



3046

REPRINT

Gianpiero Meloni, Dimitri Paolini and Juan de Dios Tena

American beauty: Trade flows and export cots of US movies

Journal of Cultural Economics, 42, 701-716, 2018

CORE
CENTER FOR OPERATIONS RESEARCH AND ECONOMETRICS

CORE

Voie du Roman Pays 34, L1.03.01

B-1348 Louvain-la-Neuve

Tel (32 10) 47 43 04

Email:lidam-library@uclouvain.be

https://uclouvain.be/en/research-institutes/lidam/core/reprints.html

J Cult Econ (2018) 42:701–716
https://doi.org/10.1007/s10824-018-9324-6

ORIGINAL ARTICLE

American beauty: trade flows and export costs of US movies

Gianpiero Meloni[1,2] · **Dimitri Paolini**[1,2,3] · **J. D. Tena**[1,2,4]

Received: 1 April 2017 / Accepted: 5 April 2018 / Published online: 4 May 2018
© Springer Science+Business Media, LLC, part of Springer Nature 2018

Abstract This paper estimates a hedonic model of revenues for 1431 American movies in 45 non-American countries over the period 2002–2013. Such an extensive analysis is proven to be useful to estimate the effect of indicators of geographical and cultural distance on total revenue and the number of American films exported to non-American countries once the model specification also controls for relevant film features. It is found that indicators of cultural and geographical distance, in addition to the Human Development Index, have a significant effect on both the number of American films shown and the revenue obtained. Estimation results have been also used to appraise the importance of film features in different groups of countries from the perspective of their degree of development and distance from the USA.

Keywords American movies · Export · Geographical and cultural distance

JEL Classification Z1 · F14 · L82

1 Introduction

The American motion picture industry is of high economic interest for researchers in international trade and decision makers within the industry. For the first group, the industry offers relevant information on trade patterns along different nations for

✉ Dimitri Paolini
dpaolini@uniss.it

[1] Libreria Ubik, Cagliari, Italy

[2] CRENoS, Sassari, Italy

[3] CORE, Louvain-la-Neuve, Belgium

[4] Management School, University of Liverpool, Liverpool, UK

an American service product. This facilitates analysis of how international trade varies in terms of country-level variables, trade restrictions, and different film features. Moreover, this investigation is of high cultural importance as it appraises how countries with different cultures and traditions pay attention to American products.

The analysis of American films is also relevant from a decision management perspective. The importance of the industry can be appraised, for example, from information on global box office for all films released around the world that reached $38.3 billion in takings in 2015 (2015 theatrical statistics summary). In such a big industry, it is highly relevant to estimate how different film features, such as genre, rating, and sequel, affect consumption preferences for American films.

This paper estimates a hedonic model of revenues for 1431 American movies in 45 non-American countries over the period 2002–2013. Hedonic models have been used in the motion picture literature to estimate the contribution of different film features to revenues; see Walls and McKenzie (2012). Here, this basic framework is augmented by considering the impact of relevant country-level variables for a comprehensive group of nations. Such an extensive analysis is proven to be useful in estimating how indicators of geographical distance (GD) and cultural distance (CD) affect total revenue and the number of American films exported to non-American countries, once the model specification also controls for relevant film features. The literature on international trade, which will be more specifically outlined in the next section,[1] has typically studied these issues at the aggregate country-level base. However, among the problems of this type of aggregate study, it is noteworthy that failing to develop an analysis at the film level could result in a Simpson's paradox problem, as the important inherent heterogeneity that is unobserved at the aggregate level is not taken into consideration. This is a typical problem in social science, whereby a characteristic that is present in disaggregate data disappears, or even reverses, for the aggregate counterpart; see Wagner (1982) for a discussion.

To preview, we find that indicators of CD and GD, in addition to the macroeconomic indicator of the Human Development Index (HDI), significantly affect both revenue and the number of American films. The latter is indirectly appraised based on the probability of theatrical release. This result is, in principle, not consistent with Hanson and Xiang (2011) who, in their highly insightful paper based on an extension of the theoretical model proposed by Melitz (2003), indicate that service products, unlike manufacturing products, are dominated by global, rather than bilateral, fixed costs and are more likely to experience adjustment by revenues, while the number of films exported is relatively fixed. However, it is also found that there is not a unique pattern of adjustment for all countries. In particular, countries with low levels of the HDI, there is a negative relationship between the probability of the arrival of a given film and the expected revenue. This result can be explained if we consider that there could be restrictions on the demand for American films in these types of countries. Conversely, there is a positive correlation between the probability of film arrival and expected revenue in developed countries. The analysis also

[1] See, for example, Melitz (2003) and Hanson and Xiang (2011).

appraises the importance of film features in the different groups of countries from the perspective of their degree of development and distance from the USA.

