q_j | 1, X_j, \tau] \equiv E_m^k[q_j | 1, \tau]$.

Zero-level moviegoers do not think about the studio’s actions of cold opening a movie. They act as if the movie’s quality is average, $E_m^0(q_j | 1, \tau) = 50$.²⁰

¹⁹ The parsimony of the single parameter specification is the reason we have chosen this approach over using a level- k model (see Stahl and Wilson 1994; Costa-Gomes et al. 2001; Costa-Gomes and Crawford 2006; Crawford and Iriberri 2007a, b). Also Camerer et al. (2004) found that the Poisson restriction fit almost as well as models with several free parameters for different level frequencies. Given that we have only box office data and studio decisions to cold open, we would be unable to identify the proportion of levels in the population without some type of distributional assumption. The Poisson version of CH does give such an assumption. Other approaches are possible.

²⁰ Assuming that 0-level players choose randomly across possible strategies is natural in many games. However, the more appropriate, general interpretation is that 0-level players are simple, or heuristic, rather than necessarily random. For example, in “hide-and-seek” games a natural starting point is to choose a “focal” strategy (see Crawford and Iriberri 2007a). In auctions a natural starting point is to bid one’s value (Crawford and Iriberri 2007b).

They will go to any movie with probability

$$p(X_j, E_m^0(q_j | 1, \tau)) = \frac{1}{1 + \exp[\gamma - \beta X_j - 50\alpha]}. \quad (7)$$

A 0-level studio will best respond to the 0-level moviegoer.²¹ The 0-level studio calculates the expected revenue from cold opening a movie as

$$\bar{R}_0(1, X_j, q_j, \tau) = Mp(X_j, E_m^0(q_j | 1, \tau)). \quad (8)$$

It will therefore cold open movie j with probability

$$\begin{aligned} \pi_0(X_j, q_j, \tau) &= \frac{1}{1 + \exp[Mp(X_j, q_j) - \bar{R}_0(1, X_j, q_j, \tau)]} \\ &= \frac{1}{1 + \exp[M[p(X_j, q_j) - p(X_j, 50)]]}. \end{aligned} \quad (9)$$

Proceeding inductively, for $k > 0$ moviegoers will consider the expectations of all moviegoers of lower types ($k' < k$). They will form a conditional expectation using τ of lower level types and assume studios only cold open movies with quality lower than that expectation (as in our motivating example). Their expectation about the quality of cold-opened movies will be the average quality of movies below that threshold. Formally,

$$\begin{aligned} E_m^k(q_j | 1, \tau) &= \frac{\int_0^{\sum_{n=0}^{k-1} P(x=n|\tau)E_m^n(q_j|1,\tau)} qP(q) dq}{\int_0^{\sum_{n=0}^{k-1} P(x=n|\tau)E_m^n(q_j|1,\tau)} P(q) dq} \\ &= \frac{1}{2} \sum_{j=0}^{k-1} P(x = n | \tau) E_m^j(q_j | 1, \tau). \end{aligned} \quad (10)$$

Notice that (10) fits our motivating example well. A 0-level moviegoer believes cold-opened movies have a quality of 50. A 1-level moviegoer knows this fact, assumes studios will only cold open movies below a quality of 50, and given a uniform distribution of quality infers cold-opened movies have a quality of 25. A 2-level moviegoer averages 50

In our game, random choice by moviegoers would mean random attendance at movies. That specification of 0-level play does not work well because it generates far too much box office revenue. It is admittedly not ideal to have special ad hoc assumptions for different games. Eventually, we expect there will be a theory of 0-level play that maps the game structure and a concept of simplicity or heuristic behavior into 0-level specifications in a parsimonious way.

²¹ An alternate specification, more in line with the spirit of experimental work, would have 0-level studios cold open at random. The issue with this specification is that then both 0 and 1-level moviegoers believe cold openings have an expected quality of 50. Depending on values of τ , this can lead to a pattern of two successive levels of moviegoers or studios behaving in the same way, creating an identification issue. For this reason, we avoid this specification. See Brown et al. (2009) for a specification more in line with the experimental spirit.

