Consequences of omitting advertising in demand estimation: An application to theatrical movies

Charles C. Moul*

June 2006

Abstract

Given the difficulties of properly estimating demand in the presence of advertising, advertising has often been omitted. This paper uses a Monte Carlo experiment of price-and-ad-setters to show that the omission of advertising can cause substantial biases in estimation, even in markets with relatively low ad-sales ratios and when advertising appears to be statistically insignificant. I then validate the appropriateness of these Monte Carlo simulations with respect to advertising with an analysis of domestic theatrical movies. Results using actual movie data when advertising is omitted and included mimic those of the simulations. Omitting advertising from this context overstates the impact of theaters by 25% and substantially understates the impact of starring cast members (by 80%).

Keywords: Advertising, motion pictures, differentiated products

Despite being an active research area from the 1940s through the 1960s, the estimation of advertising’s impact on demand and resulting strategic implications has not received much attention from economists in the last thirty years. Much of this lack of attention can perhaps be attributed to Schmalensee (1972) which simultaneously provides a thorough literature review of the economics of advertising and many novel (but negative) findings in the field, especially with respect to the cigarette industry. The fundamental empirical problem is that advertising expenditures typically show little time-series variation within products. Consequently, controlling for unobserved product characteristics absorbs these relatively constant levels of advertising, and insufficient variation remains for the demand analysis, leaving statistically insignificant advertising estimates. Ignoring advertising, however, may itself have negative impacts, as I show in this paper that its omission can substantially distort estimation results even when instruments are used for all other endogenous variables.

*Washington University, Campus Box 1208, St. Louis, MO 63130-4899, USA. Tel: 314-935-5836. Fax: 314-935-4156. E-mail: moul@artsci.wustl.edu
The responsiveness of demand to price and other endogenously determined variables is central to any structural empirical research. This fact has driven the emphasis on identification and innovations in new instrumental variables. Few identification assumptions, however, can withstand the omission of an important endogenous variable. Consider advertising. The reduced form of the optimal level of advertising will depend upon a product’s cost and its competitive environment, both popular instruments for price. If advertising is important and is omitted from the regression, these instruments are invalid. The direction of this bias is straightforward if instruments are cost-based: equilibrium advertising is decreasing in cost, and this will bias price coefficients downward (away from zero). The bias from using instruments based upon the competitive environment is more ambiguous, but will also be downward for typical market shares in the logit model. I use Monte Carlo experiments of price-and-advertising setting duopolists to illustrate the potential extent of this problem. These experiments confirm that omitting advertising in demand biases price coefficients downward, i.e., estimated to be more elastic than they are. Typical applications of such demand estimates will lead researchers to overstate the extent of competitive forces in a market. While the magnitude of this bias depends upon advertising’s impact on demand, price coefficients are biased downward (away from zero) by 20% when implied advertising-sales ratios are only 6%, even though estimates indicate that advertising is statistically insignificant. An upward bias of similar magnitude occurs regarding the effect of observable characteristics. I conclude the price-setting experiments with a test (Creel, 2004) that allows one to apply Hausman-like tests to situations in which neither estimator is efficient, detecting advertising as an important omitted variable when advertising’s t-statistic fails to do so.

While prior work has shown that Monte Carlo experiments similar to those above can successfully capture the important aspects of price-setting, advertising is arguably a more complex and subtle topic. Advertising is often thought to have relatively long-lasting effects, and so firms may be forward-looking in their advertising choices. Prices obviously enter consumer decisions through a disutility of lost money, but the manner in which advertising affects consumer decisions remains uncertain. Signaling high unobserved quality, generating prestige, serving as a complement to the product itself, or informing consumers of the set of potential products are all possibilities. Given this rich uncertainty, it is an open question if reality can be adequately captured by simulations that assume advertising directly affects consumer utilities and firms that maximize static profits.
In the presence of these potential complications, the results from the simulations of price-and-advertising-setting firms would be substantially enhanced if they could be confirmed using real world data. An ideal empirical application would take this intuition to a price-setting industry to confirm the nature of these biases, but the problems identified by Schmalensee still hold for most applications that lack individual data. I instead try a different tack and consider the domestic theatrical demand for motion pictures. In the case of movie demand (where the price that consumers pay is fixed and decidedly not endogenous), a parallel to endogenous price exists with respect to the endogenous number of exhibiting theaters. Within studies of the motion picture industry, estimated values such as own-theater elasticities of demand are critical to questions of market conduct as well as any serious policy experiment. My goal then is to compare the biases that arise when advertising is omitted from both the Monte Carlo experiment and the real-world empirical application.

A Monte Carlo experiment of a theater-and-advertising industry shows that omitting advertising overstates the impact of theaters to such an extent that estimates suggest widespread violations of necessary supply-side conditions. This result may partially explain the finding of underscreening in foreign markets of Elberse and Eliashberg (2003). The experiments also reveal that omitting advertising substantially understates the impact of observable characteristics, perhaps explaining why documentation of such impacts has been so difficult in structural work. Using data from the movie industry, actual estimates largely match the predictions of a Monte Carlo experiment appropriate for the industry. When advertising is omitted from estimation using the real-world data, the average responsiveness of demand to the number of exhibiting theaters is overstated by 25%. This upward bias drives 15% of the observations to theoretically implausible levels, while the estimates that include advertising are all well within acceptable bounds. Omitting advertising also makes cast appeal appear to be a statistically and economically unimportant variable, having less than 20% of the impact as estimated when advertising is included. Curiously, the estimated impact of director appeal shows little difference when advertising is excluded, suggesting that studios do not base their decisions on this variable despite its strong estimated impact on consumer utility. While useful, Monte Carlo experiments remain an imperfect substitute for empirical analysis of actual data.

The structure of this paper is as follows. Section 1 briefly reviews the relevant literature in advertising and movie economics. Section 2 uses a logit model to illustrate the source of bias when an endogenous variable such as advertising is omitted from IV estimation of discrete-
choice models and then shows Monte Carlo results at varying levels of advertising importance. Section 3 considers important characteristics of theatrical motion picture distribution and lays out the structure of the simulations, while Section 4 explains the logit extension that I use to estimate movie demand. Section 5 introduces the data, and explicit identification assumptions and estimation algorithms are laid out in Section 6. Section 7 presents the results of the Monte Carlo experiments and the movie demand estimation, and Section 8 concludes.

1 Literature Review

The study of advertising has largely fallen to marketing analysts, but the topic of how advertising affects consumer behavior has received some attention from economists. Ackerberg (2001) examines consumer-level data on yogurt purchases and television-watching to show that advertising is better characterized as being informative rather than convincing consumers that products impart higher prestige. Ackerberg (2003) uses the same data to quantify the value of this information as consumers learn through experience what products they like. Horstman and MacDonald (2003) examine a panel data set of CD players for ten years after their introduction to test (and reject) time-series implications of the simple explanation of advertising as a signal of quality. Shum (2004) combines scanner data of ready-to-eat cereal purchases with quarterly national brand-level advertising expenditures and store-level promotional variables to show that advertising works to counteract brand loyalty and is therefore pro-competitive. Research that is akin to the pre-Schmalensee literature includes Scott Morton (2000), who looks at advertising as a potential barrier to entry in pharmaceuticals but finds no statistically significant impact of advertising.

The idiosyncrasies of the movie industry have attracted research attention along other dimensions as well. Einav (2005) decomposes the observed seasonality of admissions into underlying seasonality and the amplification that arises when distributors time their best releases for high-demand periods. Moul (2006b) uses a similar model of demand to measure the speed and magnitude of information transmission (i.e., word of mouth) among consumers. Corts (2001) examines how the several divisions of a single studio coordinate their release decisions and concludes that these divisions behave like an integrated whole. Moul (2001) looks at the introduction of synchronous sound recording technology to find

\[^1\] Dube, Hitsch, and Manchanda (2004) offer a thorough review of this literature as well as an interesting look at the dynamics of advertising for frozen entrees.
evidence consistent with movie producers making higher quality films as they became more experienced. Emphasizing exhibition, Davis (2005) uses price variation between theaters to consider how cross-price elasticities change with spatially differentiated products. There has also been research that is more particular to the industry. Ravid (1999) considers the movie’s entire revenue stream and measures factors of profitability, finding that G-rated movies are historically the most profitable and that stars (whether cast or director) fully capture their own rents. Goettler and Leslie (2004) look at various explanations of why studios jointly finance movies, concluding that such co-financing alleviates coordination problems in release scheduling. Finally, Moul (2006a) examines aggregated payments from theaters to distributors in the context of demand estimates and models of exhibition-clearing, concluding that the level of these payments is best explained through collusion among studios.

2 Model

2.1 Logit and omitted advertising bias

Like much of the literature, I model demand using the discrete-choice techniques introduced by McFadden (1974) and applied to markets of differentiated products by Berry (1994). I will make use of the simple logit model for my Monte Carlo experiments that capture traditional price-setting markets. My Monte Carlo experiments and actual demand estimation of the theatrical movie industry will allow for consumer heterogeneity along product characteristics with its resulting segmentation.

Discrete-choice models of demand are based upon a population of M consumers making choices from a set of J products each period. Including the outside option (i.e., the option of choosing no product at all for that period) within the choice set, a consumer will purchase a product if her valuation from such a purchase exceeds the valuation from any other option available that period. Considering the entire market at time t, individual i’s optimal choice j and valuation for choice j can be respectively expressed as

\[ j \text{ s.t. } U_{itj} \geq U_{itk}, \forall k \]  

\[ U_{itj} = \delta_{tj} + \epsilon_{itj} \]

where \( \delta_{tj} \) is the mean valuation of consumers considering product j, and \( \epsilon_{itj} \) is the individual-specific deviation about that mean. Ties between these consumer-specific expectations occur with zero probability. Because this valuation is an ordinal concept, normalizing the mean
valuation of purchasing nothing to zero \( \delta_{t0} = 0 \ \forall \ t \) is necessary for identification. The empirical goal is then to explain variation in quantities (scaled as purchase probabilities \( s = Q/M \)) with variables explaining mean utilities \( \delta \).

Assume that mean utility takes the following form (market subscripts have been suppressed):

\[
\delta_j = \beta_0 + \beta_1 x_j + \alpha p_j + \beta_{ADS} \ln(A_j) + \xi_j
\]

where \( x \) denotes an observed exogenous product characteristic, \( p \) denotes price and \( A \) advertising (both endogenously determined), and \( \xi \) denotes the mean valuation of characteristics unobserved by the econometrician. If one ignores advertising and lets \( \Theta \) denote the sum of advertising’s impact and \( \xi \), it appears that a single instrument is necessary for the model to be just identified.

I first consider what happens if one uses a cost-based instrument \( w \) for price but ignores advertising. For simplicity assume a linear cost relationship: \( c_j = w_j \gamma + \omega_j \). The instrument matrix is then \( Z = [x \ w] \). While the normal identification assumption \( (E(Z'\xi) = 0) \) continues to hold, consistent IV estimates will require that \( E(Z'\ln(A)) = 0 \) as well. An application of Cramer’s Rule to the firm’s profit function \( \Pi_j = (p_j - c_j)Ms_j - \tau A_j \) and appropriate advertising first-order condition reveals that \( \frac{dA_j}{dw_j} < 0 \). Furthermore, \( \frac{dA_j}{dx_j} > 0 \) so long as demand is convex, which is true for the logit whenever purchase-probabilities are less than half. These equilibrium relationships imply that omitting advertising when using cost-based instruments will bias the price coefficient away from zero, overstating demand’s price elasticities. For reasonable market outcomes, there is also a likely upward bias on the coefficients of product characteristics. The magnitudes of these biases will obviously depend upon the magnitude of \( \beta_{ADS} \).

