Preliminary: Please do not quote or cite without permission

Product Differentiation and Film-Programming Choice: Measuring Product Similarity in a Dynamic Panel Data Setting

Darlene C. Chisholm
Department of Economics
Suffolk University
Boston, MA 02108

George Norman
Department of Economics
Tufts University
Medford, MA 02155

November 12, 2003

Abstract

This paper proposes a general empirical measure of the degree of product differentiation between two products. The robustness of this measure is tested using a rich new panel data set on film-programming choice in a major U.S. metropolitan motion-pictures exhibition market. The degree of similarity between two products is captured by the extent to which the two products share the same attributes. We propose a similarity index equal to one for cases in which the attributes of two products are measurably identical, equal to zero when the set of common attributes is empty, and equal to values ranging between zero and one as the degree of product similarity moves from completely dissimilar to perfectly identical. Further, we propose a model of strategic interaction between and among firms that leads to predictions of the factors that will influence the similarity index. We apply the model and predictions to the choice motion-pictures exhibitors make when deciding on weekly film offerings at their theatres, relative to the expected equilibrium choices of their competitors sharing the market. Our dataset covers a year's worth of weekly movie offerings and screening times between 2000 and 2001 in the Boston metropolitan market. We find significant evidence of stability in the degree of product differentiation within specific theatre pairs over time. Discrepancies in relative market power, in percentage terms, lead to increased differentiation. The further from a holiday the particular week under study, the more differentiated the film offerings are for a given theatre pair. Finally, we explore the significance of ownership patterns in the context of strategic product differentiation and propose an alternative similarity index to capture differences in physical capacity and thus strategic market power.

JEL Classifications: L11, C33, L82

We are grateful to the DeSantis Center for Motion Picture Industry Studies of Florida Atlantic University College of Business and Economics for providing a grant to fund this research and to Synergy Retail for compiling the dataset used in this paper. We thank Sanjiv Jaggia for invaluable discussions and suggestions and are grateful to Vicky Huang, Anna Kumysh, and Vidisha Vachharajani for their careful research assistance and to Sorin Codreanu for his data-management programming assistance.

I Introduction

The seminal analysis by Hotelling (1929) on "Stability in Competition" has sparked a large and growing theoretical literature concerned with the following deceptively simple questions. Should we expect to find, as Hotelling claimed, that firms offering horizontally differentiated products choose product designs with minimum differentiation? Or, should we expect to find, as D'Aspremont *et al.* (1979) claim in their critique of Hotelling, that these same firms seek maximal differentiate in their product designs in order to soften competition? The simple answer to these theoretical questions, as we indicate in more detail below, is "it depends". Specifically, it depends upon the specific modeling assumptions that we make.

When theory is inconclusive, we might hope that empirical or experimental investigation would shed light on the circumstances that are most conducive to minimum or maximal differentiation. Unfortunately, there are almost no such investigations available. Significant exceptions are Borenstein and Netz (1999) and Netz and Taylor (2002). In both cases, however, the area of application is essentially spatial rather than one of product design.

Our paper examines a type of product differentiation in the movie industry – the weekly program choice by the first-run movie theatres within a well-defined geographic area – that allows us to address specific product design choices and how these choices relate to theoretical analysis. We can think of first-run movie theatres as offering a product with multiple characteristics – the number of different movies being shown and the number of screenings of these movies. On this basis, movie theatre i is less differentiated from movie theatre j in a particular period the more movies, or screenings, they have in common in that period.

Our analysis allows us to shed light on two important questions. First, do first-run movie theatres that are in more direct competition with each other adopt product designs – movie selections – that are more, or less similar? Secondly, to what extent does ownership matter in product design? It is to be expected that a multiproduct firm will make different design decisions from those of single-product non-cooperative oligopolists.[1] Full coordination of product design choice by a particular multiproduct firm is unlikely to be feasible, however, since this firm is in competition with other single and multiproduct firms.

A novel feature of our data is that they allow us to examine the dynamics of movie selection and, in particular, the effect of the specific contractual system within which the movie theatres operate. The major studios typically release what they hope to be "blockbuster" movies close to important holidays and negotiate with the distributors and exhibitors to secure extensive coordinated release of these movies: the recent release of the final film in the "Matrix" trilogy is just one case in point. This implies that we should expect to find greater similarity in movie selection nearer to major holidays but, if there is indeed, a desire to differentiate to at least some extent, this similarity should be expected to decline as we move further from such holidays.