Table 4 Moviegoer Inferred Quality and Predicted Attendance by Level of Thinking for *When a Stranger Calls* at $\tau_m = \tau_s = 1.638$

Moviegoers steps of thinking	Inferred quality of <i>When a Stranger Calls</i> , given it is opened cold, ^a $\tau_m = 1.638$	Inferred quality of <i>When a Stranger Calls</i> , given it is opened cold, ^a $\tau_m = 100$	Proportion of types out of maximum possible that will attend movie, ^b $\tau_m = 1.638$	Probability of studio cold-opening movie, knowing $q_j = 6.179$, ^b $\tau_s = 1.638$
0	50.0000	50.0000	0.0054	0.5843
1	25.0000	25.0000	0.0047	0.5339
2	17.2383	12.5000	0.0045	0.5195
3	14.3325	6.2500	0.0044	0.5143
4	13.2184	3.1250	0.0044	0.5123
5	12.8228	1.5625	0.0044	0.5116
6	12.6995	0.7813	0.0044	0.5114
7	12.6664	0.3906	0.0044	0.5113
8	12.6586	0.1953	0.0044	0.5113
9	12.6570	0.0977	0.0044	0.5113
10	12.6567	0.0488	0.0044	0.5113

^aInferred quality by levels of thinking is the same for all movies. By assumption, it does not depend on X_j .

^bMoviegoer attendance and probability of cold opening do depend on X_j , so these values are specific to this movie.

and 25 using τ and believes expected quality is half this average as studios best respond to a distribution of 0 and 1-level moviegoers. A k -level moviegoer will attend a movie using the same equation as before:

$$p(X_j, E_m^k(q_j | 1, \tau)) = \frac{1}{1 + \exp[\gamma - \beta X_j - E_m^k(q_j | 1, \tau)\alpha]}. \quad (11)$$

For $k > 0$ a studio best responds to a distribution of $k' \leq k$ determined by τ . Their choice to cold open is also dependent on their movie's specific characteristics through expected revenue. They will calculate expected revenue using

$$\bar{R}_k(1, X_j, q_j, \tau) = \frac{M}{P(x < k | \tau)} \sum_{n=0}^k P(x = n | \tau) p(X_j, E_m^n(q_j | 1, \tau)). \quad (12)$$

This leads to k -level, movie-specific, probability of cold opening,

$$\pi_k(X_j, q_j, \tau) = \frac{1}{1 + \exp[Mp(X_j, q_j) - \bar{R}_k(1, X_j, q_j, \tau)]}. \quad (13)$$

As an example, Table 4 shows values for the first 10 steps of thinking for a cold-opened movie, *When a Stranger Calls*, when $\tau_m = 1.638$ and 100. Moviegoers' inference is determined by (10). They make a decision whether to go to the movie from (11), which determines the proportion of moviegoers that attend. A k -level studio best responds to a distribution of moviegoers $k' \leq k$. Because they know the quality of their movie, they make a decision about whether to cold open by comparing the expected revenues given moviegoers' inferred quality (conditional on a cold-open choice) and the true quality and including an

idiosyncratic error term (to model stochastic choice). Notice that the values of inferred quality are the same for all movies given the steps of thinking (they do not depend on X_j by Assumption 2'), but the proportion of moviegoers that see the movie and the cold-opening probabilities depend on X_j , so those values are unique to this movie.

3.2. Estimation

Before the estimation procedure is explained, a few of the numbers used in the process must be clarified. The logic of the model and our data (see §2 and Table 2) suggest that cold opening most strongly affects the first weekend's revenue (which may then affect cumulative revenue). Therefore, we use the first weekend's revenue to calibrate the models' revenue equations and studio decisions. Thus, our representation of revenue, $R(X_j, E_m(c_j, X_j))$, will use weekend box office revenue normalized to 2005 dollars. Movie ticket prices are also in 2005 dollars. The value M , the maximum possible box office, is chosen as double the highest weekend gross over the set being evaluated.²² Movie quality, q_j , is the standardized version of the average metacritic rating used in §2. Movie characteristics X_j are the independent variables used in the initial regressions on weekend box office in §2, excluding cold opening and critic rating. The term c_j has the same value as the cold dummy in §2.