Now consider instrumenting for price using a product’s competitive environment \( \bar{x} \) but again ignoring advertising. A similar application of Cramer’s Rule shows that, so long as demand is convex, \( \frac{dA_j}{dx_j} < 0 \) and \( \frac{dA_j}{dw_j} > 0 \). Omitting advertising when using competitive-environment instruments will typically bias the price coefficient away from zero and the product characteristic coefficient upward, though it is easy to envision scenarios where concave regions of demand would reverse these predictions. As the model of movie demand that I estimate below is identified entirely with competitive environment instruments (as is common in the literature), the direction in which my endogenous variable is biased can only
be answered with the data.

### 2.2 Monte Carlo simulations

Table 1 displays Monte Carlo results for estimation of the logit model with and without the inclusion of log-advertising. While some parameters have been altered, this is essentially an extension of the experiment presented in Berry (1994). Each simulated sample consists of 500 duopoly markets and the utility of each consumer in each market is given by

$$u_{ij} = \beta_0 + \beta_x x_j - \alpha p_j + \beta_{Ads} \ln(A_j) + \xi_j + \varepsilon_{ij}$$

with $\varepsilon_{ij}$ being the appropriate logit error.\(^2\) The utility of the outside option is given by $u_{i0} = \varepsilon_{i0}$, where $\varepsilon_{i0}$ is also drawn from the logit distribution. I added a second cost variable to Berry’s original so that my estimation would (like his) have a single overidentifying restriction. The log of marginal cost is given by

$$\ln(c_j) = \gamma_0 + \gamma_1 x_{1j} + \gamma_2 w_{1j} + \gamma_3 w_{2j} + \sigma_c \xi_j + \sigma_\omega \omega_j$$

Firms maximize profits of $\Pi_j = (p_j - c_j) M s(\delta_j, \delta_{-j}) - \tau A_j$. Notationally, $x$ and $w$ denote exogenous product characteristics and costs, $p$ and $A$ denote endogenous price and advertising, $\xi$ captures the mean valuation to consumers of product characteristics unobserved by the econometrician, and $\sigma_c \xi_j + \sigma_\omega \omega_j$ captures costs unobserved by the econometrician, some of which may be related to unobserved product characteristics.

Exogenous variables ($x$, $\xi$, $w$, $\omega$) are created as independent standard normal random variables. The parameters to be estimated are $\beta_0$, $\beta_x$, $\alpha$, and $\beta_{Ads}$, while $\sigma_c$ and $\sigma_\omega$ capture the variance of the unobserved variables. Parameter values were chosen by experiment to yield sufficient variation in purchase-probabilities without having too many markets in which one firm’s purchase-probability went to zero. The chosen values common to all regressions are $\beta_x = 2$, $\alpha = 1$, $\gamma_0 = 1$, $\gamma_1 = 0.5$, $\tau = 100$, and $\gamma_2 = \gamma_3 = \sigma_c = \sigma_\omega = 0.25$. The presence of advertising necessitates choosing a population size, and I assume $M = 1,000,000$. As the feature of interest is the extent of bias that arises under varying levels of advertising importance, I separately consider the values $\beta_{Ads} = 0.1$, 0.25, and 0.5. Intercept parameter values vary so that market shares are roughly equal across simulations.

For each market, I first calculate the equilibrium values of price, advertising, and purchase-probabilities by jointly solving each duopoly market’s four first order conditions. I assume

\(^2\)Nevo (2001) essentially adds brand fixed effects and random coefficients to this demand specification.
that these variables and the exogenous variables \( x \) and \( w \) are observed, but that \( \xi \) and \( \omega \) are unobserved by the econometrician. As usual, the mean utility of good \( j \) can be found as
\[
\delta_j = \ln(s_j) - \ln(s_0),
\]
where \( s_0 \) denotes the probability that the typical consumer purchases nothing.

The first column in each of the three scenarios of Table 1 presents Monte Carlo estimation results from samples of markets when log-advertising is included as an endogenous regressor. The second column shows results when log-advertising is excluded. Scenarios are ordered from advertising being relatively unimportant to important. I have also included the mean firm-level advertising-to-sales ratio implied by the simulations beneath the first column to ease comparison to actual industries.

Several results are apparent. First, the inclusion of log-advertising yields consistent but relatively imprecise estimates. Only when \( \beta_{Ads} \geq 0.4 \) (a value that generates the relatively high \( A/S \approx 0.10 \)) is the advertising’s coefficient significantly greater than zero. Industries with such high ratios such as distilled and blended liquor (SIC #2085) at 10.4% and soaps and detergents (SIC #2840) at 10.0% have traditionally been the centers of economists’ advertising studies.\(^3\) For all but the lowest magnitude of advertising’s impact, however, ignoring advertising systematically overstates the values of the price and product characteristic coefficient. Whenever \( \beta_{Ads} \geq 0.25 \) (a threshold consistent with the more common \( A/S \approx 0.06 \)), estimates that ignore advertising falsely reject the true parameter values. Therefore demand analyses of industries such as books (SIC #2731) at 7.3% and sugar and confectionery products (SIC # 2060) at 5.9% could be affected by omitting advertising. The bias is even noticeable, though not highly significant, at \( \beta_{Ads} = 0.1 \) when the implied \( A/S \approx 0.025 \) is quite low. A large number of industries where advertising is rarely examined are therefore conceivably affected, though the magnitude of these biases may well be slight.

Of course, economists typically lack the knowledge of parameter’s true values in empirical applications. If the individual t-statistic on advertising is not useful for determining whether advertising should be included, what statistical tests would be useful? One candidate test might be the J-statistic that coincides with the minimized GMM objective function when the weighting matrix is \( (Z\hat{\Omega}Z)^{-1} \). Under the null hypothesis of a correctly specified model, this statistic should be distributed \( \chi^2 \) with degrees of freedom equal to the number of overidentifying restrictions. Unfortunately, this test is so broad that rejection fails to be constructive.\(^4\) An alternative test that follows the intuition of Hausman (1978) has recently

\(^3\)All ad-dollar to sales ratios are taken from Schonfield and Associates, Inc. (2005).

\(^4\)Indeed, it is quite uncommon for models that estimate random coefficients not to be rejected using
been proposed by Creel (2004). While Hausman’s insight of comparing an always consistent estimator to a potentially inconsistent estimator applies only to situations when one of the estimators is efficient, Creel’s test allows for the comparison of two inefficient estimators as are typically available in GMM estimation. The approach is to include the two estimators in a single supermatrix and then empirically calculate covariances of the two estimators. Table 1 includes both J-statistics and these pseudo-Hausman statistics. The first Hausman statistic uses the displayed point estimates and is inefficient, while the second re-estimates parameters using the efficient weighting matrix and is thus efficient. No test statistic can reject the null hypothesis that advertising has no impact on impact for the lowest level of advertising, but the efficient Hausman statistic can reject that null hypothesis for the intermediate level of advertising with 10% confidence. Recalling that advertising’s t-statistic can reject the unimportant-advertising null for the high level of advertising, it is unsurprising that both pseudo-Hausman test statistics also reject the null for that scenario.

Structural point estimates themselves are rarely the object of interest, and Table 2 displays true and estimated elasticities, derivatives, and Lerner Indices (i.e., $P_{x-c}P$) under each of the three advertising levels. Despite the inability to statistically reject the unimportant-advertising hypothesis for low levels of advertising, the mean estimated own-price and cross-price elasticities when advertising is omitted overstate their true counterparts by about 8%. The implied Lerner Index is understated by about the same amount, though in this case the bias is about the same magnitude as when advertising is included (albeit in the opposite direction). Similarly, the mean own-characteristic derivative $\frac{ds}{dX}$ is overstated about 8% when advertising is omitted. 5 The magnitudes of bias for these implications under the intermediate level of advertising are substantially greater. Own and cross price elasticities as well as own-characteristic derivatives are overstated by 25% when advertising is omitted; furthermore, firm Lerner Indices are understated by 20%. These substantial biases occur when the advertising t-statistic was only 1.2 and emphasize that caution must be taken when deciding to exclude advertising to reduce collinearity and enhance precision. Implications under the high-advertising case are striking mostly for their magnitudes. Price elasticities and characteristic derivatives are overstated by over 40% and Lerner Indices are understated by over 30% when advertising is excluded. The concern of ignoring advertising understating market power appears to be valid and important at lower levels than one might have thought a

5 Derivatives rather than elasticities are employed because characteristics are drawn from a mean-zero distribution.
Monte Carlo experiments can shed light on the possible problems in estimation, but these results are by construction driven entirely by the assumptions of the simulated model. The appropriateness of these assumptions is an obvious bridge to whether these problems are relevant in practice. The strong assumptions greatly simplify solving for profit-maximizing outcomes, but one has no way to gauge whether these assumptions are particularly onerous without an empirical comparison. Firms behave as if they believe that the impact of ads lasts beyond a single period; how much is lost and distorted by assuming that firms choose advertising levels to maximize static profits? An ongoing question in economics regards the best way to model advertising’s effect on consumer behavior; what is distorted by assuming that advertising appears as a direct component of utility rather than through any of the possible information avenues that might be considered more realistic? Given these uncertainties, there are rightly more questions regarding how advertising should be incorporated than there are regarding price; whether actual data and subsequent estimates will behave as simulations predict is an open question. I now turn my attention to considering whether an experiment appropriate for the theatrical movie industry generates results similar to those that are found from demand estimation using actual data.

3 Theatrical Movie Distribution

3.1 Reality

Theatrical distribution entails two primary costs: the physical replication of the reels of film and their transportation to theaters (i.e., prints and shipping), and the purchase of advertising. Most contemporary estimates suggest that it costs about $2000 to copy and ship each additional print. Whether this amount is substantial depends upon how long a print is used. Advertising, however, is clearly an important expenditure. Industry reports suggest that it is now common for a studio to spend an amount equal to 30-50% of a movie’s production cost on domestic advertising and promotion. Most of these advertising expenditures (about 60%) occurs on television (network, spot/local, cable, and syndicated),

---

6 These possible interpretations range from the signaling models of Nelso (1974) and Milgrom and Roberts (1986) to the complements in consumption model of Becker and Murphy (1993).

7 See Moul and Shugan (2005) for a far more thorough examination of the institutions of and factors determining the theatrical launch of movies.

8 The cost translates to the box office receipts of about 350 admissions. The median per-theater weekly admissions of my sample is about 420.
with the majority of the remainder consisting of newspaper ads.\textsuperscript{9} Information on advertising and marketing budgets is occasionally released, but only aggregated over a movie’s theatrical run. For example, reports/rumors suggested that New Line spent $40M promoting “The Spy Who Shagged Me” in the summer of 1999.\textsuperscript{10} For this expenditure New Line (owned by Time-Warner) probably received about $110M in rental payments from exhibitors. Given that available data conceivably allow one to estimate weekly demand, these aggregated figures are of limited use even when they can be found.

Distributor-exhibitor negotiations primarily specify payments to the distributor through a rental rate (i.e., the percentage of box office receipts that the exhibitor cedes to the distributor). While contracts apply for several weeks (four weeks was common for my sample’s time period), negotiated terms are allowed to vary across weeks, and declining rental rates are the norm.\textsuperscript{11} An opening week rental rate of 70% is common for highly anticipated movies, and rental rates will often reach 30% (occasionally lower) as a movie plays out.

