Is Category Spanning Truly Disadvantageous?

New Evidence from Primary and Secondary Movie Markets

Marc Keuschnigg[1,*] and Thomas Wimmer[2]

[1]Institute for Analytical Sociology, LIU, Norra Grytsgatan 10, 601 74 Norrköping, Sweden
[2]Department of Sociology, LMU Munich, Konradstrasse 6, 80801 Munich, Germany
* Corresponding author; marc.keuschnigg@liu.se

Supplementary Material

doi:10.1093/sf/sox043

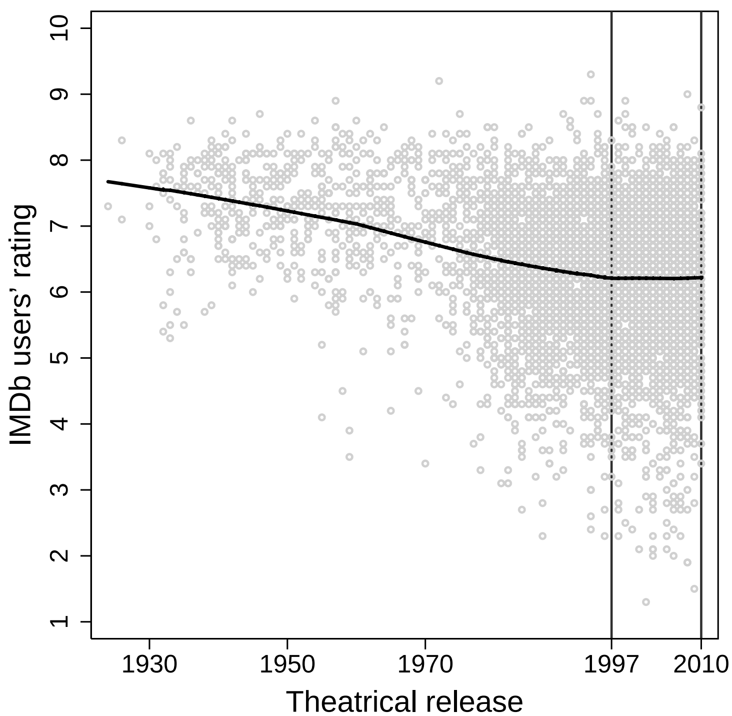
Abstract

Genre assignments help audiences make sense of new releases. Studies from a wide range of market contexts have shown that generalists defying clear mapping to established categories suffer penalties in market legitimacy, perceived quality, or audience attention. We introduce an empirical strategy to disentangle two mechanisms, reduced niche fitness and audience confusion, causing devaluation or ignorance of boundary-crossing offers. Our data on 2 971 feature films released to U.S. theaters and subsequently made available on digital versatile disc (DVD) further reveal that consequences of category spanning are subject to strong moderating influences. Negative effects are far from universal, manifesting only if (a) combined genres are culturally distant, (b) products are released to a stable and highly-institutionalized market context, and (c) offers lack familiarity as an alternative source of market recognition. Our study provides ramifications as to the scope conditions of categorization effects and modifies some widely acknowledged truisms regarding boundary crossing in cultural markets.

A1 Choosing the Period of Observation

In February 2015, a total number of 6 130 films traded on *Amazon Marketplace* for which additional data would have been available from *Box Office Mojo* (BOM), the *Internet Movie Database* (IMDb), and *Rotten Tomatoes* (RT). These films, currently available on DVD, were released to U.S. theaters as early as 1924. Figure A1 plots theatrical release dates against average IMDb ratings. We included a lowess smoother to highlight the non-linear association. Clearly, this larger cut set of films biases strongly toward movies with more favorable evaluations. For films released before 1997, the year of DVD introduction, both variables correlate with $r = -.373; p < .001$, indicating that from the large range of pre-digital movies mostly "better" pictures were re-released as DVDs. The correlation reduces to $r = -.009; p < .001$ for 1997–2010. The period currently studied is thus free from selection bias.