We jointly estimate the parameters using box office revenue data and studio-cold-opening decisions in a maximum-likelihood-estimation procedure. Equation (4), which represents the expected box office revenue in our model, is nonlinear and requires a transformation to fit a linear model. We estimate the

²² For 2000–2005 data, $M = 249.46$; for 2006–2009 data, $M = 283.82$.

model equation with movie specific error term ξ_j , which is normally distributed, $N(0, \sigma)$. That is,

$$\log\left(\frac{M}{R(X_j, E_m(c_j, X_j))} - 1\right) + \gamma = \alpha E_m(q_j | c_j) + \beta X_j + \xi_j. \quad (14)$$

Denoting the residuals of this linear model as e , we have a log-likelihood function,

$$L_m(\alpha, \beta) = -\frac{n}{2} \log(2\pi\sigma^2) + \left(-\frac{1}{2\sigma^2} e'e\right). \quad (15)$$

Because σ^2 is unknown, it will be estimated by $(e'e)/(N - Q)$, where N is the number of movies and Q is the number of movie characteristics including quality.²³

The log likelihood for the studio decisions is calculated using the estimated predicted probabilities of cold opening. For each set of parameter values, there is a predicted probability that a cold-opened movie *would have been* cold opened ($\pi(X_j, q_j)$). Similarly, for each set of parameter values, there is a predicted probability a movie that was screened for critics *would have been* screened for critics ($1 - \pi(X_j, q_j)$). The studio log-likelihood function is the product of these values, logged:

$$L_s(\alpha, \beta) = \sum_{j \in N} \log(c_j \pi(X_j, q_j) + (1 - c_j)(1 - \pi(X_j, q_j))). \quad (16)$$

The partial log likelihoods are summed to form a likelihood function that incorporates both box-office revenue and studio decisions. Estimates for the parameters in the model are obtained by maximizing the function $L(\alpha, \beta)$ defined by Equation (17):

$$L(\alpha, \beta) = L_m(\alpha, \beta) + L_s(\alpha, \beta). \quad (17)$$

A given set of values, (α, β) ,²⁴ is put into (15) and (16) and logged, and then summed to get a likelihood value in Equation (17). Maximum-likelihood parameter estimates, (α^*, β^*) , are obtained using an optimization algorithm (Nelder and Mead 1965) that begins at the origin. Standard errors of all coefficients are obtained by 100 random bootstraps of the data set using the same algorithm. For the bootstraps, the algorithm is started at the parameter estimates (α^*, β^*) instead of the origin.

The CH model uses the same basic maximum-likelihood-estimation procedure as the standard

model. The difference is the addition of the parameter τ . Because moviegoers and studios are different types of players with different objective functions and levels of experience, the two types of players are allowed to have differing τ parameters. Moviegoers have parameter τ_m and studios have parameter τ_s .²⁵ This will allow us to infer whether the degree of limited thinking by both studios and moviegoers is similar (i.e., if $\tau_m \approx \tau_s$).

As with the standard model, we jointly estimate the parameters of the CH model. For a given set of values $(\alpha, \beta, \tau_m, \tau_s)$, where $\tau_m, \tau_s \geq 0$, α, β , and τ_m are used to compute the expected box office of cold-opened (those where $c_j = 1$) movies.²⁶

The parameters α, β , and τ_s are then used to compute the probability of each movie being cold opened using Equation (13). Those values are put into Equation (16) to compute the other partial log likelihood. The values are then logged and summed to give a full maximum-likelihood value for the parameter values $(\alpha, \beta, \tau_m, \tau_s)$. Maximum-likelihood parameter estimates, $(\alpha^*, \beta^*, \tau_m^*, \tau_s^*)$, are obtained using a Nelder-Mead optimization algorithm that begins at the origin. Standard errors of all coefficients are obtained by 100 random bootstraps of the data set using the same algorithm. For the bootstraps the algorithm is started at the parameter estimates $(\alpha^*, \beta^*, \tau_m^*, \tau_s^*)$ instead of the origin. The results of both estimations are given in the next section.