A common question in the theatrical movie sector is why admissions prices to consumers are constant, both across movies at a given theater and over time for a given movie. While Orbach and Einav (2005) offer some insights on the matter, no definitive answer has yet been reached. This constancy of prices has now been written into the standard distributor-exhibitor contract (exhibitors are allowed a single discount night during the week). The fixed price is unsettling from an economic perspective, but quite useful from an econometric one. I will throughout assume that admissions prices are exogenous to both distributors and exhibitors on a weekly basis.

\subsection*{3.2 Monte Carlo simulations}

There are four substantive ways in which a Monte Carlo of the theatrical movie industry will differ from the stripped-down price-setting industry described earlier: the importance of vertical relationships, the primacy and capacity implications of the number of retail outlets (i.e., theaters) as the foremost market-clearing variable, the possibility of a studio choosing not to advertise in a given market, and the absence of any cost-shifting instruments. I address each in turn.

\textsuperscript{9} Figures from January to December 2000, Competitive Media Reports (2001)
\textsuperscript{10} For comparison, production costs were estimated to be $33M (www.imdb.com).
\textsuperscript{11} The declining length of movies’ theatrical runs as consumers are shifted to the opening week has generated a new contract that specifies a single rental rate that is constant over the run. These contracts were first introduced in 2001 and appear to be replacing the former contract structure. Filson, Besocke, and Switzer (2004) examine explanations of these contracts in the context of risk-sharing.
As stated above, distributors and exhibitors typically contract using revenue-sharing. This can be most readily modeled as in Moul (2006a) through distributors setting rental rates and then an exhibition sector selecting what movies to show until an equilibrium condition is satisfied. The primary problem with this approach is the number of necessary assumptions such as the extent that theaters “steal” business from one another and the extent to which the number of exhibiting theaters can vary with entry. Rather than arbitrarily choose among these different assumptions, I simplify the matter by assuming a vertically integrated industry: Studios receive all box office receipts and rent theaters and buy advertising, both at constant marginal cost. This assumption generates the same reduced form relationships that are documented in the data (e.g., the number of exhibiting theaters increases with favorable movie characteristics such as cast appeal).

This simplistic vertically integrated industry has theoretical implications regarding the interpretations of elasticities. Similar to the Dorfman-Steiner condition which relates profit-maximizing price and advertising elasticities to advertising-to-sales ratios, the distributor’s first-order condition regarding the number of theaters can be rearranged to show that a movie’s own-theater elasticity will equal its theater expense-to-sales ratio. A similar result holds with respect to advertising. Neither elasticity nor their sum should therefore exceed one in any observations. This same boundary can be derived under the more realistic vertically disintegrated model described above, but there it stems from necessary existence conditions for an equilibrium in exhibition.

The number of exhibiting theaters is similar to price in that both presumably affect consumer utility directly. This can most readily by seen as the number of theaters capturing the distance that a consumer must travel to see a movie. Theaters, however, also have a relationship with capacity ... given fixed prices, the most popular movie showing at a single theater will have a limited box office in any given week. I incorporate this capacity interpretation as I model advertising, specifically allowing advertising to affect consumer utility only through an interaction with the number of exhibiting theaters. I use both this interaction specification and advertising alone in separate regressions in my empirics to the actual data.  

It is common for movies to have no advertising. I therefore include ln(1 + A) rather than ln(A) in my specifications. I ensure that advertising has an everywhere positive effect by interacting log-advertising with ln(1 + T) so that small values of T do not lead to advertising

12 Allowing advertising to affect mean utility independently as well as through the interaction generated too much multicollinearity for precise estimates.
having a negative impact on utility. To determine optimal studio behavior, I attempt to solve the theater and advertising first-order conditions under the four possible scenarios: 

\[ \min(A_1, A_2) > 0; \ A_1 = 0, \ A_2 > 0; \ A_1 > 0, \ A_2 = 0; \text{ and } A_1 = A_2 = 0. \]

The equilibrium outcome is then determined by a Nash condition on the implied profits. Finally, there are few cost factors that would credibly affect movie exhibition in ways that vary across movies. Identification must consequently come entirely from variations in the competitive environment. I therefore expand the number of continuous observed exogenous variables from one to three.

Preliminary efforts to mimic the earlier Monte Carlo experiments under these new conditions showed the multicollinearity between characteristics, theaters, and advertising is severe. Reasonably precise estimates require two changes: a larger number of duopoly markets and a more powerful set of instruments. While the former is trivial to fix (aside from computation time), I address the latter by expanding the model from a simple logit to a one-level nested logit (\( \sigma \) lies on the unit interval, with \( \sigma = 0 \) denoting no segmentation). The larger number of observations pins down the impact of exogenous variables, while the new instruments provided by the nested logit pin down the effects of theaters and advertising. While the nesting in a simulation need have no real-world counterpart, I find it useful to think of the nest as capturing the segmentation between Family (\( Fam = 1 \)) and non-Family movies that is driven by the presence or absence of children in a household. Whether a movie is Family or not also affects mean utility and is therefore added as a fourth exogenous characteristic.

\[
\delta_{tj}(\sigma) = \kappa_0 + \kappa_1 X_{tj1} + \kappa_2 X_{tj2} + \kappa_3 X_{tj3} + \gamma Fam_{tj} \\
+ \alpha \ln(T_{tj}) + \beta \ln(1 + T_{tj}) \ln(1 + A_{tj}) + \xi_{tj}
\]  

(2)

Notation is similar to before, but \( T \) now denotes the number of exhibiting theaters.

Price-setting models are relatively self-controlled. Firms selling goods with higher innate demand will charge higher prices, and this will mitigate the impact on sales of outliers. The movie example above, though, has no such mechanism: Studios with movies facing higher innate demand will exhibit them at more theaters and secure more advertising, exacerbating outliers. I will therefore consider a much smaller range of exogenous variables than I did in the price-setting case.

\[^{13}\text{In situations in which not advertising is optimal, it is common for no solution to exist to the full set of first order conditions. In these cases, answers minimize the sum of squared errors from the logged conditions, e.g., } e_1 = \ln(MB^{Ads}) - \ln(MC^{Ads}).\]
4 Empirical Model and PD GEV

The one-level nested logit described above is sufficient for the Monte Carlo exercise, but actual empirical work may require even richer demand specifications. To allow for this richness, I follow Bresnahan, Stern, and Trajtenberg (1997) by using a principles of differentiation generalized extreme value (PD GEV) model. The PD GEV has a parameterization similar to the multi-level nested logit without the restrictions that are generated by (typically arbitrary) ordering of the nests of the consumer decision process. The primary advantage of this specification over estimation using random coefficients is its relative robustness without imposing a supply-side model (e.g., joint pricing equation) and in the absence of microlevel data such as demographics.

Since Quandt’s (1956) work, economists have known that the distribution of the individual-specific deviations $\varepsilon$ determines much of the nature of aggregate demand and the resulting substitution patterns. Deviations that are independently drawn (such as those from the Type 2 extreme value distribution that generates logit probabilities) yield the familiar Independence of Irrelevant Alternatives property. Allowing such deviations to be independent across consumers but correlated for a given consumer across products can yield more intuitive substitution patterns. The nested multinomial logit and PD GEV are both attempts to relax this independence assumption and generate more plausible substitution among products.

McFadden (1978) formalizes a framework to allow for this sort of correlation in his generalized extreme value model. In this application, individual-specific deviations are realized from a distribution that depends upon segmentation parameters $\rho$ that lie on the unit interval. The econometrician chooses the potential pattern of segmentation to be estimated by selecting principles of differentiation, and each such principle is assigned its own $\rho$. As any dimension’s $\rho$ approaches 1, consumers give less consideration to substituting to similar products along that dimension. Conversely, as any $\rho$ approaches 0, consumers consider only similar products along that dimension when substituting among competing products. In cases of a single dimension, the PD GEV simplifies to the nested logit where $\rho = 1 - \sigma$.

I focus upon two such dimensions for substitution within the PD GEV framework. They roughly correspond to the questions, Should I see any movie?, and Should I see a family movie?14 These dimensions should capture likely effects of consumer heterogeneity in competition between movies and the non-movie option and between movies of different audience

---

14I also considered a Principle of Differentiation for action vs non-action movies. All estimates indicated that $\rho^A \approx 1$, that is, segmentation along this dimension is unimportant.
appeal. The formal specification is presented in Appendix 1.

With this framework established to allow for more flexible substitution patterns, I return to the components of the mean utility $\delta$. Like the Monte Carlo specification, I assume that consumers have preferences about the number of exhibiting theaters and advertising. Because one expects advertising to have a much larger effect on demand if the movie is available nationwide instead of showing only at a few theaters, I interact theaters with advertising. I also include advertising by itself as a robustness check.

Perhaps the most obvious exogenous characteristic of theatrical movie demand is the way in which demand decays over time. I follow Einav (2005) and capture this decay with a movie’s age, specifically the quadratic of $\ln(Age)$. I use three additional exogenous variables that vary by week and by movie. The first is a binary variable for whether the movie faces an abbreviated week. While most movies are released on Friday, Wednesday openings are also common around certain holidays (e.g., Thanksgiving). As the revenue data are collected from Friday to Thursday, such an abbreviated week is likely to have a large and negative impact on admissions. I also account for the announcement of the six most prominent Academy Award (Oscar) nominations and winners by including two indicator variables that denote the prior announcement of each.

The richness of my data allows me to utilize movie fixed effects for all movies observed more than three times. The impacts of movie-specific characteristics such as genres and the appeal of the cast and director can then be recovered by regressing the estimated fixed effects upon these variables (Nevo, 2001). I proxy for seasonal differences in demand by including binary variables for each month from February to December and for the eleven major holidays. Finally, I also include a measure of the unemployment rate and a crude measure of average admissions price.

Below is the specification for mean utility and the result of inverting the observed purchase probabilities.

$$
\delta_{tj}(\rho) = \alpha \ln(T_{tj}) + \beta g(A_{tj}) + \gamma_1 \ln(Age_{tj}) + \\
\gamma_2 \ln(Age_{tj})^2 + X_{tj}\kappa + A_j + W_t\psi + \xi_{tj}
$$

(3)

The advertising function $g(\bullet)$ is then either $\ln(1 + A_{tj})$ or $\ln(1 + T_{tj}) \ln(1 + A_{tj})$. The number of theaters exhibiting movie j in week t is denoted by $T_{tj}$, while the advertising of movie j

---

15 Movies that I observe three or fewer times have the product characteristics (including intercept) of the second-stage regression included in lieu of a movie fixed effect.

at week \( t \) is denoted by \( A_{tj} \). The particular functional forms were chosen after preliminary polynomial results indicated increasing and concave relationships. The vector \( X_{tj} \) includes whether the movie’s week is abbreviated and whether the movie has received any Oscar effect by that week, movie fixed effects are captured by \( \Lambda \), and the vector \( W_t \) includes week-specific effects such as seasonality and macroeconomic conditions.

5 Data

5.1 Quantities, screens, and advertising

I assemble the movie data from a number of independent and publicly available sources. Like much of the recent empirical work in the movie-economics literature, I turn to Variety’s listings for revenue measures, from which I gather each week’s movies. Each week Variety magazine publishes a ranking of the prior week’s domestic (U.S. and Canada) box-office revenues for the fifty highest grossing movies. These figures are compiled by Variety and EDI from studio revenue reports. The week of August 24, 1990, was the first where this direct reporting by the studios fully supplanted the sampling techniques that Variety had previously used, and it is with this week that I begin my sample.\(^{17}\) In addition to revenues and rankings, the Variety listings include each movie’s number of exhibiting theaters. The sample ends the week of December 27, 1996, spanning 332 weeks and yielding 16600 observations.