Figure A1: Data Window 1997–2010



A2 Measuring Category Spanning

Our operationalization of category spanning follows Hsu, Hannan, and Koçak (2009) and Zhao, Ishihara, and Lounsbury (2013). We, first, collected genre information from two archival sources, IMDb and RT, and identified 15 categories common to both sources: action, adventure, animation, comedy, documentary, drama, family, fantasy, horror, musical, romance, science fiction, sport, thriller, and Western. Note that Hsu et al. use three websites, IMDb, RT, and *Showbizdata.com*. Similar to their approach, we only consider categories used in each of our sources and carefully re-classify films with alternative assignments: (a) We refer a film to one of our 15 larger categories when in one source it appears in a discounted subcategory. RT, for example, does not use the label 'thriller,' but the subcategory 'mystery and suspense.' (b) We allocate hybrid cases such as 'action and adventure' or 'science fiction and fantasy' to each higher-level category.

Second, we determined each film's grade of membership (GoM $\in [0, 1]$) in a category based on the proportion of archival sources listing it under a certain genre. Average GoM is largest for drama (.616) and smallest for sport (.024), indicating that audiences mostly agree on the assignment of tragic films, but disagree when classifying movies depicting sporting scenes. To prevent any genre specifics biasing our estimates, we control for films' GoM in all models.

Third, we calculated a film's category spanning based on the dispersion of its graded memberships over categories $c = 1, 2, ..., C$ ($C = 15$ in our case). Considering multiple archival sources $a = 1, 2, ..., A$, film i's grade of membership in category c is given by $\text{GoM}_{i,c} = \frac{1}{A}\sum_{a=1}^{A} I_{a,c}$, where $I_{a,c}$ indicates whether archive a classifies i as a member of category c ($I_{a,c} = 1$) or not ($I_{a,c} = 0$) and summation runs over all selected archives ($A = 2$ in our case).

Fourth, we transformed each graded membership $\text{GoM}_{i,c}$ into a relative frequency $\mu_{i,c}$ by dividing by the sum of film i's grades of membership: $\mu_{i,c} = \text{GoM}_{i,c}/\sum_{j=1}^{C}\text{GoM}_{i,j}$. Each weighted GoM, $\mu \in [0, 1]$, reflects a film's degree of category engagement, increasing with both fewer categories spanned and stronger consensus among archives. 1 minus the sum of $\mu_{i,c}^2$ over all categories c, $S_i = 1 - \sum_{c=1}^{C}\mu_{i,c}^2$, returns the standard measure of category spanning, ranging from 0 to 1. The complement secures intuitive interpretation, with $S_i = 0$ representing identical pure-type category assignments in all archives and $S_i > 0$ denoting boundary crossing. S_i increases with both the number of categories assigned and the disagreement across archives.

A3 Exhibitors' Allocation Behavior

Exhibitors are crucial intermediaries in the film industry (DeVany 2004; Hirsch 1972). Their allocation decisions increase both the visibility and the availability of new releases and certainly reduce audiences' uncertainty as to a film's underlying features (Goldman 1982, 81; Zuckerman and Kim 2003, 34). Exhibitors are quite sympathetic to category-spanning films (see Table A1). Generalists receive relatively more screenings upon release than pure-type films (models 1 and 2). Again, category spanning is irrelevant in the independent segment (models 3 and 4). One interpretation suggests that this allocation strategy rests on exhibitors' belief that generalists appeal to different target groups and thus attract larger overall audiences.

Table A1: Number of Opening Theaters for Category-Spanning Films

	Opening theaters, 3 yrs relative							
	1 Majors		2 Majors		3 Independents		4 Independents	
	β	se	β	se	β	se	β	se
Category spanning	.111*	(.039)	.120**	(.039)	.024	(.034)	.026	(.034)
IMDb rating			.016***	(.004)			.017**	(.005)
Budget	.424***	(.027)	.424***	(.027)	.536***	(.034)	.534***	(.034)
Budget data missing	−.087***	(.016)	−.086***	(.016)	−.221***	(.013)	−.217***	(.013)
Star power	.021**	(.007)	.014	(.008)	.036***	(.006)	.032***	(.006)
Season	−.009	(.009)	−.011	(.009)	−.016	(.008)	−.017*	(.008)
Foreign production	−.218***	(.028)	−.223***	(.028)	−.072***	(.010)	−.079***	(.010)
Constant	.177***	(.026)	.181***	(.026)	−.019	(.035)	−.015	(.035)
N movies	1452		1452		1519		1519	
Pseudo R^2	.205		.208		.196		.199	

Note: Tobit regressions. Non-standardized coefficients, robust standard errors in parentheses. All models include genre GoMs and year dummies. *** $p<0.001$, ** $p<0.01$, * $p<0.05$.