4. Structural Estimation Results

Table 5 presents the results of the estimation of both baseline and CH models separated over the periods 2000–2005 and 2006–2009. For studio choices,

²⁵ An alternative approach would be to have only one τ for moviegoer and studio behavior and jointly estimate it based on studio decisions and box office data. The trouble with this approach is that because cold opening occurs so infrequently, the number of observations that determine the studio's parameter τ_s are roughly ten times as great as the number of observations that determine moviegoers' behavior τ_m . For this reason, any joint estimation of this type will be highly biased toward studio behavior (which already resembles the standard model) and neglect the cold-opening premium, the primary motivation for this exercise.

²⁶ In all estimations, for a given parameter value τ , Equation (13) is approximated up to the level $k = 100$. All probability for values $k > 100$ was assigned to $k = 100$. A τ value of 100, the upper limit, was an entire distribution of 100-level thinkers. To allow an identical maximum-likelihood-estimation procedure with the baseline model, we use a single-quality dimension, \hat{q} , such that

$$\bar{R}(0, X_j, \hat{q}) = \lim_{k \rightarrow \infty} \bar{R}_k(1, X_j, q_j, \tau_m).$$

Basically, \hat{q} functions as the single value of expected quality that would generate the same expected revenue as the CH model with parameter τ_m . This value is used for all cold-opened movies, and q_j is used for all regularly released movies, to calculate the partial log likelihood in Equation (15).

²³ The constant $Q = 24$ for the 23 movie characteristics plus critic rating. For 2000–2005 data, $N = 778$; for 2006–2009 data, $N = 558$.

²⁴ An alternate baseline specification treats $E(q_j | 1)$ as a free parameter instead of setting it to zero; see the Web appendix for the results of that specification.

Table 5 Parameter Estimates for Jointly Estimated Baseline and CH Models by Time Period, Using Weekend Box Office Revenue and Cold-Opening Decisions Data

Dependent variable:	<i>Moviegoer attends movie</i>			
	Baseline		Cognitive hierarchy	
	2000–2005 (1)	2006–2009 (2)	2000–2005 (3)	2006–2009 (4)
<i>Moviegoer mean steps of thinking</i> (τ)	—	—	1.638 (9.927)	0.000 (0.000)
<i>Studio mean steps of thinking</i> (τ)	—	—	100.000*** (13.229)	5.022 (9.636)
<i>Metacritic rating</i>	0.008*** (0.001)	0.004*** (0.000)	0.008*** (0.001)	0.006*** (0.001)
<i>IMDb user rating</i>	−0.023 (0.027)	−0.012 (0.029)	−0.018 (0.028)	−0.014 (0.029)
<i>Log theaters opened</i>	1.061*** (0.100)	1.299*** (0.146)	1.057*** (0.099)	1.284*** (0.142)
<i>Log production budget</i>	−0.066* (0.036)	0.117** (0.056)	−0.065* (0.036)	0.131** (0.056)
<i>Log advertising expenditures</i>	0.621*** (0.081)	0.406*** (0.091)	0.634*** (0.083)	0.426*** (0.091)
<i>Average log competitor budget</i>	−0.040 (0.036)	−0.069* (0.036)	−0.044 (0.036)	−0.071* (0.036)
<i>Average log competitor advertising expenditures</i>	−0.079 (0.049)	0.010 (0.053)	−0.074 (0.049)	0.019 (0.054)
<i>Average log star ranking</i>	−0.020** (0.008)	−0.025 (0.015)	−0.020** (0.008)	−0.026* (0.015)
<i>Adaptation or sequel</i>	0.072 (0.050)	−0.014 (0.057)	0.157*** (0.039)	−0.332*** (0.081)
Genre dummy variables included	Yes	Yes	Yes	Yes
MPAA rating dummy variables included	Yes	Yes	Yes	Yes
Release date timing variables included	Yes	Yes	Yes	Yes
Average quality of cold-opened movie ^a	19.172*** (0.837)	29.987*** (1.252)	19.023*** (0.805)	26.205*** (1.623)
Predicted cold-opening percentage ^b	0.108*** (0.006)	0.179*** (0.013)	0.107*** (0.005)	0.182*** (0.017)
Predicted cold premium ^c	−0.101*** (0.014)	−0.075*** (0.007)	0.105 (0.103)	0.177*** (0.040)
Observations	778	525	778	525
Log likelihood	−755.19	−676.61	−752.49	−664.08

Notes. Tables displaying coefficients for all regressors are available in the Web appendix. Standard errors are calculated from 100 bootstraps for each model and time period.