While revenues are broadly consistent with quantities and sometimes instructive on their own, demand analysis requires quantities and therefore a measure of price. As explained above, the admissions price does not vary substantially across movies at a given theater, and so I turn to the National Association of Theatre Owners (NATO) annual average admissions price. As Davis (2002) notes, virtually no one pays this average price, but it is arguably the best available when considering the domestic market as a whole.\(^{18}\) I use a linear reconstruction from these eight observations to approximate the market’s weekly admissions price. Quantities are derived from the Variety revenues and this nominal price. Additionally, the transformation of quantities into purchase-probabilities \((s = \frac{Q}{M})\) requires a measure of market size. I use the combined population of the U.S. and Canada for the first week of July each year and linearly reconstruct weekly counterparts.\(^{19,20}\) Conditional purchase-probabilities are the

\(^{17}\)Throughout my data, weeks begin on Fridays and end on Thursdays. Dates refer to the starting Friday.
\(^{18}\)See Switzer (2004) for a microlevel examination of the potential biases that this can introduce, especially regarding how the quantities for G-rated movies are undercounted.
\(^{19}\)U.S. Census Bureau (2001).
likelihood that a consumer who saw any movie saw a particular movie ($s^M = Q_{t,j} / \sum Q_{t,k}$) and that a consumer who saw any movie of specific family-ness (e.g., a non-Family movie) saw a particular movie of that same family status ($s^F = Q_{t,j} / \sum_{Fa(k)=Fa(j)} Q_{t,k}$).

Other than the number of exhibiting theaters, the dominant component of demand that varies across both movies and weeks is advertising. While total advertising budgets are occasionally released after movies’ runs are complete, I could find no broadly available measures of a movie’s weekly advertising expenditures. The measure I consider instead is newspaper advertising from that week’s Sunday *Chicago Tribune*. Sunday papers receive a disproportionately large level of newspaper advertising.\(^{21}\) Chicago is a sufficiently large market that distributors themselves purchase movies’ display ads. This avoids the problem of separating national distributor ads from local exhibitor ads. I measure these ads in column-inches: the product of an ad’s width in columns and its height in inches. Such a measure will be a good proxy for weekly advertising expenditures if distributors spend a constant proportion of total advertising expenditures at the *Tribune*. Given the difficulty of verifying this assumption, I will consider reduced form estimates to confirm its appropriateness before turning to the demand estimation.

### 5.2 Exogenous variables

Even though I am exploiting movie fixed effects, movie-specific characteristics are useful to get a sense of what drives these fixed effects and are essential to create sufficiently powerful instruments. Appendix 2 describes in detail my measures of cast and director appeal, but both essentially make use of the box office history of movies within the prior five years (in billions of dollars). Demand estimation using the PD GEV requires some measure of a movie’s family status, and I use a movie’s action status as a regressor and source of instruments as well. A movie falls within the family genre if it is rated either G or PG, and a movie is categorized as within the action genre if it is listed as Action or Adventure in the Internet Movie Database (www.imdb.com).

Instruments that will exploit portfolio effects demand some measure of ownership, even though this ownership does not appear in the mean utility specification or the second-stage regression. I construct ten binary variables to indicate whether a movie is released by one of the notable distributors: Disney, MGM, Paramount, Sony, Twentieth Century Fox,

\(^{21}\) Friday advertisements are about as large, but similar data from the Friday *New York Times* proved too unrepresentative for use.
Universal, Warner Bros., Orion, Miramax, and Savoy King. Following Corts (2001), all subsidiaries are collected under a single variable (e.g., Disney encompasses Buena Vista, Hollywood Pictures, and Touchstone Pictures, as well as Disney Pictures).

To this point, the only variables I have described that vary across movies and across weeks are theaters and advertising, but I also consider four exogenous variables of this type. My binary variable $Ab$ equals 1 for the opening week of any movie released on a day other than Friday, and zero otherwise. Academy Award announcements also fall in this category. I use the announcements to create the binary variables $OscNom$ and $OscWin$. These variables equal one if the movie in question prior to that week has been nominated or won in any of the six major categories (Best Picture, Best Director, Best Actor, Best Actress, Best Supporting Actor, and Best Supporting Actress). I conclude by defining $Age$ as the number of weeks since the movie’s initial entry into the Top 50 listings that the movie has appeared in those rankings, so that a movie in its opening week has $Age = 1$.

5.3 Describing the demand data

The descriptive statistics of these variables can be found in Table 3. Perhaps most striking from the table are the large variances. *Jurassic Park* in its opening week of June 11, 1993, had admissions of over 19 million, while *Life and Nothing But* in the week of December 27, 1990, had 593 admissions. That same disparity is apparent with respect to the number of exhibiting theaters and advertising. *Batman Returns* opened at 3700 theaters the week of June 19, 1992, while many movies showed at a single theater. *Toy Story, The Godfather Part III*, and *101 Dalmatians* (1995) all had over 120 column-inches during their opening weeks, even though 2/3 of the observations had no such advertising. Given the large number of the observations with no advertising in the *Tribune*, I provide statistics conditional upon positive advertising as well as unconditional statistics. This concentration of admissions, theaters, and advertising at the top suggests that the convexity in revenues found by earlier studies applies to other aspects of the industry as well. This convexity poses a severe challenge to estimation since the model must explain both the bulk of observations and the outliers from which most revenues come.

The binary variables capturing Oscar announcements and abbreviated weeks are straightforward, though the small number of observations for Oscar wins (304) and abbreviated weeks

---

22 My Oscar variables thus capture the impact of the Oscar announcement rather than the movie being good enough to warrant an Oscar.
(81) suggests that it might be difficult to obtain precise estimates of their impacts. The typical movie had been in theaters for 7 weeks, but a few outliers running for longer inflate the mean. From the summaries of variables that capture movie-specific variables, one can see that about 5% of movies faced an abbreviated opening week. Noting that genres are not exclusive, 20% of the sample’s movies are categorized as family movies and 25% as action movies.

Patterns similar to those observed in admissions, theaters, and advertising are also apparent when attention turns to the cast appeal variable. Of the observations, 44% have no cast appeal. As with the advertising variables, I also provide statistics conditional upon an observation’s movie having positive cast appeal. Top actors include Arnold Schwarzenegger, Julia Roberts, Paul Hogan (Crocodile Dundee), Tom Cruise, and Macauley Culkin (Home Alone). The director appeal variable is also heavily affected by outliers, with 41% of observations having no director appeal. Top directors include Steven Spielberg, Chris Columbus (Mrs. Doubtfire), Robert Zemeckis, and Tim Burton. The plausibility of the top directors, especially when compared to the outliers of the cast appeal proxy, suggests that directors’ past experience may be more important an explanatory variable than the starring cast.

I more systematically consider the potentially market-clearing variables in Table 4, where simple correlations between the transformations of the six (presumably) endogenous variables and age are presented. The near-unity correlation between $\ln(s)$ and $\ln(s^M)$ is not unexpected, but it does suggest that precision when estimating $\rho^M$ is likely to be relatively poor. The importance of age in explaining market outcomes is clear from the large negative correlations with all variables except $\ln(T)$, an exception that meshes with the declining rental rate structure that characterized distributor-exhibitor contracts during the period. The general extent of multicollinearity is also apparent and will presumably work against precise estimation of the model.

Given the centrality of advertising, it makes sense to examine the plausibility of using newspaper advertising as a measure of total ad expenditures. As mentioned above, almost two-thirds of the sample receives no advertising in the Chicago Tribune, suggesting that there exist a minimum level of innate demand that is required for advertising to be worthwhile. For example, the marginal benefit of advertising may be less than the marginal cost of advertising for all relevant theater and advertising levels. To address this concern, I estimate a Heckman two-step regression for the binary issue of whether a movie is advertised in a given week and the continuous variable of the amount of advertising (conditional upon being
advertised). Table 5 displays least squares and Heckman results using column-inches in the Sunday *Chicago Tribune* as the measure for advertising. Conclusions are similar across the four regressions, and the Heckman results are reasonable and internally consistent. Both the latent profit difference and observed advertising are decreasing in a movie’s age, but increasing in a movie’s cast and director appeal. Receiving an Oscar nomination in one of the main six categories increases both dependent variables, but actually winning an Oscar does not. Family movies generate (marginally) higher dependent variables than non-family movies, but there is no comparable finding for action movies. In sum, there is support for Tribune Sunday column-inches being an adequate proxy for total advertising.

6 Identification, Instruments, and Estimation

The primary concern in estimating demand is that variables that are omitted by the econometrician will be substantially correlated with an included variable. In the above framework, the segmentation parameters $\rho$ as well as the endogenous variables of theaters and advertising warrant special attention. I address the identification of all with instruments that capture a product’s competitive environment. My demand estimates based upon fixed effects then rely upon the following two identifying assumptions:

1) The individual’s idiosyncratic addition to expected utility $\varepsilon_{itj}$ is drawn from a distribution characterized by McFadden’s (1978) assumptions and my own choice of segmentation dimensions.

2) The movie-and-week-specific unobserved component of utility $\xi$ is uncorrelated with product characteristics, including the strength of that movie’s competitive environment.

Assumption 1 is the typical functional form assumption. Assumption 2 supports the use of competitive environment variables as instruments. The implicit exogeneity of movie release dates upon which this assumption relies seems at odds with the release-timing game played by studios, but this is lessened greatly when one recalls that movie fixed effects will control for the most obvious examples (e.g., highly anticipated movies released in early summer and at the end-of-year holidays).

Following the literature and especially Bresnahan, Stern, and Trajtenberg (1997), I construct seventeen instruments that capture the strength of a movie’s competitive environment.

---

23 This is not the case when I examined advertising in the Friday *New York Times*. For example, high cast movies were less likely to be advertised but received more advertising conditional upon receiving positive advertising.
as well as the predictable decay in demand as the market saturates. These IVs for movie j at week t are shown in Appendix 3, and make use of age-adjusted exogenous variables of competing movies that are no older than 4 weeks. They are based upon the general, action, and family principles of differentiation as well as studio portfolios.24

Estimation of the full model using Generalized Method of Moments (GMM) follows a standard two-step process.25 After making a guess for $\rho$, I invert observed purchase probabilities to obtain the implied mean utilities $\delta$ using a contraction mapping. I then find the 2SLS point estimates and the resulting objective function. The initial guess is then updated using a Nelder-Mead simplex search to find the vector of parameters that attains the global minimum of $\xi'Z(Z'Z)^{-1}Z'\xi$. I use these consistent estimates of $\xi$ to construct $(\tilde{Z}\tilde{\Omega}Z)^{-1}$, which serves as as the new and efficient weighting matrix when I repeat the exercise. Standard errors for point estimates are calculated using finite perturbation of the non-linear parameters, and the construction of $\tilde{Z}\tilde{\Omega}Z$ allows for arbitrary heteroskedasticity and a serial correlation within a movie’s time-series (Newey-West with three lags).26 The minimized objective function value can be interpreted as a J-statistic, which under the null hypothesis of valid moment conditions is distributed chi-squared with degrees of freedom equal to the number of overidentifying restrictions.

7 Results

7.1 Monte Carlo

The Monte Carlo experiment makes use of a utility specification that generates the one-level nested logit and thus provides competitive environment instruments that should have more power. For purposes of exposition, the dimension of segmentation can be interpreted as whether or not a movie is in the Family genre. Heterogeneity among consumers is assumed to be real but not overwhelming: $\rho = 0.8$ in the PD GEV notation or $\sigma = 0.2$ in the traditional nested logit notation. I assume that there is a one in three chance that any movie is a Family movie and the average consumer reaction to a Family movie is negative. Besides the Family designation, movies are defined by three observable characteristics and one unobservable characteristic. There are thus eight potential competitive environment instruments.