The next section discusses the previous economic literature pertaining to cinema. Section 3 presents the data used in the analysis and provides some relevant descriptive statistics. Section 4 shows and discusses the estimation of the proposed hedonic model. Finally, Sect. 5 concludes the paper.

2 Literature review

Depending on the proposed theory and on the nature of the database used in the empirical analysis, economic papers related to the cinema industry, and, in general, to any service product, can be broken down into at least two main strands of the literature: (1) international trade and (2) industrial analysis of the different film features to explain success.

The first group of papers focuses on analyzing how macroeconomic indicators, different measures of distance, and trade restrictions affect exports of service products. As suggested by Bernard and Jensen (1999) and Melitz (2003), a general paradigm for manufacturing goods holds that, if firms are heterogeneous, the presence of high export costs would imply that only the most productive companies find that it is profitable to sell manufacturing goods abroad. This is consistent with the empirical evidence in Bernard et al. (2007) and Helpman et al. (2008), among others. In this context, Marvasti and Canterbery (2005) estimate a gravity-iceberg model of US export of motion pictures and find that spatial distance, together with other economic and cultural variables, strongly influences the US motion picture trade.

Hanson and Xiang (2011) emphasize the differences between manufacturing and the service trade by extending the model in Melitz (2003) with both global and bilateral fixed export costs. According to their framework, in an industry where global costs are more important than are bilateral costs, the ratio of import sales to product variety would decrease with trade barriers. On the other hand, if bilateral costs dominate global costs, the main adjustment would be through output variety. Hanson and Xiang (2011) test this hypothesis using a sample of US motion picture exports for 46 countries over the period 1995–2006. Their results suggest that global fixed export costs dominate. The importance of trade restrictions for audio-visual services is also studied by Hellmanzik and Schmitz (2015) by means of a gravity model. They consider an indicator of cultural proximity based on the interest that other countries show in the web content of a country. Unlike Hanson and Xiang (2011), they cannot assess whether the adjustments to trade barriers occur in output variety or total sales. However, they also find that cultural proximity is an important determinant of trade.

Compared to the literature above, an important contribution of this paper is that it considers film-level information, rather than aggregate measures at the country level. This distinction is especially relevant because, as mentioned previously, it allows us to take the different heterogeneity sources observed at the film level that

could lead to misleading results at the aggregate level into account in the estimation.[2] Moreover, the use of disaggregate information at the film level allows us to control for the different country-level shocks in the regression.

The second big strand of the literature analyzes how different film features affect consumer preferences. For example, Lee (2009) estimates how cultural differences influence the reception of US movies in East Asian countries. Although the research is specific to the Asian region, it provides a highly insightful and rich analysis of the impact of different types of nominations and awards on the success of American movies for each country. Fu and Lee (2008) investigate the determinants of audience acceptance of foreign movies in Singapore. Interestingly, rather than focussing on the introduction of American films in other countries, as most of the previous literature does, Fu and Lee analyze the patterns of consumption of movies with different nationalities in Singapore. They find that both cultural proximity and the price of a ticket play an important role in determining movie success. Palsson et al. (2013) study how film rating influences consumer and producer decisions for all domestically distributed films in the USA for the period 1996–2009. They find that a mature rating reduces a movie's box office revenue by 20%. Walls and Mckenzie (2012) study the demand for Hollywood movies in the USA and Canada and in six additional countries: Australia, France, Germany, Mexico, Spain, and the UK. They identify the main determinants of the demand for Hollywood movies in each of these countries and find support for the hypothesis that the supply has been adapted to global demand after a decrease in the relative size of the US movie market.

An important difference between this research and the literature discussed above is that the analysis of consumer preferences for films is extended to 45 countries, appraising the main determinants of consumption patterns, especially how they are influenced by indicators of CD, social distance, and GD. Moreover, having such a big sample allows us to run separate regressions for different groups of countries, depending on their HDI and their CD from the USA. CD is time invariant and HDI can take different values in different years, but they are very stable through time. This allows us to classify countries according to their mean HDI values. This is especially useful in identifying different consumption patterns in different groups of countries.