In the next section we develop our measure of similarity in product design. Section III describes the data, their sources and presents some summary statistics and our empirical analysis. Section IV provides concluding remarks.

II Measuring Product Differentiation in a Strategic Market Setting

The dependent variable throughout most of our analysis is the degree of product similarity between pairs of first-run movie theatres. Specifically, we develop a measure, S_{ij}, of the degree to which the movie selection at theatre i is similar to that at theatre j. Two such measures suggest themselves, both of which we consider in the analysis: similarity in the films being shown in the two theatres, ignoring the number of screenings of a specific film, or similarity in screenings, on the basis that a theatre with eight screenings of a particular film is not the same as a theatre with six screenings. More generally, our metric can be applied to measure the degree of similarity between pairs of multi-attribute products.

A complication in measuring similarity in our specific context is that the number of screens at a particular theatre affects the characteristic (movie) mix that the theatre can offer. This suggests the following approach for any pair of theatres *i* and *j*. Begin by determining the movies that are playing at each theatre on a particular day: given our data sources, we chose the Friday of each week. This is used to generate a similarity metric:

$$S_{ij} = \left(\frac{M_c^2}{\overline{M}_i \cdot \overline{M}_j} \right)^{\beta} \tag{1}$$

where M_c is the number of movies the two theatres have in common, $\overline{M}_h$ is the number of screens at movie theatre $h = i, j$ and β is a parameter greater than 0.

An obvious limitation of S_{ij} is that, while it is distributed on the interval [0, 1], it is concatenated in this interval whenever $\overline{M}_i \neq \overline{M}_j$. To correct for this potential bias we normalize S_{ij} by the maximum degree of similarity $\overline{S}_{ij} = \left(\left(\min\left(\overline{M}_i, \overline{M}_j \right) \right)^2 / \overline{M}_i \cdot \overline{M}_j \right)^{\beta}$ to give the similarity measure:

$$S_{M,ij} = \left(\frac{M_c}{\min\left(\overline{M}_i \cdot \overline{M}_j \right)} \right)^{2\beta} \tag{2}$$

[1] See, for example, Chisholm and Norman (2003) and Norman and Pepall (2001).

As we noted above, measuring similarity in movies shown ignores the possibility of there being differences in the number of times that a particular movie is shown and whether or not a movie is shown on multiple screens in one theatre but not in another. This suggests that we develop an analogous measure using the total number of showings of each film that the two theatres have in common, rather than the total number of films they have in common. That is, for each film playing at theatre *i*, we determine if that film is also playing at theatre *j*. If so, and if the film is playing three times at theatre *i* and four times at theatre *j*, the number-of-showings matches for this film is three. This number is then added to all of the other number-of-showings matches for all other common films across both theatres to derive the total number of showings in common, S_c. We use these counts to derive the alternative normalized measure of similarity:

$$S_{S,ij} = \left(\frac{S_c}{\min\left(\bar{S}_i \cdot \bar{S}_j\right)} \right)^{2\beta} \qquad (2)$$

where now $\bar{S}_h$ is the number of showings that is possible at theatre $h = i, j$.

The similarity indices (1) and (2) can be thought of as count measures, reflecting the number of "successes" (or matches) the two products mutually possess, relative to the maximum potential for success (or matches). This formulation of the similarity index suggests an underlying binomial process, which motivates the logistic estimations presented in the empirical analysis of Section IV. It further suggests that in our analysis we should confine our attention to β = 1/2.

We implement these similarity indices using data from the first-run exhibition market in the Boston metropolitan area. The market includes 13 theatres in and around Boston: see Figure 1 for their locations. For each theatre, for each week from June 30, 2000 through the week of

June 22, 2001, we have information from Nielsen EDI on which films were playing, and on the revenues generated by each film for that week. We supplemented these data by recording screening times on the Friday of each week, for each film, for each theatre, for each week of our data set. Screening-time information was determined by reviewing *Boston Globe* movie advertisements on microfilm. This screening information is the basis for constructing the similarity indices.

During this time period no first-run theatres in this market opened or closed. Thus we can treat the spatial structure of the market as essentially constant throughout the period of study. Further, when we examine the theatre *i* and theatre *j* pairs using panel-data techniques, we work with a balanced data set.[2]

III. Empirical Analysis

We apply panel-data techniques to estimate equations of the following general form, where $S_{K,ijt}$ represents the normalized similarity index of interest, where $K=M, S$:

$$S_{K,ijt} = \alpha + x_{ijt}\beta + \nu_{ij} + \varepsilon_{ijt} \qquad (3)$$

Note that in this formulation, each theatre pair is treated as the *ij*th group.