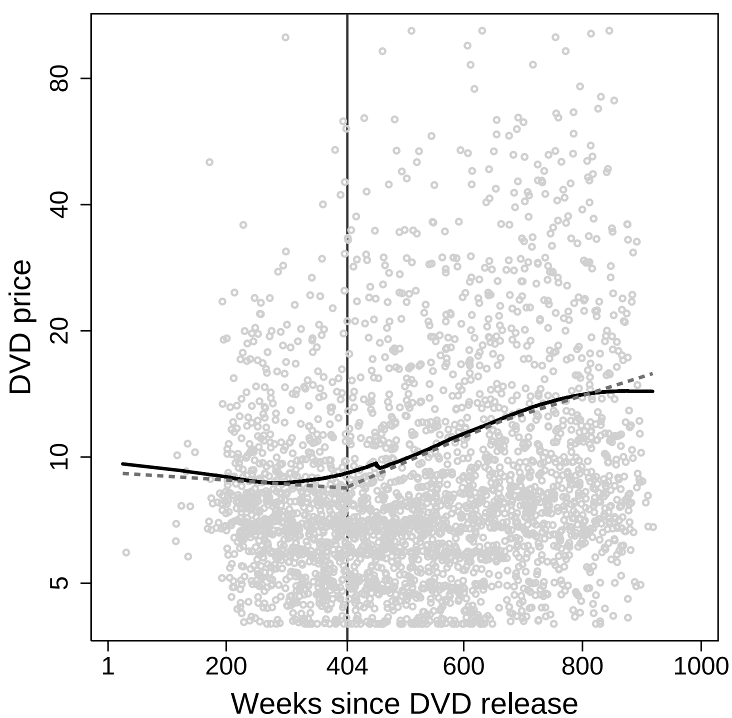
A4 Modeling the DVD Price Trend

We derived minimum prices from a total of 89 098 bids at *Amazon Marketplace*, on average 30 per sampled film. Sampled minimum price is US$4.00. Nominal prices of DVDs may be as low as US$.01. Marketplace products, however, ship at a fee of US$3.99 which Amazon shares with sellers and adds to DVDs' final price. We thus restrict our analysis to real prices (i.e. nominal price + shipping cost). We excluded all films trading for more than US$100.

Similar to the primary market, audience attention in the secondary market is large shortly after release and declines thereafter. However, the composition of market offers varies over time, resulting in a complex relationship between the time since a DVD's release and its current price. A glut of cheap offers reduces average price for aging DVDs which reaches a minimum approximately seven years after original secondary market introduction. Following this sell-off period, average price recovers substantially. Older films appear to benefit from limited supply and collectors' interest.

We employ a linear spline approach to control for this time trend. We estimated a cubic model ($\log(price) = \hat{\beta}_0 + \hat{\beta}_1 t + \hat{\beta}_2 t^2 + \hat{\beta}_3 t^3 + \hat{\beta}_4 \log(offers) + \hat{u}$; $R^2 = .539$) to identify the optimal location of knots connecting piecewise linear functions. It turned out that a single knot at the polynomial's minimum (404 weeks after DVD release) is sufficient to model the trend effectively. Figure A2 shows the resulting piecewise specification (dashed lines) we adopt in our secondary-market models and compares it to a non-parametric lowess smoother capturing mean price over time (solid curve).

Figure A2: Time Trend in DVD Prices



Release dates are only available from Amazon for 68% of sampled DVDs. Availability is unrelated to both primary market success ($r = -.028; p = .124$) and the year of theatrical release ($r = .007; p = .722$). For the remaining 956 DVDs we imputed release dates based on other films' median time span between theatrical release and DVD release ("video window"). The median out-of-market gap decreased during the time of study from 78 to 17 weeks and these out-of-market gaps differ between major and independent films. We thus conditioned imputation on the year of theatrical release and on distributor type. Imputation does not change secondary market results, which show similarly for a reduced dataset ($N = 2\,015$).