3 Data

We collect information about 1392 US movies in 45 countries (excluding the US) over the period 2004–2013, in addition to indicators at the country level. However, as it will be explained in the following section, 2 years of variables will be used in the paper with the purpose of model identification reducing the total number of observations to 1152. The panel is unbalanced regarding the first group of variables, as not every American film is shown in each of the countries under analysis. Data on

[2] A remarkable exception to this is Holloway (2014), who, using information at the film level, estimate the impact of quality on the propensity of a given movie entering a foreign market.

box office revenues are available from boxofficemojo.com. In addition, we consider several quantitative and qualitative variables for each of the films in the sample. In particular, production budget (adjusted for inflation, based on 2013) is obtained from numbers.com at the URL http://www.the-numbers.com/movie/budgets/all. The variable *nomination* takes the value 1 if the film has received an Academy Awards' nomination in one or more of the main categories: best movie, best director, best actor or actress in a leading or supporting role, or best animation movie. Movie ratings provide parents with advance information about the content of movies to help them determine what is appropriate for their children. They are used in our sample as a proxy of the content of a movie in terms of violence, sex, etc. We consider the following rating indicators: *G* stands for General Audiences, PG stands for Parental Guidance Suggested, PG-13 stands for Parental Guidance Strongly Cautioned; and *R* stands for Restricted, with Under 17 requiring accompaniment by a parent or adult guardian. Genre information is captured with the following set of binary dummies: *drama*, *comedy*, *action*, and *thriller*. The variables *sequel* and *sequel2* indicate whether the film is either the first sequel or a second or successive sequel of a previously released movie, respectively. Finally, the variable *original* indicates whether the movie is based on the writer's own idea. We gather these data from opusdata.com and imbd.com.

Same basic descriptive statistics of these variables, apart from budget and box office, are all dichotomous variables. Interestingly, budget varies greatly across films, with a minimum value of 0.16 million dollars for the film Once and a maximum of 336.9 million dollars for the film Pirates of the Caribbean 3. All groups of genres and ratings are represented in the sample, although the most common genres are comedy and action, and 80% of the films are either *R* or PG-13. This suggests that the sample contains a big range of different types of films. Sequels make up approximately 10% of the sample, nominated films make up approximately 10%, and original screenplays make up approximately 50%.

Table 1 reports some descriptive statistics for budget and log-box office in each of the 45 nations in our sample. The dynamic evolution of these two variables is not shown for the sake of brevity they show a stable behavior through time. When we turn our attention to country differences, it can be observed that western developed countries such as the UK, Spain, Australia, and Germany are the main importers of American films while developing countries such as China, India, Jamaica, and Kuwait are the ones who import the fewest number of American films.

We consider two alternative country-level indicators to classify countries: the HDI and the CD. The HDI is an index created by the United Nations that summarizes measures of average achievement in key dimensions of human development: standards of living, education, life expectancy, quality of life, and the per capita GDP for each country. This variable could be considered as an indicator of the potential demand for cultural products in a given country.

As discussed in the previous section, the literature on international trade considers CD indicators to account for trade between two countries. To capture information about the CD between the USA and each of the countries in the sample, we follow Lee (2009) and implement a value-based index developed by Hofstede (1980) that is built around four dimensions: (1) power distance, which expresses the degree