Equation 3 states that the similarity index for a given theatre pair, on a given week, will be a function of strategic factors, captured in general by a vector of independent variables that vary over time within this theatre-pair relationship. Further, we allow for the possibility that the similarity measure will be affected by time-invariant effects ν_{ij} specific to each *ij* pair. Finally, ε_{ijt} is the usual disturbance term.

[2] One first-run theatre in Quincy advertised in the *Boston Globe* from June 30, 2000 through September 28, 2000, then did not advertise for the remainder of the time period. We have excluded Quincy from our analysis under the assumption that it belongs to a market south of Boston and thus it is not reasonable to treat it as being in competition with the theatres that were closer to Boston and advertised in the same medium.

Fixed and Random Effects.

In order to estimate (3) we must determine whether the equation should be estimated using fixed- or random-effects estimation techniques. Fundamental to this choice is whether or not it is appropriate to treat ν_{ij} as a random variable. If we think that our data set essentially reflects the population of interest to our study, then the fixed-effects approach is appropriate, since the results will be derived conditional on the observations within our data set.[3] Alternatively, if the data comprise a relatively small sample from a relatively large population, then, conceptually, the random-effects model would be the more appropriate model.[4] While we examine data on every first-run theatre in the Boston market, for virtually every film, and every screening time, during our entire sample period, we can think of our estimation as capturing population parameters relating to U.S. metropolitan motion-pictures exhibition markets in general, with the Boston market representing a relatively small portion of the entire national market. The nature of our data then suggests that indeed the random-effects approach is the more suitable of the two. We apply the Hausman specification test to determine the appropriateness of the random-effects model by, in particular, checking for correlation between ν_{ij} and the independent variables in x_{ijt}.

Factors Influencing the Similarity Index.

We turn now to the specific strategic and institutional factors that influence the degree of similarity of film offerings between two theatres. In doing so, we distinguish between time-invariant effects that are likely to affect similarity in programming across weeks and time-variant effects that are likely to affect the dynamics of programming choice.

---

[3] See Kennedy (1998), p. 227.
[4] *Ibid.*

Given the nature of our data there are three obvious time-invariant effects that can be expected to influence the degree of similarity in movie selection for each theatre pair. First, theatres that are located more closely to each other are likely to be in competition much more directly than those that are geographically separated. In order to test for this effect we measure $DISTANCE_{ij}$, the distance in miles between theatre *i* and theatre *j*. Our expectation is that more closely related theatres will seek to differentiate themselves in order to soften competition for customers.[5] Thus we expect the similarity index to increase with $DISTANCE_{ij}$.

Second, we should expect to find that "ownership matters" but in this case precisely how is not clear *a priori*. To capture the different incentives that might arise when two theatres are owned by the same company, we create the dummy variable $SAMEOWN_{ij}$, which equals 1 if theatre *i* and theatre *j* are owned by the same company, and 0 otherwise. If companies negotiate better contracting terms with distributors when movies are acquired in bulk, we would expect $SAMEOWN_{ij}$ to be positive. Similarly, if programming decisions are centralized and affected by the "center's" reading of the market, we would expect $SAMEOWN_{ij}$ to be positive. By contrast, if programming decisions are centralized and dominated by the desire to avoid direct competition between theatres under common ownership we would expect $SAMEOWN_{ij}$ to be negative. Finally, if individual theatres behave autonomously, with inter-firm competitive forces dominating programming choice, and with few economies from large-scale distribution contracts, then $SAMEOWN_{ij}$ should be have little or no effect.

Third, it is perhaps to be expected that programming choice will be affected at least in part by demographics, to the extent that movie-going choices differ by the precise characteristics

[5] This is consistent with Irmen and Thisse (1998) who conclude that in a multi-characteristic Hotelling space firms seek to differentiate themselves in one characteristic.

of the movie-going population "close to" and thus within the natural catchment area of a particular movie theatre. As a result, we test for the importance of a number of demographic variables, each measured within a five-mile radius of the particular movie theatre.

We noted in the introduction that the contractual context in which the movie theatres operate is likely to affect programming choice. In particular, since our focus is on first-run theatres in a major metropolitan area, we should expect similarity in programming choice to be greatest in the vicinity of major holidays when many of the theatre owners are contractually committed to allocate multiple screens to the typical holiday films.