A5 Robustness Analyses

27% of samples films lack budget information. Instead of dropping these films we introduced a binary variable with value 1 for films without budget data (cf., Hsu et al. 2009). Regressed on niche fitness (Table 2), the indicator for missing budget data has a negative sign, meaning that production costs are underreported for poorly rated films. Exclusion of all films without budget information would thus induce success bias. Still, our main finding as to which category spanners' evaluations reduce only in the mainstream segment does not change if we exclude films with missing budget information. Point estimates are –.509 (model 1) and –.221 (model 2) for expert critique and –.550 (model 3) and –.199 (model 4) for audience ratings, respectively. Coefficients are only significant for major movies. The indicator also correlates negatively with primary market success (Table 3). Again, our results are robust against exclusion. For a reduced sample of 1 272 majors and 892 independents, estimates for boundary crossing are –.096 (model 1) and –.078 (model 2) in the mainstream segment and –.048 and –.040 in the independent segment, respectively. Estimators are significant only for major movies. We also get similar results for Table 3 if we substitute log-transformed cardinal measures for the relative indicators of box-office totals, opening screens, and production budgets. Marginal effects of category spanning then reach –.660 (model 1) and –.533 (model 2) for majors and –.576 (model 3) and –.466 (model 4) for independents. Again, estimators are significant only in the major segment.

We ran additional secondary market regressions weighting observations by films' (log) number of offers at Amazon Marketplace. This analysis puts greater emphasis on films with larger trade volumes such that we measure minimum prices with increased precision. Weighting, in other words, reduces the bias imprecise price measurement might exert on our results. Our findings remain similar to those reported in Table 4.

Finally, one might think of alternative strategies to estimate familiarity's impact on spanning effects. One possibility, exemplified by Jensen (2010), includes differentiation of movies based on star power, as "film stars grant status to films [...] which makes it more likely that they are perceived as normatively legitimate" (Jensen 2010, 43). This strategy, however, neglects that stars carefully screen film projects before accepting roles (DeVany 2004) biasing treatment assignment toward promising and highly budgeted pictures. Similar limitations apply if one defines familiarity on the grounds of opening week attendance and estimates its interaction with category spanning solely based on primary market data. This strategy ignores familiarity garnered in later stages of the theatrical run and, perhaps more importantly, distorts familiarity measurement toward manageable aspects of the film business such as a film's number of opening theaters, advertising, budget and, again, the employment of stars.

Taking into account the "nobody knows principle" (Caves 2000; Goldman 1983), our approach benefits from the near-randomness of ultimate commercial success in cultural markets: "That is, producers and executives know a great deal about what has succeeded commercially in the past and constantly seek to extrapolate that knowledge to new projects. But their ability to predict at an early stage the commercial success of a new film project is almost non-existent" (Caves 2000, 371). Of course, even if one accepts that commercial success relies heavily on randomness, it seems plausible that, at least on average, "better" films attract larger audiences. To secure a familiarity measurement free from quality distortions, our models control for audience ratings of film appeal.

Nonetheless, we tried to reproduce Jensen's (2010) alternative strategy estimating familiarity's influence on spanning effects based on actors' star power. We find no such interaction effect indicating that stardom does not work as a substitute for primary market success in the home video market. Finally, we defined familiarity on the grounds of opening-week attendance and estimated its interaction with category spanning based on primary market data (the dependent variable is then a film's remaining box-office revenue). This approach reproduces our finding in that opening-week attendance $\times$ category spanning is positive for majors, but negative for independents.

References

Caves, Richard E. 2000. *Creative Industries: Contracts Between Art and Commerce*. Cambridge, MA: Harvard University Press.

DeVany, Arthur. 2004. *Hollywood Economics: How Extreme Uncertainty Shapes the Film Industry*. New York, NY: Routledge.

Goldman, William. 1983. *Adventures in the Screen Trade*. New York, NY: Warner Books.

Hirsch, Paul M. 1972. "Processing Fads and Fashions: An Organization-set Analysis of Cultural Industry Systems." *American Journal of Sociology* 77(4):639-59.

Hsu, Greta, Michael T. Hannan, and Özgecan Koçak. 2009. "Multiple Category Memberships in Markets: An Integrative Theory and Two Empirical Tests." *American Sociological Review* 74(1):150-69.

Jensen, Michael. 2010. "Legitimizing Illegitimacy: How Creating Market Identity Legitimizes Illegitimate Products." *Research in the Sociology of Organizations* 31:39-80.

Zhao, Eric Yanfei, Masakazu Ishihara, and Michael Lounsbury. 2013. "Overcoming the Illegitimacy Discount: Cultural Entrepreneurship in the US Feature Film Industry." *Organization Studies* 34(12):1747-76.

Zuckerman, Ezra W. and Tai-Young Kim. 2003. "The Critical Trade-Off: Identity Assignment and Box-office Success in the Feature Film Industry. *Industrial and Corporate Change* 12(1):27-67.