---

24 Estimates derived without instruments based upon the action-ness genre were similar but substantially less precise.
26 Newey and West (1994).
Table 6 displays Monte Carlo results for estimation of the movie-specific nested logit model with and without the inclusion of advertising. Each simulated sample consists of 5000 duopoly markets and the utility of each consumer in each market is given by

$$u_{ij} = \kappa_0 + X_{ij1}\kappa_1 + X_{ij2}\kappa_2 + X_{ij3}\kappa_3 + Fam_j\gamma$$

$$+ \alpha_T \ln(T_j) + \beta \ln(1 + T_j) \ln(1 + A_j) + \xi_j + \mu_{ij}$$

with $\mu_{ij}$ being the nested logit error which differs from the logit error depending upon $\sigma$. Firms maximize profits of $\Pi_j = p_jM s(\delta_j, \delta_{-j}) - cT_j - \tau A_j$. Other notation is similar.

Exogenous variables ($x, \xi$) are created as independent normal random variables with mean zero. Each of the three observed variables has standard deviation of 0.015, and the unobserved variable has standard deviation of 0.045. Observed exogenous variables therefore comprise only 25% of exogenous variation. The parameters to be estimated are $\sigma (= 1 - \rho)$, $\kappa$, $\gamma$, $\alpha$, and $\beta$. Parameter values were chosen by experiment to mimic real-world outcomes and are $\kappa_0 = -11.3$, $\kappa_{1-3} = 1$, $\gamma = -0.1$, $\alpha_T = 0.6$, $\beta = 0.02$, $\tau = 100$, and $c = 3000$. To further facilitate comparison to actual numbers, I assume an admissions price of $p = 5$ and a population of $M = 300,000,000$.

For each market, I calculate the equilibrium values of theaters, advertising, and purchase-probabilities (considering the possibility that not advertising might be profit-maximizing) by solving the relevant first order conditions, keeping in mind that zero-advertising may be optimal. I assume that these variables and the exogenous variables $x$ and $Fam$ are observed, but that $\xi$ is unobserved by the econometrician. Given the nested logit assumption, the mean utility of good $j$ can be found as $\delta_j = \ln(s_j) - \ln(s_0) - \sigma \ln(s_{j/G})$, where $s_0$ denotes the probability that the typical consumer purchases nothing and $s_{j/G}$ denotes the probability of seeing movie $j$ conditional upon seeing a movie of $j$’s genre $G$.

These parameters yield a sample in which 56% of observations have no advertising and the industry-wide advertising-to-sales ratio is 5.8%. The average number of theaters showing a movie is 1600, and the average admissions are about 1.2 million (slightly over $6M in revenues). Roughly 70% of these revenues are allocated to theater rental expenses.

Table 6 shows 2SLS estimates when ads are included and excluded as well as some characteristics of the generated data set and estimated implications. All estimates when advertising is included are less than one standard error from their true parameter values and most estimates are very precise. The exceptions are the continuous exogenous variables. Implied elasticities match up with actual elasticities quite well. The J-statistic does not reject
the null hypothesis of valid instruments, and no observations have own-theater elasticities that exceed one. Estimates when advertising is omitted, however, have none of these desirable qualities. Segmentation is understated \((\hat{\sigma} \approx 0.06 < 0.2 = \sigma)\), and this in turn somewhat complicates the comparison of other estimates. The average implied derivative of share with respect to \(X_1\) is one-fifth of the true derivative \(\frac{ds_1}{dX_1}\). The average own-theater elasticity when inferred from estimates that exclude advertising exceeds the average true own-theater elasticity by 36%, and 23% of observations exceed the theoretical upper bound of unity. The J-statistic decisively rejects the null of a correct model, but, as stated earlier, this test is unlikely to be useful in practice. Given the rejection of advertising’s impact being zero with its individual t-statistic, there is no need to employ the pseudo-Hausman statistic offered by Creel.

There are thus three predictions from this Monte Carlo experiment:

1) When advertising is omitted, demand’s responsiveness to theaters will be overstated, perhaps to a theoretically unacceptable extent.

2) When advertising is omitted, consumer heterogeneity and resulting segmentation will be understated.

3) When advertising is omitted, demand’s responsiveness to exogenous characteristics will be understated.

How well the actual empirics match these predictions will then shed light upon the general value of the Monte Carlo experiments as they apply to advertising.

7.2 Movie Demand

Table 7 displays the GMM estimates of movie demand when advertising is included as a regressor (first alone and then interacted with theaters) and then omitted. I will first consider the general implications of the full model’s demand estimates in the first two columns and then turn to how excluding advertising substantively affects these conclusions.

Consumer heterogeneity leads to segmentation between movies and non-movie entertainment options and between family and non-family (G or PG-rated) movies.\(^{27}\) These results are qualitatively similar to those of Moul (2006b) which uses one-level nested logits in partial reduced form estimation on this data set. Examining the second-stage results, one can also see that the typical consumer views family movies as markedly inferior to non-family movies.

\(^{27}\) Results which allowed for segmentation between action and non-action movies (not shown) indicated that \(\rho^A \approx 1\).
While the hypothesis that $\rho^M = \rho^F$ cannot be rejected statistically in the second specification (/t-statistics/ are 2.83 and 1.56 respectively), the aforementioned multicollinearity concerns suggest that segmentation between movies and the outside option is likely to be more severe than that between genres.\(^{28}\) The estimated impacts of both theaters and advertising are positive and significant, though I defer the interpretation of these estimates for the moment. In neither model does the J-statistic reject the overidentifying restrictions.

As expected, demand for a movie declines as it ages. Movies with abbreviated weeks intuitively face lower demand than they would have in the counterfactual full week. The average effect of this abbreviation in either specification is 26%, that is, a movie can expect to receive one-fourth of the admissions in its abbreviated opening week that it would have garnered had it opened on Friday.\(^{29}\) Contrary to the descriptive advertising results and Moul (2006b), Oscar nominations are not estimated to have a significant positive impact on admissions. The estimated impact of Oscar wins, however, is greater than zero at 90% confidence levels. Presumably the higher percentage impact of a win for an older movie is less than the smaller percentage impact for a younger movie.

The model’s fit is quite high, but that is unsurprising given the presence of movie fixed effects. Making use of GMM estimates and observed theaters and advertising, the model’s predicted quantities explain about 89% of the observed variation in quantities. When fixed effects are regressed on the appeal and genre variables, $R^2$ is fairly low at less than 5%. Both cast and director appeal are shown to have positive impacts on consumer expectations in this second-stage regression, though cast is significant at only 90% confidence levels in the first specification. These are of course only the direct impacts. In equilibrium, these effects are magnified by endogenous advertising and the number of exhibiting theaters.

The summary statistics regarding the implied own-theater and own-ad elasticities (shown at the bottom of Table 7) are the best way to interpret the coefficients on the endogenous variables. Considering all observations, a 1% increase in the number of exhibiting theaters is expected to increase quantity demanded by about 0.7%. No observations in either specification are estimated to face the theoretically impossibility of a 1% increase in theaters generating a greater than 1% increase in quantity demanded. This mean own-theater elastic-

\(^{28}\) Whether the consumer heterogeneity regarding movie purchases (e.g., preference differences between teenagers and working adults) is inherent to the population or driven by the type of movies that are produced cannot be answered using these techniques.

\(^{29}\) Assuming that these non-Friday releases come primarily on Wednesdays and allowing for the possibility of Thursday holidays, this is within the bounds [0.15, 0.40] suggested by the daily data sets of Davis (2005) and Switzer (2004).
ity is similar to the estimated own-theater elasticity for the U.S. (which includes advertising) in Elberse and Eliashberg (2003) using entirely different data and methods.

Interpreting own-ad elasticities is more informative if one only considers the set of observations in which advertising is positive. Within this subset, a 1% increase in advertising is expected to increase quantity demanded by 0.25%. Given the industry’s institutions, the first-order condition for advertising implies that a profit-maximizing monopolist will (when he advertises) choose advertising so that the ratio of ad expenditures to rental payments equals the own-ad elasticity of demand: \( \eta^A = \frac{z_A}{\tau_A} \). As there are many weeks without positive advertising, this elasticity can serve as an upper bound of the industry’s advertising-to-sales ratio. The robustness of the implications across the two specifications suggests that the data and not functional form are driving the results. For comparison, Ackerberg (2001, 2003) using individual-level television commercial exposure finds an average own-ad elasticity of 0.15 for yogurt. Nevo (2001) using quarterly brand level expenditures of ready-to-eat cereals finds an own-ad elasticity of 0.06 with his preliminary logit results, though he concedes that the result seems low and he does not instrument for advertising.

What happens when demand is estimated without advertising? The Monte Carlo experiment’s first prediction is intuitive, that theaters will pick up advertising’s impact. Indeed, one of the most obvious changes when advertising is omitted is the increased coefficient on \( \ln(TH) \). This upward bias appears to overstate implied own-theater elasticities by about 0.15 percentage points, and 16% of the sample’s observations are now estimated to have \( \eta^T > 1 \). It is noteworthy that the regressions in Elberse and Eliashberg (2003) using European data that yield much higher estimated own-theater elasticities lack advertising variables. Before strong international conclusions regarding business and optimality can be drawn, it is necessary to ensure that comparisons are apt. The distorted implied own-theater elasticities also have implications for other movie-relevant lessons. In terms of one application (Moul, 2006a), the use of the estimated parameters when advertising is omitted would lead one to conclude that distributors are setting rental rates well above even the fully collusive level of wholesale prices.

The Monte Carlo’s other predictions stem from the first, namely that the heightened estimate of theaters’ impact comes at the expense of other variables. While there is some change to the segmentation parameters (now \( \rho_M \ll \rho_F \), \( /t − \text{statistic}/ = 3.84 \)), the prediction that segmentation along the Family dimension will be understated does not appear to
be significant.\textsuperscript{30} The cast appeal estimate in the second-stage regression, however, suffers exactly the substantial downward bias that the Monte Carlo illustrated. When advertising is omitted, one (wrongly) concludes that cast appeal is unimportant and it is no longer significant at any level of confidence. When advertising is included and interacted with theaters, a 10\% increase in cast appeal (for movies with positive cast appeal) on average is estimated to generate a 0.67\% increase in admissions. When advertising is omitted, the same increase in cast appeal is estimated to generate only a 0.11\% increase in admissions. The estimate of director appeal, on the other hand, seems to be little affected by the exclusion of advertising, despite its highly significant impact on movie fixed effects. Estimates that include the interacted advertising generate (on average) a 0.77\% increase in admissions from a 10\% increase in director appeal. Excluding advertising lowers that impact only slightly, to a 0.71\% increase in admissions. This difference between cast and director presumably stems from distributors’ ability to convey appeal information in advertising. The results are consistent with directors greatly affecting the quality/popularity of their movies but studios being limited in their ability to use advertising to exploit this.

\section*{8 Conclusions}

This paper has shown that, despite marginally significant impacts on demand, the omission of advertising from demand can substantially distort inferences, especially regarding the impacts of other endogenous variables. While the direction of these biases will vary across applications, one can reasonably conclude that strong efforts should be made to include advertising in demand estimation whenever advertising-sales ratios exceed 0.05. Furthermore, whenever data are available, initial regressions should include properly instrumented advertising and discard it only after showing that doing so has no substantive impact on other estimates. Low t-statistics are insufficient evidence that advertising can be safely excluded, but the Creel statistic may be quite useful.