One approach to measuring this effect would be to create a dummy variable dependent upon whether or not a particular week is "close to" a holiday. Given that contractual commitments override individual strategic considerations, however, we might expect to find that similarity decreases more smoothly with "distance" from such holidays. We test for this effect by $HOLIDAYDISTANCE_{ijt}$, defined as the number of weeks the current week is away from the nearest holiday. If the current week is a holiday week, $HOLIDAYDISTANCE_{ijt}$ equals zero. If the current week is between two holidays, the total number of weeks between the two holidays is divided in half. As a result, *HOLIDAYDISTANCE* increases with the number of weeks away from the first holiday until it reaches the half-way point between the two holidays, then declines incrementally until it reaches zero again at the next holiday. We use Memorial Day, the Fourth of July, Thanksgiving, and Christmas as the holidays in our sample, to correspond to the historical importance of these major holidays to revenue generation for motion-pictures exhibitors. Our general expectation is that theatres will offer more similar product choices closer to holidays.

Changes in programming at a particular theatre in a particular week are likely to be determined in part by the previous week's revenues at the theatre. By the same argument, we are likely to find that the evolution in programming similarity between any theatre pair will be affected by differences in their revenues. This effect is measured by $\%REVDIFF_{ijt\text{-}1}$, the difference in total weekly revenue between the two theatres during the previous week, divided by the average weekly revenue generated by the two theatres during the previous week. We anticipate that larger differences in market share and revenue generation will be related to smaller degrees of similarity across the two theatres.

There is likely to be some degree of inertia in programming choice. If two theatres were similar last week they will be similar this week, if they were similar last week they will have been similar the week before, and so on. As with our holiday measure, however, this inertia will be offset by strategic considerations that lead theatres to try to differentiate themselves. The stronger are the strategic considerations the shorter will be the period over which inertia in programming is likely to be important. This leads us to introduce lagged values of the dependent variable.

Finally, in order to allow the dependent variable to range from negative to positive infinity, we perform a log-odds transformation of the similarity index, with appropriate adjustments for the case of an index value equal to zero or one. To summarize, we estimate the reduced form equation:

$$Log(S_{K,ijt}/1\text{-}S_{K,ijt}) = \alpha + \beta_1 DISTANCE_{ij} + \beta_2 SAMEOWN_{ij} + \beta_3 \%INCDIF_{ij} + \beta_4 \%REVDIFF_{ij,t\text{-}1} + \beta_5 \%HOLIDAYDIST_{ijt} + \beta_6 LOGODDS_{K,ijt\text{-}1} + \beta_7 LOGODDS_{K,ijt\text{-}2} + \beta_8 LOGODDS_{K,ijt\text{-}3} + \nu_{ij} + \varepsilon_{ijt} \quad (4)$$

We estimate (4) with random effects, corrected for serial correlation, using both the movie-count index and the showings-count index. The results for the showing count index are reported in Table 1, Regression I. Before discussing these results we consider an alternative, measure of similarity between two theatres.

Our measures of similarity in (1) and (2) are symmetric: theatre *i* is as similar to theatre *j* as *j* is to *i*. An argument can be made, however, in favor of the idea that the degree of similarity between any two theatres could be asymmetric: theatre *i*, for example, could be "more like" theatre *j* than *j* is to *i*. The implication is that we should allow for the possibility that within a given *ij* pair, the perception of degree of similarity will differ depending on the strategic position each player has relative to the other. To see why, consider the following simple example. Suppose that theatre *i* is a 5-plex and theatre *j* is a 20-plex, that all five films showing at *i* are also showing at *j* but that the remaining films showing at *j* are not showing at *i*. It could be argued that the 5-plex perceives the two theatres to be 100% similar whereas the 20-plex considers them to be only 20% similar. This leads us to propose the following measure, from the point of view of theatre *i*, when theatre *i* is considering its position relative to theatre *j*.

$$S_{AM,ij} = M_{cij}/M_i \qquad (5)$$

$$S_{AS,ij} = S_{cij}/S_i \qquad (6)$$

The numerator in each index counts the number of matches (movies or showings) between theatre *i* and j, and the denominator limits the total number of matches (movies or showings) to the capacity of theatre *i*. These alternative measures of similarity get to the heart of the importance of *perceived* market interdependence in strategic decision making. Equation 4 was re-estimated, using the log-odds of these two alternative measures of similarity. The results are presented in Table 1, Regression II.