^aActual values: 13.47, 24.49.

^bActual values: 0.055, 0.181.

^cActual values: 0.059, 0.301.

*Significant at the 10% level; **significant at the 5% level; ***significant at the 1% level.

the baseline equilibrium model (columns (1) and (2)) with $E_m(q_j | 1) \equiv 0$ predicts a general reluctance of studios to cold-open movies (only 11% in 2000–2005) but it is forced to use idiosyncratic error to explain the times studios do cold open.²⁷ The CH

model with best fitting $\tau_s = 100$, the upper bound, for 2000–2005 (column (3)) and $\tau_s = 5.022$, for 2006–2009 (column (4)) (values much higher than what is typically observed in laboratory studies) can also account for the low rate of cold openings.

The one parameter for which the models differ is, importantly, the predicted cold premium. In the baseline model, movies are cold opened only because of studio idiosyncratic error, and moviegoers think their quality is zero. This means, provided movie quality is above zero, that those movies should make *less* box office revenue than if they had been screened for

²⁷ Table 4 assumes $E_m(q_j | 1) = 0$. However, if that conditional expectation is a free parameter, the best-fitting values are 0 and 9.12 in the two time periods (see the Web appendix for that model). The log likelihood in the later period is −673.79, only two points better than the restricted model, and the predicted cold premium is −0.054 (0.015), so the restriction $E_m(q_j | 1) = 0$ is not producing the poor premium prediction.

critics (i.e., the predicted box office premium is negative). The baseline model's predictions reflect this effect, predicting that the average cold-opened movie should do 10.1% and 7.5% worse in 2000–2005 (column (1)) and 2006–2009 (column (2)), respectively. The CH model parameterizes moviegoers as relatively naïve, they do on average 1.636 (column (3)) and 0.00 (column (4)) steps of thinking in each of the two periods. This is equivalent to an expected quality ($E_m(q_j | 1)$) of 25.313 and 50.000, respectively. Because the cold-opened movies usually have qualities below these values, there is a positive cold-opening premium of 10.5% and 17.7% in each period, as moviegoers overestimate the quality of cold-opened movies. The CH predictions are therefore closer to actual cold-opening premiums of 5.9% and 30.1%.

It is true that the log likelihood is only slightly better for CH than for the baseline model. However, the only difference in fit comes from explaining a small percentage of cold openings (around 10%) and a modest premium (around 20%). Furthermore, the baseline model clearly misestimates the sign of the box office premium, so although the overall fit is not bad, the adequate fit comes from an idiosyncratic error explanation that gets the economics wrong. And, the difference in explanatory power does increase between CH and baseline from 2000–2005 to 2006–2009 as the frequency of cold openings increases (5.53% versus 18.1%) (the log likelihoods are -676 and -664).

Note that because of the $E_m(q_j | 1) = 0$ assumption, in this particular baseline model, moviegoers do not have correct Bayesian expectation of cold-opened quality. Moviegoers believe cold-opened movies have a quality of 0, when in fact they have an average quality of 19.17 and 29.98 because of studio idiosyncratic error. However, such a model that uses these correct expectations (see the Web appendix) has log-likelihood values of $-1,150$ and -828 for two respective periods, much worse than the Table 4 baseline or CH model fits.

In general, the estimated τ values for moviegoer behavior are much closer to those observed in laboratory experiments than the studio estimates. The value 1.638 is close to other experimental estimates (generally around 1–2.5). The estimated value of 0 for 2006–2009 implies a pure naïveté, not usually found in experimental data (see Camerer et al. 2004, Östling et al. 2011). However, the low value may be more understandable because 0- and 1-level moviegoers behave identically (see Footnote 21), so the fitted 0-level play may be capturing 1-level play as well. Studio estimates suggest Nash play in 2000–2005 (high estimated τ_s) and a high number of steps of thinking in 2006–2009 (high estimated τ_s). Because

studio executives making the decision to cold open think a great deal about their strategy, and have experience in these decisions (i.e., they are not new to these games), their higher sophistication compared to moviegoers may make some sense. At the same time, the value of 5.02 instead of 100 in the second time period suggests studios may be learning that moviegoers are more naïve than they thought in 2000–2005.