The theatrical movie application illustrates that Monte Carlo experiments capture critical forces that stem from advertising and the simulation’s usefulness is not limited to price-setting markets. The relatively precise estimates of theatrical movie demand with respect to advertising bode well for future analysis of this industry. Beyond the demand-side issues

\textsuperscript{30}Given that segmentation between movie-goers and non-movie-goers is identified entirely from functional form, I hesitate to draw any conclusions from heterogeneity along that dimension being somewhat overstated when advertising is omitted.
that have received the most attention, interesting issues regarding supply remain. The
frequency of observation suggests that theatrical movies could prove fruitful for the study of
firms’ responses to the information transmission among consumers inherent to new products.
While the newspaper advertising used in this paper is clearly not ideal, it might be sufficient
for cost-side inferences regarding advertising, an application that has yet to receive attention
in empirical industrial organization.

9 Appendix

9.1 PD GEV framework

From McFadden (1978), suppose $G(\cdot)$ is a nonnegative, homogeneous of degree one function
that approaches $+\infty$ as any argument approaches $+\infty$, its first partials are nonnegative,
and its mixed partials alternate in sign. Then $F(\delta_{t0}, ..., \delta_{tJ}) = \exp(G_t(e^\delta))$ is the cumulative
distribution function of a multivariate extreme value distribution during week $t$, and

$$s_{tj} = \frac{e^{\delta_{tj}} \left( \frac{\partial G}{\partial e^\delta} \right)}{G_t(e^\delta)}$$

is the probability a consumer will choose product $j$ in week $t$. This flexible form reduces to
the multinomial logit when $G_t(e^\delta) = e^{\delta_{t0}} + \sum_{j=1}^{J(t)} e^{\delta_{tj}}$

Recall that $\delta_{t0}$, interpreted as the population-level mean utility of no purchase during week
$t$, is normalized to 0 for identification purposes. The one-level NML is another special case.
Let $\rho$ be a parameter measuring the degree of consumers’ idiosyncratic preference for family
or non-family movies. If $\rho = 1$, there is no idiosyncratic preference and such movies are
perfectly substitutable along that dimension. These movies become less substitutable as
$\rho \to 0$. For such a framework, the function $G_t(\cdot)$ is

$$G_t(e^\delta) = e^{\delta_{t0}} + \sum_{k \in \text{Family}(t)} e^{\delta_{tk}/\rho} + \sum_{k \not\in \text{Family}(t)} e^{\delta_{tk}/\rho}$$

Each summation term includes all the movies (in week $t$) that have the family-ness
indicated by the subscript. Continuing for an analytic expression of purchase-probabilities
yields

$$s_{tj} = \left( \frac{e^{\delta_{tj}}}{G_t(e^\delta)} \right) \left( \frac{(e^{\delta_{tj}/\rho})^{1-\rho}}{\left( \sum_{k \in \text{Family}(tj)} e^{\delta_{tk}/\rho} \right)^{1-\rho}} \right)$$
Summing these probabilities for all movies that share family-ness yields

$$
\sum_{k \in Fa(tj)} s_{tk} = \left( \frac{\left( \sum_{k \in Fa(tj)} e^{\delta_{tk}/\rho} \right)^\rho}{G_t(e^\delta)} \right)
$$

so the within-cluster purchase-probabilities in this framework can be expressed as

$$
s_{Fa(tj)} = \frac{e^{\delta_{tj}/\rho}}{\sum_{k \in Fa(tj)} e^{\delta_{tk}/\rho}}
$$

Using this new notation, the probability that a consumer sees movie j in a given week is

$$
s_{tj} = \frac{e^{\delta_{tj} s_{1-\rho}^{Fa(tj)}}}{G_t(e^\delta)}
$$

One of the larger benefits of these analytical expressions is the direct inversion they allow between observed purchase probabilities and unobserved mean indirect utility levels. The inversions for both the multinomial logit and the one-level nested multinomial logit are well-known:

$$
\delta_{tj} = \ln(s_{tj}) - \ln(s_{t0}) \text{ for logit}
$$

$$
\delta_{tj} + (1 - \rho) \ln(s_{Fa(tj)}) = \ln(s_{tj}) - \ln(s_{t0}) \text{ for nested logit}
$$

Note that, in the latter case, \( \delta_{tj} \) (and therefore unobserved characteristics) still exist in the within-nest market share. Econometric techniques are needed to handle potential endogeneity bias along that front.

Alternatively, one might believe that consumers have idiosyncratic preferences regarding multiple dimensions. To address these concerns simultaneously, I derive a version of Bresnahan, Stern, and Trajtenberg’s PD GEV application (1997). The two PDs that I consider are movie-ness (M) and family-ness (F). Again using McFadden’s notation, I specify \( G_t(\cdot) \) as follows. The family cluster enters in the form

$$
Fa_t(\cdot) = \left( \sum_{k \in Fa(t)} e^{\delta_{tk}/\rho_F} \right)^{\rho_F} + \left( \sum_{k \notin Fa(t)} e^{\delta_{tk}/\rho_F} \right)^{\rho_F}
$$

while the movie cluster takes the form

$$
Mov_t(\cdot) = \left( \sum_{\forall k} e^{\delta_{tk}/\rho_M} \right)^{\rho_M}
$$

As any principle’s \( \rho \to 0 \), the above elements become more segmented. Conversely, as every cluster’s \( \rho \to 1 \), each of the terms becomes a simple sum of \( \exp(\delta_{tk}) \). To combine these two
terms, I follow Bresnahan, Stern, and Trajtenberg (1997) and use a weighted average where the weights depend upon the different \( \rho \)s. Each weight \( (a) \) is

\[
a_{\text{Dim}} = \frac{1 - \rho_{\text{Dim}}}{2 - \rho_M - \rho_F}
\]

By construction, these weights are positive and sum to one, thereby meeting McFadden’s sufficient properties. My specification for \( G_t(\cdot) \) is then

\[
G_t = a_M Mov_t(\cdot) + a_F F_t(\cdot) + e^{\delta_{t0}}
\]

Next I generate analytical forms of the market shares/purchase-probabilities. To this end, I introduce some notation. In the one-level nested logit with family movies, a product’s within-nest market share has the elegant form

\[
s_F(tj) = \exp(\delta_{tj}/\rho) \sum_{Fa(k)=Fa(j)} \exp(\delta_{tk}/\rho)
\]

This result is no longer true in the GEV model, as purchases depend on factors outside any one nest and the market as a whole. I maintain a slightly altered notation for convenience.

\[
\tilde{s}_{\text{Dim}(tj)} = \frac{\exp(\delta_{tj}/\rho_{\text{Cluster}})}{\sum_{k \in \text{Dim}(tj)} \exp(\delta_{tk}/\rho_{\text{Cluster}})}
\]

Using this notation, predicted purchase-probabilities take the form

\[
s_{tj} = \frac{\exp(\delta_{tj})[a_M s_{M(tj)}^{1-\rho_M} + a_F s_{F(tj)}^{1-\rho_F}]}{G_t}
\]

Given guesses for the \( \rho \)s and the parameters within mean utility, predicted market shares depend exclusively upon \( N \) \( \xi \)s, and there exist \( N \) observed unconditional purchase probabilities. I find the \( N \) mean valuations of the unobservable characteristics using a contraction mapping. To ensure that the contraction mapping reaches a solution, it is necessary to weight the difference of the log-shares by the smallest value of \( \rho \): \( \xi_{t+1} = \xi_t + (\ln(s) - \ln(\tilde{s}(\xi_t, \Gamma))) \ast \min(\rho) \).

### 9.2 Appeal definitions

I define an actor or actress’s revenue history at the time of a specific movie to be the sequence of final grosses of all recently released movies in which he or she was a starring cast member. (Chen and Shugan, 2002, use a similar measure.) In my data, this time frame ends with the release of the specific movie and begins five years prior to that release. Various statistics can then be used to summarize this revenue history, and, in the case of multiple-star movies, other
statistics aggregate the several revenue histories and summarize the cast appeal of the movie as a whole. Examining directors is simpler, as movies almost always have a single director. There is correspondingly no issue regarding aggregation of multiple directors or regarding whether a director is a “star”. Directors of movies and their performance histories can be found in The Annual Index to Motion Picture Credits and the Internet Movie Database.

Recognizing that distributors have an interest in informing the public about the presence of celebrities, I use movie advertising in the Friday New York Times as a guide to which cast members are the stars of the movie. Just as theaters of the past used their marques to advertise star presence, many Times ads early in a movie’s run include the names of a few members of the cast. To maintain tractability, I limit this to printed names and not pictures of cast members. The one exception is for sequels, for which a picture of a recurring character is the equivalent of a printed name. I assume that these individuals are the starring cast of those movies (unless certain names are more prominent than others, in which case only the most prominent names are the movie’s stars). While a few movies listed many names, most ads that list any cast display only two or three names.

I build my measure of cast appeal by summing up the revenue histories of the marquee cast and dividing by the number of movies released in the prior five years in which those cast had starred. This measure is intuitive in that it tends to be lower when a movie’s stars are numerous but not very successful at the box office (e.g., Robert Altman films) and when a starring actor in a movie is prolific but not very lucrative (e.g., post-GhostBusters Dan Aykroyd). I measure the appeal of the director by simply summing that individual’s revenue history. Both these appeal variables are then scaled in billions of dollars. As an example, 1993’s Jurassic Park’s director appeal is calculated as the total domestic box office of Steven Spielberg’s three movies in the five years prior (Hook, 1991; Always, 1989; and Indiana Jones and the Last Crusade, 1989), or $0.36B.

9.3 Instrument definitions

All instruments are based upon age-adjusted measures of the number of competing movies as well as their aggregated cast appeal and director appeal. I further break down the set of competing movies into those that are released by the same distributor and those that share the same genre (family or non-family, action or non-action). This leaves a maximum of 18 instruments: 3 variables x 2 portfolio dimensions (general vs. intraportfolio) x 3 clusters (general vs. action vs. family). As the number of available movies each week is constant by
construction at 50, an instrument based upon the number of competing movies regardless of
genre or ownership is constant at 49. There are therefore 17 instruments. Let

\[ Port(t,j) = \{ \text{other movies available in week } t \text{ that were also released by the } \\
\text{distributor of movie } j \text{ and are no more than 4 weeks old} \}, \]

\[ A(t,j) = \{ \text{other movies available in week } t \text{ that share Action-ness status with movie } j \text{ and } \\
\text{are no more than 4 weeks old} \}, \]

\[ F(t,j) = \{ \text{other movies available in week } t \text{ that share Family-ness status with movie } j \text{ and } \\
\text{are no more than 4 weeks old} \}. \]

The 17 instruments are then (ordered first by genre regardless of ownership and then
within portfolios by genre)

\[
\sum_{k \in A(t,j)} \left( \frac{\text{CastApp}_k}{\text{Age}_{tk}} \right), \sum_{k \in A(t,j)} \left( \frac{\text{DirApp}_k}{\text{Age}_{tk}} \right), \\
\sum_{k \in F(t,j)} \left( \frac{\text{CastApp}_k}{\text{Age}_{tk}} \right), \sum_{k \in F(t,j)} \left( \frac{\text{DirApp}_k}{\text{Age}_{tk}} \right) \\
\sum_{k \in Port(t,j)} \left( \frac{1}{\text{Age}_{tk}} \right), \sum_{k \in Port(t,j)} \left( \frac{\text{CastApp}_k}{\text{Age}_{tk}} \right), \sum_{k \in Port(t,j)} \left( \frac{\text{DirApp}_k}{\text{Age}_{tk}} \right) \\
\sum_{k \in A(t,j) \cap Port(t,j)} \left( \frac{1}{\text{Age}_{tk}} \right), \sum_{k \in A(t,j) \cap Port(t,j)} \left( \frac{\text{CastApp}_k}{\text{Age}_{tk}} \right), \sum_{k \in A(t,j) \cap Port(t,j)} \left( \frac{\text{DirApp}_k}{\text{Age}_{tk}} \right) \\
\sum_{k \in F(t,j) \cap Port(t,j)} \left( \frac{1}{\text{Age}_{tk}} \right), \sum_{k \in F(t,j) \cap Port(t,j)} \left( \frac{\text{CastApp}_k}{\text{Age}_{tk}} \right), \sum_{k \in F(t,j) \cap Port(t,j)} \left( \frac{\text{DirApp}_k}{\text{Age}_{tk}} \right)
\]

Instruments that looked back more than 4 weeks provided no better fit for separate first-stage
regressions with \( \ln(s_M) \) and \( \ln(s_F) \) as dependent variables.