Whether the symmetric or asymmetric index of similarity is used, the qualitative results of both models are similar. Movie theatres that are located more closely to each other tend to be less similar in their program selection, consistent with an attempt by these theatres to soften competition between them. It is also clear that ownership does, indeed, matter, but we noted above that the expected sign of SAMEOWN is ambiguous. This is consistent with programming decisions being centralized rather than decentralized to the individual theatres.

As expected, program selection is more similar the nearer we are to a major holiday. This result is consistent with the industry pattern of wide release of holiday films expected to be blockbusters, followed by more limited releases of a larger number of films expected to succeed in niche markets. The former "force" similarity while the latter gives more play to strategic, theatre-specific effects.

Larger differences in last period's revenue lead to smaller degrees of similarity between two theatres. This result suggests that dominant firms choose fundamentally different product attributes relative to smaller players.

Finally, we find that there is, indeed, some inertia in the program selection but that the significance of the lagged dependent variable declines rapidly. This is consistent, once again, with there being important strategic effects at work in program selection, encouraging individual theatres to seek some degree of individuality in the movie selection that they choose.

Grouped Logit Estimation.

The log-odds transformation in (4) served the purpose of ensuring that the dependent variable falls in the range (-∞, +∞) rather than [0, 1]. Two criticisms of the log-odds transformation are that somewhat arbitrary adjustments must be made for 0 and 1 values of the "raw" similarity index and that the interpretations of the estimated coefficients are indirect (Papke and Wooldridge (1996)). On the first point, since the current data set involves relatively few similarity values of 0 and 1, the results are not significantly impacted by the adjustments made for these observations. With respect to the second point, estimating (4) still produces the correct standard errors for the stated specification, thus we can say something meaningful about the significance of the sign of the coefficient; the coefficient itself, however, does not offer a direct predictive interpretation (that is, we cannot, from the coefficient, determine directly by how much the degree of similarity between two theatres' offerings will increase per increase in *DISTANCE* of one mile, for example.)

An alternative specification of the model presented in (4) would take advantage of the proportional nature of the similarity index, while at the same time producing coefficients with direct interpretations and preserving index values of zero and one. To illustrate the point, consider a subsample of the current data set, comprising each theatre pair from a particular period, t. The similarity index for a given theatre pair measures the number of successes (matches), M_c or S_c, relative to the total possible number of successes (matches). The index can thus be thought of as a proportion, drawn from a population of size N, where N represents the minimum of movies (showings) between theatre i and theatre j.

In the case where we are comparing movie matches, consider creating N new observations for a given theatre pair, M_c of which have a dependent variable equal to 1, $N - M_c$

of which have a dependent variable equal to 0, all of which have the same values of the independent variable for that theatre pair in that period. If we thus expand each observation in this period, we will have effectively transformed the proportional dependent variable into a dichotomous choice variable, with a data set made up of generated observations, and with a total sample size equivalent to the sum of the *N* values across all theatre *i* and theatre *j* pairs. The logit model can then be used to estimate the following equation, amounting to a grouped-logit analysis of the theatre pairs in period *t*:

$$Y_{ij} = \alpha + \beta_1 DISTANCE_{ij} + \beta_2 SAMEOWN_{ij} + \beta_3 \%INCDIF_{ij} + \beta_4 \%REVDIFF_{ijt-1}$$
$$+ \beta_5 S_{K,ijt-1} + \beta_6 S_{K,ijt-2} + \beta_7 S_{K,ijt-3} + \varepsilon_{ij} \qquad (7)$$

where Y_i equals 1 or 0 in accordance with the weights implied by the population from which the movie (showings) matches are drawn. Note that we can now include lagged values of the similarity index directly, rather than the log odds of these values, for a direct interpretation of the impact a change in last period's degree of similarity will have on this period's degree of similarity.

The grouped-logit estimation implied by (7) was implemented using the symmetric similarity index (Table 2) and the asymmetric similarity index (Table 3). The results in each table contain two typical one-period analyses for non-holiday weeks (Regressions I and II) and for holiday weeks (Regressions III and IV). Complete cross-sectional grouped-logit estimations for all periods for the symmetric similarity index are summarized and presented in Appendix 1 for reference.[6] Qualitatively, the results presented in Tables 2 and 3 mirror those found in the

[6] While the dataset covers 52 periods, we report the results for only 49 periods, starting with period four, since the thrice-lagged dependent variable requires omitting the first three periods under study.

panel-data random-effects estimation presented in Table 1. A notable exception is that *%REVDIF*$_{-1}$ is consistently positive and often significant in the cross-sectional analyses for both the symmetric and asymmetric indices, but is significant and negative in the panel-data analysis for the asymmetric index.