Table 1 Countries descriptive statistics: box office

Country	No. of movie	% movies	Mean	SD	Min	Max
Argentina	995	0.70	1,251,744.19	2,185,664.781	1464	23,556,116
Australia	1225	0.86	6,519,565.82	8,429,739.898	1605	114,876,536
Austria	1021	0.71	1,295,698.23	1,710,308.112	2330	15,339,482
Belgium	1008	0.70	1,400,454.83	1,911,476.309	1729	14,744,213
Brazil	947	0.66	4,038,623.61	6,616,694.647	1374	64,863,380
Bulgaria	826	0.58	158,687.73	228,226.804	1343	3,699,024
Chile	898	0.63	653,480.89	1,163,902.986	2808	11,394,242
China	176	0.12	16,771,464.81	26,707,876.414	8599	197,911,296
Colombia	675	0.47	1,063,486.09	1,670,273.202	2067	14,791,967
Croatia	659	0.46	157,600.70	217,043.118	2346	2,740,992
Czech Republic	860	0.60	439,665.29	815,728.457	1635	13,541,869
Denmark	799	0.56	1,281,129.15	2207,207.449	1338	24,857,536
Dominican Republic	134	0.09	120,134.64	165,901.790	2269	1,420,529
Ecuador	611	0.43	393,628.83	582,329.570	1329	5,491,452
Egypt	764	0.53	128,567.28	175,055.301	1526	2,482,628
Estonia	487	0.34	97,227.91	131,733.654	1775	1,313,026
Finland	887	0.62	646,796.70	1,006,001.645	2438	9,016,192
France	1182	0.83	7,835,170.94	12,445,908.565	1350	171,871,504
Germany	1198	0.84	7,706,033.92	13,200,710.352	2332	171,115,344
Greece	924	0.65	997,674.87	1290,114.475	2808	12,775,577
Hungary	819	0.57	500,500.34	661,597.267	2161	7,928,378
India	360	0.25	1,349,806.03	2,778,435.522	2665	26,299,510
Indonesia	404	0.28	1,093,129,40	1,641,125.635	30,073	15,308,150
Israel	530	0.37	898,788.26	1,083,924.519	11,016	14,224,082
Italy	1178	0.82	4,790,454.54	7,019,764.230	1337	90,679,032
Jamaica	151	0.11	75,117.66	101,000.584	1446	671,902
Japan	734	0.51	13,064,651.17	24,489,015.708	9447	243,522,416
Kuwait	151	0.11	281,045.80	256,964.726	9094	1,947,920
Malaysia	633	0.44	1,113,718.56	1,536,392.376	2557	13,909,003
Mexico	1192	0.83	4,885,422.43	7,023,475.080	11,565	63,420,020
Netherlands	1049	0.73	1,839,929.09	2,768,811.902	1150	25,723,286
New Zealand	1093	0.76	980,112.78	1,264,528.786	2382	13,669,993
Norway	935	0.65	1,212,640.31	1,883,928.630	2096	15,774,829
Portugal	970	0.68	807,240.66	973,256.034	1844	10,094,128
Russia	1015	0.71	5,372,056.80	8,556,759.968	5895	127,174,120
Slovakia	602	0.42	154,889.99	265,665.508	1172	3,167,044
Slovenia	620	0.43	149,104.45	176,354.553	2430	1,981,953
South Africa	1056	0.74	738,156.47	851,264.411	2042	8,449,588
South Korea	800	0.56	5,270,132.87	9,576,162.332	1306	114,557,272
Spain	1265	0.88	5,601,730.24	7,409,987.907	1859	119,452,120
Sweden	851	0.59	1,612,282.09	2,757,706.609	1372	30,540,692
Thailand	759	0.53	852,686.85	1,239,324.887	1139	9,138,544

Table 1 (continued)

Country	No. of movie	% movies	Mean	SD	Min	Max
Turkey	931	0.65	976,549.55	1,354,475.709	4523	19,075,024
United Kingdom	1277	0.89	12,945,515.41	19,378,329.138	3899	163,594,016
Uruguay	637	0.45	111,337.73	151,339.844	813	1,298,218

to which less powerful members of a society both accept and expect that power is distributed unequally; (2) uncertainty avoidance, which expresses the degree to which the members of a society feel uncomfortable with uncertainty and ambiguity; (3) individualism versus collectivism; and (4) masculinity versus femininity. We gathered this information from Hofstede (2001), and each country's CD from the USA was computed using Kogut and Singh (1988) formula:

$$\mathrm{CD}_j = \sum_{I=1} \left\{ \left(I_{ij} - I_{iu}\right)^2 / V_i \right\} \tag{1}$$

where CD_j is the CD of country j from the USA, I_{ij} is the value for country j on the ith cultural dimension (I_{iu} for the US), and V_i is the variance of the ith cultural dimension. Finally, to evaluate GD we use data from CEPII's GeoDist database[3] and weighted distances between countries calculated on bilateral distances of the biggest cities of two countries and on those inter-city distances being weighted by the share of the city in the country's overall population.

Table 2 shows CD, GD, and HDI values for each of the nations considered in our analysis.