5. Conclusion

This paper is the first to apply a parametrized behavioral model to a game of disclosure in the field, an example of “structural behavioral economics.” We study a market in which information senders (movie studios) are strategically withholding information (the quality of their movie) from information receivers (moviegoers), by not showing movies to critics in time for reviews to be published before opening weekends. Contrary to the simple Nash equilibrium, there is a “box office premium”—cold-opened movies earn more than screened movies with similar characteristics.

We provide two structural models to explain the environment being studied. The baseline model has moviegoers expect cold openings to have the worst possible quality and critics to cold open entirely though idiosyncratic error. The CH model with a low number of thinking steps τ_m to represent moviegoer naïveté, and a high τ_s to represent studio oversophistication has the same general qualities of the baseline model but is also able to predict the cold-opening premium. Furthermore, the best-fitting τ_m values for moviegoers, derived from box office data, are relatively similar to those observed in laboratory studies. The studio's τ_s are much closer to Nash levels than those observed in laboratory experiments, but the shift of values from 2000–2005 to 2006–2009 suggests that studios may be learning to better respond to relatively naïve moviegoers.

The question remains why moviegoers have become more naïve about cold openings and appear to be regressing rather than learning. Cold openings appear to have *increased* in profitability in the later part of the decade, suggesting if anything consumers are inferring less about their quality than before. Although factors like repeated play and reputation of studios may explain the reluctance of studios to cold open, the continued naïveté of moviegoers is difficult for standard game-theoretic models to explain.²⁸ One

²⁸ Economic intuition and experiments on lemons (e.g., Lynch et al. 2001) suggest consumers will ultimately infer that goods whose quality is not disclosed have low quality.

explanation we prefer, though cannot prove, concerns the population of moviegoers. A third of tickets are sold to young moviegoers (ages 12–24) (Motion Picture Association of America 2010). Although moviegoers might learn over time that cold-opened movies are surprisingly bad, the market has overlapping generations. New, naïve moviegoers are always entering the population; there is no reason to expect rapid convergence of beliefs across such a population. Of course, this explanation is just speculation, and it is an open research topic to determine the dynamics of relatively young consumer populations.

This paper has connected a major area of economics and management research, models of limited strategic thinking, with a major area of economic research, games of selective disclosure. From a management perspective, companies may have some intuitions that it can be advantageous to selectively withhold bad quality information, not strictly following the equilibrium analysis of Grossman (1981) and Milgrom (1981) (see Brown et al. 2009 for examples). However, this paper begins to uncover *what level* of bad information should be withheld, and can help explain why. We find that in the movie industry, it appears that studios are withholding too little, although they appear to be learning quickly. Although the industry studied here, major movie studios, is quite unique, the main parts of the industry—products of unknown quality and critical review—are found in other industries. Moreover, many industries involve concentrated sellers than can learn to withhold and diffuse rotating consumers that will likely have difficulty learning. This suggests our approach could be applied to other industries: models of strategic thinking could be used in any industry that involves disclosure to examine what level of disclosure is optimal for managers given the limited strategic thinking of consumers.

Acknowledgments

The authors thank audiences at California Institute of Technology, the University of Chicago Booth School of Business, the University of California at Berkeley, Yale Graduate Student Conference on Behavioral Science, the 2007 North American Economic Science Association meeting, the 2007 Society for Judgment and Decision Making meeting, the 2007 Southern Economic Association meeting, the 2010 American Economic Association meeting, and the 8th Triennial Invitational Miami Choice Symposium, especially Stefano Della Vigna, David Grether, Stuart McDonald, Tom Palfrey, Charles Plott, Robert Sherman, and Leat Yariv. The authors thank Esther Hwang, Carmina Clarke, Ferdinand Dubin, Zachary Bethune, Cheng Cheng, Xinrong Li, and Jonathan Garrity for help with data collection. The authors also thank Matt Shum and the Bush School Empirical Work-

shop, especially Stephanie Houghton, Steve Puller, and Lori Taylor, for help during revision. This paper has benefitted from the very helpful comments of the department editor and three referees.