IV. Conclusion

An avenue for future research includes modifying either the existing grouped-logit or fractional-response estimation techniques (Papke and Wooldridge (1996)) to accommodate a panel-data setting. The present analysis, which combines the log-odds panel-data analysis with the grouped-logit cross-sectional analysis, does provide a robust set of results indicating that spatial attributes, relative ownership status, market position, and inertia influence the degree of product differentiation in a dynamic setting of product re-design.

Table 1. Random-Effects Panel Estimation of Showing-Count Similarity Index: Symmetric and Asymmetric Indices

Variable	(I)	(II)
CONSTANT	.250772	-.133252
	(0.50)	(-0.38)
DISTANCE	.0514482	.0329668
	(2.88)***	(2.59)***
SAMEOWN	.9748719	.6318012
	(2.11)**	(1.93)*
%INCDIF	3.046961	2.226124
	(1.90)*	(1.95)*
$\%REVDIF_{-1}$	-.0780074	-.1562977
	(-0.49)	(-1.91)*
HOLIDAYDIST	-.065884	-.0383179
	(-4.20)***	(-4.80)***
$LOGODDS_{-1}$	.2003028	.1999146
	(12.38)***	(17.46)***
$LOGODDS_{-2}$	.1036258	.1307072
	(6.32)***	(11.31)***
$LOGODDS_{-3}$	.004304	.0178753
	(0.27)	(1.56)
Sample Size	3822	7644
Overall R^2	0.1646	0.1978
Est. Autocorr. Coefficient	.39394103	.3956589

Dependent variable is log odds ratio: $\log(S_S/(1-S_S))$ for the symmetric index (Regression I), and $\log(S_A/(1-S_A))$ for the asymmetric index (Regression II), with appropriate adjustments for index values equal to 0 and 1.
%INCDIF is percentage difference in mean income within 5-mile radii of theatres i and j.
Hausman test for random effects holds with a $\chi^2(5)$ value of 15.94 for (I) and of 9.85 for (II).
Regression corrects for autocorrelation.
Data ranges over 49 of the 52 weeks; three observations dropped due to usage of lagged dependent variable.
Significance levels *.10, **.05, ***.01.

Table 2. Cross-Sectional Grouped Logit Estimation of Showing-Count Symmetric Similarity Index: Two Holiday and Non-Holiday Weeks

	Non-Holiday Weeks		Holiday Weeks	
Variable	(I)	(II)	(III)	(IV)
CONSTANT	-.7288781 (-1.74)*	-1.303285 (-5.95)***	-1.091975 (-3.40)***	-1.132842 (-4.19)***
DISTANCE	.021761 (4.97)***	.0200821 (5.61)***	.0305317 (7.51)***	.0266931 (5.37)***
SAMEOWN	.5120497 (4.25)***	.8028531 (7.44)***	.4913858 (4.40)***	.115657 (0.93)
%INCDIF	1.734704 (4.34)***	1.293083 (3.86)***	1.419497 (3.91)***	.805076 (1.90)*
%REVDIF$_{-1}$	.3903534 (2.57)**	.3507816 (3.26)***	.7082141 (5.17)***	-.0460127 (-0.34)
$S_{S,-1}$	.9138395 (2.62)***	1.432281 (6.39)***	1.398897 (4.91)***	1.562843 (7.32)***
$S_{S,-2}$	.9356307 (2.80)***	.7378169 (3.31)***	.6960205 (2.45)**	1.171095 (5.83)***
$S_{S,-3}$	-.1476309 (-0.41)	-.1965285 (-0.93)	-.4384457 (-1.48)	.2119719 (0.98)
Grouped Logit N	3333	3470	3821	3528
Likelihood Ratio	108.56	254.78	219.47	223.98

Dependent variable is 1 or 0 per grouped-logit data expansion for symmetric similarity index. %INCDIF is percentage difference in mean income within 5-mile radii of theatres i and j. Non-Holiday weeks in (I) and (II) are October 20, 2000 (T=17) and March 23, 2001 (T=39), respectively. Holiday weeks in (III) and (IV) include Christmas (T=26) and Memorial Day (T=47), respectively.
Significance levels *.10, **.05, ***.01; t-ratios in parentheses.