References


[21] Internet Movie Database (Available at http://www.imdb.com)


Table 1: Monte Carlo IV parameter estimates
100 random samples of 500 duopoly markets using logit utility

<table>
<thead>
<tr>
<th>E(A/S)</th>
<th>0.024</th>
<th>0.061</th>
<th>0.096</th>
</tr>
</thead>
<tbody>
<tr>
<td>1: $\beta_0 = 4, \beta_{Ads} = 0.1$</td>
<td>(1)</td>
<td>(2)</td>
<td>(3)</td>
</tr>
<tr>
<td>2: $\beta_0 = 3, \beta_{Ads} = 0.25$</td>
<td>(4)</td>
<td>(5)</td>
<td>(6)</td>
</tr>
<tr>
<td>3: $\beta_0 = 2, \beta_{Ads} = 0.4$</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>W/ Ads</th>
<th>W/o Ads</th>
<th>W/ Ads</th>
<th>W/o Ads</th>
<th>W/ Ads</th>
<th>W/o Ads</th>
</tr>
</thead>
<tbody>
<tr>
<td>$b_0$</td>
<td>3.751</td>
<td>4.954</td>
<td>2.743</td>
<td>5.766</td>
<td>1.792</td>
<td>6.984</td>
</tr>
<tr>
<td></td>
<td>2.631</td>
<td>0.189</td>
<td>2.473</td>
<td>0.223</td>
<td>2.179</td>
<td>0.275</td>
</tr>
<tr>
<td>$\beta_x = 2$</td>
<td>1.954</td>
<td>2.148</td>
<td>1.952</td>
<td>2.439</td>
<td>1.959</td>
<td>2.832</td>
</tr>
<tr>
<td></td>
<td>0.432</td>
<td>0.077</td>
<td>0.404</td>
<td>0.091</td>
<td>0.373</td>
<td>0.111</td>
</tr>
<tr>
<td>$\alpha = -1$</td>
<td>-0.979</td>
<td>-1.083</td>
<td>-0.978</td>
<td>-1.241</td>
<td>-0.981</td>
<td>-1.458</td>
</tr>
<tr>
<td></td>
<td>0.231</td>
<td>0.037</td>
<td>0.217</td>
<td>0.043</td>
<td>0.203</td>
<td>0.053</td>
</tr>
<tr>
<td>$b_{Ads}$</td>
<td>0.127</td>
<td>0.000</td>
<td>0.274</td>
<td>0.000</td>
<td>0.417</td>
<td>0.000</td>
</tr>
<tr>
<td></td>
<td>0.278</td>
<td>---</td>
<td>0.223</td>
<td>---</td>
<td>0.174</td>
<td>---</td>
</tr>
<tr>
<td>E(J-stats)</td>
<td>0.887</td>
<td>2.388</td>
<td>0.885</td>
<td>4.430</td>
<td>0.880</td>
<td>10.141</td>
</tr>
<tr>
<td>DoF</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>Pr(X^2&gt;X)</td>
<td>0.346</td>
<td>0.303</td>
<td>0.347</td>
<td>0.109</td>
<td>0.348</td>
<td>0.006</td>
</tr>
<tr>
<td>E(Haus1)/Pr</td>
<td>2.393</td>
<td>0.664</td>
<td>5.476</td>
<td>0.242</td>
<td>13.128</td>
<td>0.011</td>
</tr>
<tr>
<td>E(Haus2)/Pr</td>
<td>4.001</td>
<td>0.406</td>
<td>8.747</td>
<td>0.068</td>
<td>20.980</td>
<td>0.000</td>
</tr>
</tbody>
</table>

Notes: The values given in the table are empirical means and (standard errors).

The true utility function is $u_i = \beta_0 + \beta_x x_i + \alpha p_i + \beta_{Ads} \ln(A_i) + \varepsilon_i + \epsilon_i$

Marginal cost is $c_i = \exp(1 + 0.5x_i + 0.25w_1 + 0.25w_2 + 0.25\xi + 0.25\omega)$

Firm profits are $\Pi_i = (p_i - c_i)M_i - 100A_i$

M=1000000; (x, w_1, w_2, \xi, \omega) ~ N(0,1)
Table 2: Implications of Monte Carlo IV parameter estimates
100 random samples of 500 duopoly markets using logit utility

<table>
<thead>
<tr>
<th></th>
<th>1: $\beta_0 = 4$, $\beta_{Ads} = 0.1$</th>
<th>2: $\beta_0 = 3$, $\beta_{Ads} = 0.25$</th>
<th>3: $\beta_0 = 2$, $\beta_{Ads} = 0.4$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$E(\text{True} , \eta_{\text{OwnP}})$</td>
<td>-3.614</td>
<td>-3.601</td>
<td>-3.566</td>
</tr>
<tr>
<td>$E(\text{True} , \eta_{\text{CrossP}})$</td>
<td>1.467</td>
<td>1.513</td>
<td>1.606</td>
</tr>
<tr>
<td>$E(\text{True} , \eta_{\text{OwnA}})$</td>
<td>0.069</td>
<td>0.172</td>
<td>0.270</td>
</tr>
<tr>
<td>$E(\text{True ds/dX})$</td>
<td>0.354</td>
<td>0.351</td>
<td>0.351</td>
</tr>
<tr>
<td>$E(\text{True Lerner})$</td>
<td>0.376</td>
<td>0.379</td>
<td>0.386</td>
</tr>
<tr>
<td>$E(\text{Est} , \eta_{\text{OwnP}})$</td>
<td>-3.535</td>
<td>-3.915</td>
<td>-3.520</td>
</tr>
<tr>
<td>$E(\text{Est} , \eta_{\text{CrossP}})$</td>
<td>1.436</td>
<td>1.590</td>
<td>1.479</td>
</tr>
<tr>
<td>$E(\text{Est} , \eta_{\text{OwnA}})$</td>
<td>0.088</td>
<td>---</td>
<td>0.188</td>
</tr>
<tr>
<td>$E(\text{Est ds/dX})$</td>
<td>0.346</td>
<td>0.381</td>
<td>0.343</td>
</tr>
<tr>
<td>$E(\text{Est Lerner})$</td>
<td>0.402</td>
<td>0.348</td>
<td>0.404</td>
</tr>
</tbody>
</table>

Note: 
W/ Ads: With Ads
W/o Ads: Without Ads
Table 3: Summary statistics

<table>
<thead>
<tr>
<th>By observation: N=16600</th>
<th>Q</th>
<th>Mean</th>
<th>StDev</th>
<th>Median</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>441878</td>
<td>1013529</td>
<td>84904</td>
<td>593</td>
<td>19120803</td>
</tr>
<tr>
<td>Theaters</td>
<td></td>
<td>579.67</td>
<td>657.15</td>
<td>297</td>
<td>1</td>
<td>3700</td>
</tr>
<tr>
<td>Q/Theaters</td>
<td></td>
<td>991</td>
<td>3862</td>
<td>420</td>
<td>32</td>
<td>376397</td>
</tr>
<tr>
<td>CT Sunday</td>
<td></td>
<td>3.86</td>
<td>7.27</td>
<td>0</td>
<td>0</td>
<td>241</td>
</tr>
<tr>
<td></td>
<td>CT&gt;0 (34.29%)</td>
<td></td>
<td>11.27</td>
<td>8.41</td>
<td>9.66</td>
<td>0.65</td>
</tr>
<tr>
<td>OscNom?</td>
<td></td>
<td>0.05</td>
<td>0.22</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>OscWin?</td>
<td></td>
<td>0.01</td>
<td>0.11</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Ab?</td>
<td></td>
<td>0.005</td>
<td>0.070</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Age</td>
<td></td>
<td>9.00</td>
<td>7.82</td>
<td>7</td>
<td>1</td>
<td>71</td>
</tr>
<tr>
<td>Major Studio?</td>
<td></td>
<td>0.75</td>
<td>0.43</td>
<td>1</td>
<td>0</td>
<td>1</td>
</tr>
</tbody>
</table>

| By movie: N=1602       | Mean | StDev | Median | Min  | Max  |
| Ab?                    | 0.05 | 0.22  | 0      | 0    | 1    |
| CastAppeal (in $B)     | 0.016| 0.022 | 0.005  | 0    | 0.142|
| ||CastApp>0 (55.68%)     | 0.029| 0.023 | 0.025  | 0    | 0.142|
| DirAppeal (in $B)      | 0.033| 0.067 | 0.002  | 0    | 0.718|
| ||DirApp>0 (58.55%)      | 0.056| 0.080 | 0.025  | 0    | 0.718|
| Action                 | 0.25 | 0.44  | 0      | 0    | 1    |
| Family                 | 0.20 | 0.40  | 0      | 0    | 1    |
| Major Studio?          | 0.67 | 0.47  | 1      | 0    | 1    |
### Table 4: Correlations of (transformed) endogenous variables

<table>
<thead>
<tr>
<th></th>
<th>ln(s)</th>
<th>ln(sM)</th>
<th>ln(sF)</th>
<th>ln(T)</th>
<th>ln(1+A)</th>
<th>ln(Age)</th>
</tr>
</thead>
<tbody>
<tr>
<td>ln(s)</td>
<td>1.00</td>
<td>0.98</td>
<td>0.93</td>
<td>0.83</td>
<td>0.72</td>
<td>-0.34</td>
</tr>
<tr>
<td>ln(sM)</td>
<td></td>
<td>1.00</td>
<td>0.94</td>
<td>0.83</td>
<td>0.71</td>
<td>-0.34</td>
</tr>
<tr>
<td>ln(sF)</td>
<td></td>
<td></td>
<td>1.00</td>
<td>0.80</td>
<td>0.65</td>
<td>-0.28</td>
</tr>
<tr>
<td>ln(TH)</td>
<td></td>
<td></td>
<td></td>
<td>1.00</td>
<td>-0.52</td>
<td></td>
</tr>
<tr>
<td>ln(1+ADS)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>-0.52</td>
</tr>
<tr>
<td>ln(Age)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1.00</td>
</tr>
</tbody>
</table>

**Share definitions**

- \( s_{t,j} = \frac{Q_{t,j}}{M_t} \)
- \( s^M_{t,j} = \frac{Q_{t,j}}{\sum_k Q_{t,k}} \)
- \( s^F_{t,j} = \frac{Q_{t,j}}{\sum_{k, Fal(k)=Fat(j)} Q_{t,k}} \)

---

The table above presents the correlations of (transformed) endogenous variables. The variables include `ln(s)` for the share of a specific good, `ln(sM)` for the share of a specific good relative to total market share, `ln(sF)` for the share of a specific good relative to a specific family's consumption, `ln(T)` for total sales, `ln(1+A)` for the logarithm of one plus an aggregate variable, and `ln(Age)` for the logarithm of age. The correlations are measured using the natural logarithm of each variable.