At it will be discussed later, countries can be clustered according to their relative position to these CD and HDI. Thus, although these three variables, CD, GD and HDI, are highly correlated it is still possible to find developing countries who are relatively close to the American culture as well as developed countries whose cultural values are very different from the American ones.

Using information by countries, it is interesting to observe whether there is some correlation between average log revenues and the log number of movies exhibited in each country. This is shown in Fig. 1. The graph suggests a slightly positive correlation between the two variables. However, it is not possible to draw a sound economic conclusion from this figure as we are mixing a very heterogeneous group of films. A more complete analysis should control for the different film features that can affect both the probability of film arrival and box office revenues.

[3] See notes on CEPII's distance measures: The GeoDist database, Mayer and Zignano (2011).

Table 2 Countries descriptive statistics: CD and HDI

Country	CD	GD	HDI		
			Min HDI	Mean HDI	Max HDI
Group 1: Low CD and low HDI					
Argentina	1.80	9.07	0.75	0.78	0.81
Dominican Republic	2.17	8.10	0.64	0.67	0.70
Hungary	1.06	9.04	0.79	0.82	0.83
India	1.64	9.48	0.46	0.52	0.59
Jamaica	1.87	8.01	0.68	0.71	0.73
South_Africa	0.38	9.58	0.60	0.62	0.66
Group 2: Low CD and high HD					
Australia	0.02	9.60	0.91	0.93	0.94
Austria	1.49	9.00	0.85	0.87	0.89
Belgium	1.52	8.90	0.88	0.89	0.90
Czech Republic	1.08	8.99	0.82	0.86	0.87
Denmark	1.87	8.91	0.87	0.89	0.90
Estonia	1.17	8.97	0.79	0.82	0.84
Finland	1.25	8.94	0.85	0.88	0.89
France	1.52	8.92	0.85	0.88	0.89
Germany	0.47	8.94	0.87	0.90	0.92
Israel	1.73	9.26	0.87	0.88	0.90
Italy	0.58	9.02	0.83	0.86	0.88
Netherlands	1.44	8.89	0.89	0.91	0.92
New Zealand	0.25	9.45	0.89	0.91	0.92
Norway	2.00	8.86	0.92	0.94	0.95
Spain	1.90	8.91	0.85	0.87	0.88
Sweden	2.28	8.91	0.90	0.91	0.92
United Kingdom	0.08	8.84	0.84	0.86	0.89
Group 3: High CD and low HDI					
Brazil	2.32	9.00	0.67	0.70	0.74
Bulgaria	3.27	9.11	0.72	0.76	0.78
Chile	3.97	9.03	0.76	0.79	0.82
China	3.43	9.32	0.59	0.65	0.72
Colombia	3.92	8.35	0.66	0.69	0.72
Croatia	3.02	9.03	0.75	0.79	0.81
Ecuador	4.35	8.48	0.66	0.69	0.72
Egypt	3.31	9.25	0.59	0.63	0.68
Indonesia	3.82	9.65	0.54	0.59	0.68
Kuwait	4.22	9.35	0.78	0.79	0.81
Malaysia	4.26	9.61	0.71	0.74	0.77
Mexico	3.30	7.81	0.72	0.75	0.77
Portugal	4.39	8.85	0.78	0.80	0.82
Russia	4.22	9.10	0.71	0.76	0.78
Thailand	3.40	9.54	0.63	0.67	0.72

Table 2 (continued)

Country	CD	GD	HDI		
			Min HDI	Mean HDI	Max HDI
Turkey	2.65	9.18	0.64	0.69	0.76
Uruguay	3.38	9.10	0.74	0.76	0.79
Group 4: High CD and high HDI					
Greece	3.75	9.14	0.81	0.85	0.87
Japan	2.70	9.24	0.88	0.90	0.91
Slovakia	4.15	9.02	0.79	0.82	0.84
Slovenia	4.40	9.01	0.84	0.87	0.89
South Korea	3.85	9.28	0.84	0.88	0.91