References

- Brown AL, Camerer CF, Lovoal D (2009) To review or not to review? Limited strategic thinking at the movie box office. Working paper, Texas A&M University, College Station.
- Brown AL, Camerer CF, Lovoal D (2012) To review or not to review? Limited strategic thinking at the movie box office. *Amer. Econom. J.: Microeconomics* 4(2):1–26.
- Camerer CF, Ho T-H, Chong J-K (2004) A cognitive hierarchy model of games. *Quart. J. Econom.* 119(3):861–898.
- Costa-Gomes MA, Crawford VP (2006) Cognition and behavior in two-person guessing games: An experimental study. *Amer. Econom. Rev.* 96(5):1737–1768.
- Costa-Gomes M, Crawford VP, Broseta B (2001) Cognition and behavior in normal-form games: An experimental study. *Econometrica* 69(5):1193–1235.
- Crawford VP, Iriberri N (2007a) Fatal attraction: Saliency, naivete, and sophistication in experimental “hide-and-seek” games. *Amer. Econom. Rev.* 97(5):1731–1750.
- Crawford VP, Iriberri N (2007b) Level-*k* auctions: Can a nonequilibrium model of strategic thinking explain the winner’s curse and overbidding in private-value auctions? *Econometrica* 75(6):1721–1770.
- Crawford VP, Costa-Gomes MA, Iriberri N (2010) Strategic thinking. Working paper, University of Aberdeen Business School, Aberdeen, UK.
- Dranove D, Jin GZ (2010) Quality disclosure and certification: Theory and practice. *J. Econom. Literature* 48(4):935–963.
- Eyster E, Rabin M (2005) Cursed equilibrium. *Econometrica* 73(5):1623–1672.
- Fishman MJ, Hagerty KM (2003) Mandatory versus voluntary disclosure in markets with informed and uninformed customers. *J. Law, Econom., Organ.* 19(1):45–63.
- Goldfarb A, Xiao M (2011) Who thinks about the competition? Managerial ability and strategic entry in U.S. local telephone markets. *Amer. Econom. Rev.* 101(7):3130–3161.
- Goldfarb A, Yang B (2009) Are all managers created equal? *J. Marketing Res.* 46(5):612–622.
- Grossman SJ (1981) The informational role of warranties and private disclosure about product quality. *J. Law Econom.* 24(3):461–483.
- Lynch M, Miller R, Plott CR, Porter R (2001) Product quality, informational efficiency, and regulations in experimental markets. Plott CR, ed. *Collected Papers on the Experimental Foundations of Economics and Political Science, Economists of the Twentieth Century* (Edward Elgar, Cheltenham, UK).
- McFadden D (1974) *Conditional Logit Analysis of Qualitative Choice Behavior* (Academic Press, New York).
- Milgrom PR (1981) Good news and bad news: Representation theorems and applications. *Bell J. Econom.* 12(2):380–391.
- Motion Picture Association of America (MPAA) (2010) Theatrical market statistics 2010. Report, MPAA, Washington, DC. Last accessed August 8, 2012, <http://mpaa.org/resources/93bbeb16-0e4d-4b7e-b085-3f41c459f9ac.pdf>.
- Nagel R (1995) Unraveling in guessing games: An experimental study. *Amer. Econom. Rev.* 85(5):1313–1326.
- Nelder JA, Mead R (1965) A simplex method for function minimization. *Comput. J.* 7(4):308–313.

- Östling R, Wang JT-y, Chou EY, Camerer CF (2011) Testing game theory in the field: Swedish LUPI lottery games. *Amer. Econom. J.: Microeconomics* 3(3):1–33.
- Seim K (2006) An empirical model of firm entry with endogenous product-type choices. *RAND J. Econom.* 37(3):619–640.
- Stahl DO, Wilson PW (1994) Experimental evidence on players' models of other players. *J. Econom. Behav. Organ.* 25(3):309–327.
- Stahl DO, Wilson PW (1995) On players' models of other players: Theory and experimental evidence. *Games Econom. Behav.* 10(1):218–254.
- Verrecchia RE (2001) Essays on disclosure. *J. Accounting Econom.* 32(1–3):97–180.
- Wang JT (2006) Is last minute bidding bad? Working paper, California Institute of Technology, Pasadena.