Table 3. Cross-Sectional Grouped Logit Estimation of Showing-Count Asymmetric Similarity Index: Two Holiday and Non-Holiday Weeks

	Non-Holiday Weeks		Holiday Weeks	
Variable	(I)	(II)	(III)	(IV)
CONSTANT	-1.290384 (-9.04)***	-1.417244 (-11.92)***	-1.532859 (-12.43)***	-.97464 (-7.15)***
DISTANCE	.0066757 (2.85)***	.0100643 (4.64)***	.0129306 (5.83)***	.0165828 (6.70)***
SAMEOWN	.2027473 (3.17)***	.4729358 (7.87)***	.2412042 (4.04)***	.0139656 (0.22)
%INCDIF	.8267645 (3.74)***	.8664944 (4.28)***	.7606852 (3.73)***	.3268613 (1.46)
%REVDIF$_{-1}$	.3259091 (3.75)***	.2918246 (4.16)***	.2061402 (2.90)***	.1901133 (2.82)***
$S_{S,-1}$	2.513874 (15.89)***	2.03796 (14.86)***	2.745234 (19.37)***	2.208551 (16.08)***
$S_{S,-2}$	-.358202 (-2.20)**	.271085 (2.09)**	-.1043391 (-0.68)	-.2333755 (-1.57)
$S_{S,-3}$	.3279648 (2.32)**	-.133972 (-1.11)	.0566857 (0.42)	.0735626 (0.55)
Grouped Logit N	6666	6940	7642	6436
Likelihood Ratio	418.42	522.06	702.02	397.19

Dependent variable is 1 or 0 per grouped-logit data expansion for asymmetric similarity index. %INCDIF is percentage difference in mean income within 5-mile radii of theatres i and j. Non-Holiday weeks in (I) and (II) are October 20, 2000 (T=17) and March 23, 2001 (T=39), respectively. Holiday weeks in (III) and (IV) include Christmas (T=26) and Memorial Day (T=47), respectively.
Significance levels *.10, **.05, ***.01; t-ratios in parentheses.

References

Greene, William H. 2003. *Econometric Analysis,* Fifth Edition. Prentice Hall, Upper Saddle River, New Jersey.

Hsiao, Cheng. 1991 (reprinted). *Analysis of Panel Data,* in Econometric Society Monographs, Editors Jean-Michel Grandmont and Charles F. Manski. Cambridge University Press, Cambridge.

Kennedy, Peter. 1998. *A Guide to Econometrics,* Fourth Edition. MIT Press, Cambridge, Massachusetts.

Maddala, G. S. 1991 (reprinted). *Limited-Dependent and Qualitative Variables in Econometrics,* in Econometric Society Monographs, Editors Jean-Michel Grandmont and Charles F. Manski. Cambridge University Press, Cambridge.

Papke, Leslie E. and Wooldridge, Jeffrey M. 1996. "Econometric Methods for Fractional Response Variables with an Application to 401(k) Plan Participation Rates," *Journal of Applied Econometrics,* Volume 11, pp. 619-632.

Wooldridge, Jeffrey M. 2002. *Econometric Analysis of Cross Section and Panel Data.* MIT Press, Cambridge, Massachusetts.

Appendix 1. Cross-Sectional Grouped Logit Estimation of Showing-Count Symmetric Similarity Index Inclusive of All Periods with Dependent Variable Lags