The share definitions provided give a clear understanding of how each share is calculated, with `s_{t,j}` being the share of good `j` in year `t`, `s^M_{t,j}` being the share relative to the total market, and `s^F_{t,j}` being the share relative to a specific family's consumption.
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>b</td>
<td>s.e.</td>
<td>b</td>
<td>s.e.</td>
<td>b</td>
<td>s.e.</td>
<td>b</td>
<td>s.e.</td>
</tr>
<tr>
<td>( \lambda )</td>
<td>----</td>
<td>----</td>
<td>----</td>
<td>----</td>
<td>42.98</td>
<td>5.04</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age-1</td>
<td>-0.39</td>
<td>0.01</td>
<td>-0.41</td>
<td>0.02</td>
<td>-0.092</td>
<td>0.002</td>
<td>-3.58</td>
<td>0.38</td>
</tr>
<tr>
<td>Cast</td>
<td>30.28</td>
<td>2.22</td>
<td>34.44</td>
<td>4.35</td>
<td>4.824</td>
<td>0.457</td>
<td>167.37</td>
<td>22.06</td>
</tr>
<tr>
<td>Dir</td>
<td>10.97</td>
<td>0.66</td>
<td>11.68</td>
<td>1.19</td>
<td>1.502</td>
<td>0.137</td>
<td>53.27</td>
<td>6.72</td>
</tr>
<tr>
<td>AC?</td>
<td>0.37</td>
<td>0.12</td>
<td>1.18</td>
<td>0.24</td>
<td>0.004</td>
<td>0.024</td>
<td>0.94</td>
<td>0.84</td>
</tr>
<tr>
<td>FA?</td>
<td>1.10</td>
<td>0.12</td>
<td>2.11</td>
<td>0.25</td>
<td>0.106</td>
<td>0.025</td>
<td>4.68</td>
<td>0.93</td>
</tr>
<tr>
<td>Nom?</td>
<td>2.59</td>
<td>0.28</td>
<td>0.10</td>
<td>0.59</td>
<td>0.919</td>
<td>0.061</td>
<td>31.44</td>
<td>4.36</td>
</tr>
<tr>
<td>Win?</td>
<td>1.93</td>
<td>0.54</td>
<td>1.25</td>
<td>1.40</td>
<td>0.123</td>
<td>0.131</td>
<td>7.73</td>
<td>4.80</td>
</tr>
<tr>
<td>int</td>
<td>4.95</td>
<td>0.09</td>
<td>10.75</td>
<td>0.18</td>
<td>0.002</td>
<td>0.019</td>
<td>-21.94</td>
<td>3.88</td>
</tr>
<tr>
<td>n</td>
<td>16600</td>
<td></td>
<td>5692</td>
<td></td>
<td>16600</td>
<td></td>
<td>5692</td>
<td></td>
</tr>
<tr>
<td>( R^2 )</td>
<td>0.17</td>
<td></td>
<td>0.13</td>
<td></td>
<td>----</td>
<td></td>
<td>0.19</td>
<td></td>
</tr>
<tr>
<td>log(L)</td>
<td>----</td>
<td></td>
<td>----</td>
<td></td>
<td>-9119</td>
<td></td>
<td>----</td>
<td></td>
</tr>
</tbody>
</table>

Notes: All estimated standard errors are based upon i.i.d. disturbances.
Table 6: Monte Carlo IV parameter estimates
100 random samples of 5000 duopoly markets using nested logit utility

### Sample summary statistics

<p>| | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>E(A/S)</td>
<td>0.0581</td>
<td>E(T)</td>
</tr>
<tr>
<td>E(sh)</td>
<td>0.0041</td>
<td>E(cT/S)</td>
</tr>
<tr>
<td>% A=0</td>
<td>55.77%</td>
<td>% Fam</td>
</tr>
</tbody>
</table>

### Sample true elasticities and derivatives

<table>
<thead>
<tr>
<th></th>
<th>(1) W/ ads</th>
<th></th>
<th>(2) W/o ads</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>E(b)</td>
<td>E(a.s.e.)</td>
<td>E(b)</td>
<td>E(a.s.e.)</td>
</tr>
<tr>
<td>σ = 1-ρ = 0.2</td>
<td>0.1995</td>
<td>0.0035</td>
<td>0.0604</td>
<td>0.0027</td>
</tr>
<tr>
<td>κ₀ = -11.3</td>
<td>-11.3047</td>
<td>0.0312</td>
<td>-12.4913</td>
<td>0.0282</td>
</tr>
<tr>
<td>κ₁ = 1</td>
<td>0.9883</td>
<td>0.0605</td>
<td>0.2015</td>
<td>0.1148</td>
</tr>
<tr>
<td>κ₂ = 1</td>
<td>0.9887</td>
<td>0.0606</td>
<td>0.1989</td>
<td>0.1150</td>
</tr>
<tr>
<td>κ₃ = 1</td>
<td>0.9895</td>
<td>0.0609</td>
<td>0.1901</td>
<td>0.1156</td>
</tr>
<tr>
<td>γₜ = -0.1</td>
<td>-0.0994</td>
<td>0.0042</td>
<td>-0.0370</td>
<td>0.0079</td>
</tr>
<tr>
<td>αₜ = 0.6</td>
<td>0.6012</td>
<td>0.0087</td>
<td>0.9569</td>
<td>0.0053</td>
</tr>
<tr>
<td>β = 0.02</td>
<td>0.0200</td>
<td>0.0005</td>
<td>0.0000</td>
<td>----</td>
</tr>
<tr>
<td>E(J-stat)</td>
<td>5.22</td>
<td></td>
<td>470.75</td>
<td></td>
</tr>
<tr>
<td>DoF</td>
<td>5</td>
<td></td>
<td>6</td>
<td></td>
</tr>
<tr>
<td>Pr(F)</td>
<td>0.39</td>
<td></td>
<td>0.00</td>
<td></td>
</tr>
<tr>
<td>E(Est ds/dX₁)</td>
<td>1.0529</td>
<td></td>
<td>0.2046</td>
<td></td>
</tr>
<tr>
<td>E(Est η¹)</td>
<td>0.7123</td>
<td></td>
<td>0.9700</td>
<td></td>
</tr>
<tr>
<td>% (Est η¹ &gt; 1)</td>
<td>0.00%</td>
<td></td>
<td>23.00%</td>
<td></td>
</tr>
<tr>
<td>E(Est η₈)</td>
<td>0.0580</td>
<td></td>
<td>----</td>
<td></td>
</tr>
</tbody>
</table>

Notes: The values given in the table are empirical means and (standard errors)
The true utility function is \( u_i = κ₀X₁ + κ₁X₂ + κ₂X₃ + γFam_i + αTln(T_i) + βln(T_i)ln(1+A_i) + ξ_i + μ_i(σ) \)
Firm profits are \( Π = 5M_0 - 3000T_i - 100A_i \)
\( M = 300,000,000; \{x\} \sim 0.015 N(0,1); ξ_i \sim 0.045 N(0,1); Pr(Fam = 1) = 1/3 \)
Table 7: PD GEV Model of Theatrical Movie Market

<table>
<thead>
<tr>
<th></th>
<th>(1) Ads (as level)</th>
<th>(2) Ads (interacted w/ theaters)</th>
<th>(3) Ads omitted</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>b</td>
<td>s.e.</td>
<td>b</td>
</tr>
<tr>
<td>$\rho^M$</td>
<td>0.484</td>
<td>0.101</td>
<td>0.544</td>
</tr>
<tr>
<td>$\rho^F$</td>
<td>0.806</td>
<td>0.038</td>
<td>0.797</td>
</tr>
<tr>
<td>ln(T)</td>
<td>0.501</td>
<td>0.062</td>
<td>0.458</td>
</tr>
<tr>
<td>ln(1+A)</td>
<td>0.184</td>
<td>0.067</td>
<td></td>
</tr>
<tr>
<td>ln(T+1)*ln(1+A)</td>
<td>-0.294</td>
<td>0.087</td>
<td>-0.292</td>
</tr>
<tr>
<td>ln(Age)</td>
<td>-0.063</td>
<td>0.016</td>
<td>-0.068</td>
</tr>
<tr>
<td>(ln(Age))^2</td>
<td>-0.893</td>
<td>0.091</td>
<td>-0.927</td>
</tr>
<tr>
<td>OscNom?</td>
<td>0.061</td>
<td>0.045</td>
<td>0.064</td>
</tr>
<tr>
<td>OscWin?</td>
<td>0.099</td>
<td>0.050</td>
<td>0.095</td>
</tr>
<tr>
<td>GMM OBJ ($\Phi$)</td>
<td>12.94</td>
<td>13.13</td>
<td>17.52</td>
</tr>
<tr>
<td>Degrees of Freedom</td>
<td>13</td>
<td>13</td>
<td>14</td>
</tr>
<tr>
<td>Pr($\chi^2 &gt; \Phi$)</td>
<td>0.45</td>
<td>0.44</td>
<td>0.23</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>b</th>
<th>a.s.e.</th>
<th>b</th>
<th>s.e.</th>
<th>b</th>
<th>s.e.</th>
</tr>
</thead>
<tbody>
<tr>
<td>CastApp</td>
<td>0.953</td>
<td>0.523</td>
<td>1.432</td>
<td>0.533</td>
<td>0.240</td>
<td>0.618</td>
</tr>
<tr>
<td>DirApp</td>
<td>0.704</td>
<td>0.167</td>
<td>0.799</td>
<td>0.171</td>
<td>0.721</td>
<td>0.208</td>
</tr>
<tr>
<td>Action?</td>
<td>-0.043</td>
<td>0.028</td>
<td>-0.033</td>
<td>0.028</td>
<td>-0.103</td>
<td>0.033</td>
</tr>
<tr>
<td>Family?</td>
<td>-0.142</td>
<td>0.033</td>
<td>-0.126</td>
<td>0.033</td>
<td>-0.163</td>
<td>0.039</td>
</tr>
<tr>
<td>Con</td>
<td>-10.476</td>
<td>0.017</td>
<td>-10.500</td>
<td>0.017</td>
<td>-9.899</td>
<td>0.021</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>b</th>
<th>s.e.</th>
</tr>
</thead>
<tbody>
<tr>
<td>R²</td>
<td>0.040</td>
<td>0.045</td>
</tr>
</tbody>
</table>

Implications

<p>| | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>$E(\eta^T)$</td>
<td>0.710</td>
<td>0.680</td>
</tr>
<tr>
<td>range($\eta^T$)</td>
<td>[0.589, 0.822]</td>
<td>[0.581, 0.906]</td>
</tr>
<tr>
<td>% obs of $\eta^T &gt; 1$</td>
<td>0.00%</td>
<td>0.00%</td>
</tr>
<tr>
<td>$E(\eta^A</td>
<td>A&gt;0)$</td>
<td>0.248</td>
</tr>
<tr>
<td>range($\eta^A</td>
<td>A&gt;0$)</td>
<td>[0.097, 0.295]</td>
</tr>
</tbody>
</table>

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Ab %</td>
<td>0.265</td>
</tr>
<tr>
<td>OscNom%</td>
<td>1.090</td>
</tr>
<tr>
<td>OscWin%</td>
<td>1.151</td>
</tr>
</tbody>
</table>

Notes: GMM estimation uses Newey-West estimator with 3 week lag for covariance matrix
Coefficients for characteristics of movies observed three or fewer times, price, unemployment, and fixed effects for months, holidays, and movies observed more than three times not shown
2nd stage estimation uses OLS regression, estimated standard errors are robust to arbitrary heteroskedasticity