Period	DIST	SAMEOWN	%INCDIF	$\%REVDIF_{-1}$	$S_{S,-1}$	$S_{S,-2}$	$S_{S,-3}$	N
T=4	.04***	.76***	1.69***	.54***	.93***	.45*	-.08	3917
T=5	.05***	.44***	1.38***	.25*	1.91***	-.18	-.30	3771
T=6	.04***	.26**	1.58***	.61***	1.95***	-.94***	.39	3743
T=7	.03***	.37***	1.56***	.71***	1.25***	-.87**	.58	3654
T=8	.02***	.33***	1.35***	.50***	1.61***	-.15	-1.01**	3593
T=9	.02***	.31***	1.67***	.28**	1.9***	.40	-.71**	3580
T=10	.01***	.40***	1.5***	.75***	.66**	-.24	-.41	3814
T=11	.02***	.25**	1.2***	.25**	1.3***	.03	-.66***	3232
T=12	.02***	.16*	.66**	.29***	1.67***	.44*	-.73***	3163
T=13	.03***	.39***	1.06***	-.11	.96***	.75***	-.13	3289
T=14	.03***	.48***	1.25***	.00	1.47***	.91***	-.45**	3276
T=15	.02***	.19*	.98**	-.18*	1.74***	.21	-.29	3399
T=16	.03***	.15	1.56***	-.22	1.40***	.24	-.42	3415
T=17	.02***	.51***	1.73***	.39**	.91***	.94***	-.15	3333
T=18	.02***	.40***	1.24***	.59***	2.07***	.36	-.34	3388
T=19	.01*	.38***	.74**	-.08	2.23***	.37	.21	3514
T=20	.02***	.72***	1.12***	-.04	.82***	.58**	-.44	3537
T=21	**.04*** **	**.61*** **	**1.95*** **	**.20**	**.55* **	**.04**	**-.13**	**4531**
T=22	.03***	.45***	2.12***	.13	1.5***	.06	.46	3905
T=23	.03***	.41***	1.77***	.18	1.2***	.23	.19	3605
T=24	.02***	.33**	1.03**	.03	2.08***	.59	-.17	3458
T=25	.03***	.36***	.189***	.14	1.72***	.98***	.26	3560
T=26	**.03*** **	**.49*** **	**1.42*** **	**.71*** **	**1.40*** **	**.70** **	**-.44**	**3821**
T=27	.02***	.40***	1.34***	.54***	1.90***	.44	-.33	3709
T=28	.02***	.16	1.57***	.47***	1.26***	.74**	-1.49***	3412
T=30	.03***	.35***	1.45***	.65***	.81***	-.49	-1.48***	3409
T=31	.03***	.34***	1.44***	.47***	.79***	-.32	-1.52***	3490
T=32	.02***	.44***	.35	.50***	.52*	.21	.89***	3566
T=33	.02***	.45***	.07	.08	.89*	.35	-.58	3454
T=34	.02***	.55***	.73**	.43***	1.11***	.13	.18	3490
T=35	.02***	.63***	.94***	-.04	.86***	.53*	.35	3484
T=36	.02***	.67***	1.11***	.11	-.39	.20	-.50	3478
T=37	.03***	.65***	1.38***	.30***	.59**	.90***	-.69**	3430
T=38	.03***	.43***	1.58***	.21*	1.36**	.53**	-.14	3462
T=39	.02***	.80***	1.29***	.35***	1.43***	.74***	-.20	3470
T=40	.01***	.73***	.87***	-.10	1.02***	1.00***	-.59**	3484
T=41	.02***	.66***	1.10***	.12	1.35***	.55**	-.26	4461
T=42	.01**	.56***	.98***	-.04	1.20***	.72**	.12	3760
T=43	.02***	.55***	.47	.36***	1.06***	1.03***	-.44*	3780
T=44	.03***	.22**	1.25***	.59***	1.62***	.60**	-1.19***	3688

Appendix 1 Continued. Cross-Sectional Grouped Logit Estimation of Showing-Count Symmetric Similarity Index Inclusive of All Periods with Dependent Variable Lags								
Period	DIST	SAMEOWN	%INCDIF	$\%REVDIF_{-1}$	$S_{S,-1}$	$S_{S,-2}$	$S_{S,-3}$	N
T=45	.03**	.08	1.08***	.15	1.51***	1.19***	-1.25***	3518
T=46	.02***	.90	1.56***	.41***	1.22***	.80***	-.94***	3449
T=47	**.03*** **	**.12**	**.81***	**-.05**	**1.56*** **	**1.17** **	**.21**	**3528**
T=48	.04***	.18	1.14***	-.08	1.42***	-.07	-.14	3214
T=49	.04***	.48***	1.21***	.35***	.94***	.50	-.02	3523
T=50	.03**	.77**	.62	.23*	-.15	.54	.30	3521
T=51	.02***	.86***	.69*	-.01	2.23***	.93***	.23	3756
T=52	.01***	.85***	.62	-.07	2.72***	.42	.41	3936

Dependent variable is S_S. All estimations include constants; coefficient not reported here. N is population size for grouped logit, equal to the minimum number of showings between theatre i and theatre j. %INCDIF is percentage difference in mean income within 5-mile radii of theatres i and j. Holiday weekends are in bold: Thanksgiving (T=21); Christmas (T=26); and Memorial Day (T=47). Fourth of July (T=1) is excluded due to presence of lagged variables.
Significance levels *.10, **.05, ***.01.