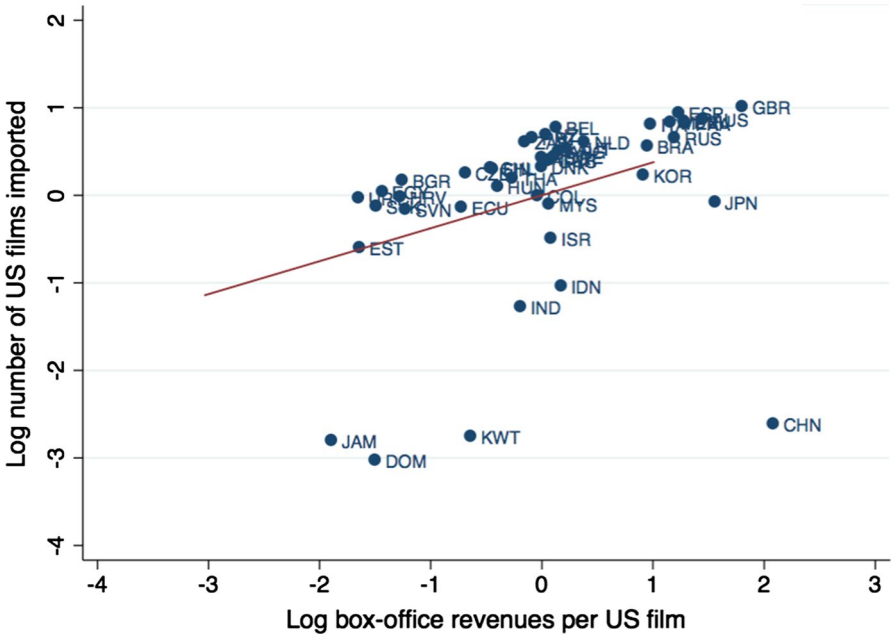


Fig. 1 Intensive versus extensive margin of movie imports, 2004–2013

4 Empirical analysis

We analyze the factors affecting both the revenues from and the arrival of different American films in 45 different countries (excluding the USA) by estimating a hedonic model of revenues; see Walls and McKenzie (2012) and the references therein. A general approach in the literature is to explain film success as a function of production budgets, awards, and different characteristics of the film, such as sequels, genre, and content. This approach is particularly sound when it is applied to

countries other than the USA, because the different explanatory indicators can generally be considered as exogenous with respect to the film revenues in each foreign country, even though they may fail to be exogenous in the American market (due to effects caused by the expected revenue).

As each movie is observed in 1 year only, it is not possible to specify a dynamic panel model. Therefore, our baseline specification explains the revenue of a film in a country j (in logs and adjusted for inflation) as a function of three main groups of variables: (1) indicators for the film quality, the budget, and award nominations, where the former is a measure of the production effort and the latter is a proxy of the artistic quality; (2) variables related to the different features of the film to determine how these features impact its success; and (3) a national trend component. Moreover, the use of a disaggregate approach allows us to control for all the potential unobserved heterogeneity at the country level, such as total production or demand in the national movie industry or the degree of protective legislation for the national industry, by including country/year fixed effects. The following model is considered:

$$\begin{aligned}\ln revenue_{ij} = {} & \beta_0 + \beta_1 \ln budget_i + \beta_2 nomination_i + \beta_3 G_i + \beta_4 PG_i + \beta_5 R_i \\ & + \beta_6 sequel_i + \beta_7 sequel2_i + \beta_8 drama_i + \beta_9 action_i + \beta_{10} thriller_i \\ & + \beta_{11} original_i + \gamma_{jt} + \varepsilon_{ij}\end{aligned} \tag{2}$$

where i, j, and t stand for film, country, and year, respectively; the terms β_r for $r = 1$–11 are parameters of the model, γ_{jt} is the specific country-year fixed effect, and ε_{ij} is an error term. All the variables in the model have been defined in the previous section but note that variables $PG13_i$ and $comedy_i$ are left outside of the age rating and genre groups, respectively, to avoid perfect multicolinearity.

Note also that the model above is particularly suitable for dealing with a potential selection bias problem as the country/year component already controls for the impact of any specific unobserved variable on the arrival of an American film at the country level. However, we explore further the possibility of an endogenous sample selection that could be due to some film-specific shadow costs or features. This may lead to inconsistent estimates of the coefficients in a model that accounts for film revenues if the shocks that affect the probability that a given movie is exhibited in a certain country are highly correlated with the shocks that determine its revenue. Based on this premise, we employ Heckman's (1979) two-step methodology. In the first step, we estimate a probit model (selection equation) for the probability that a film is exhibited in a given country. This approach allows us to obtain the Mills ratios that are necessary to correct the OLS estimates of the primary equation in stage two.

To identify the model, it is necessary to choose at least one variable to be included in the selection equation only at the first stage. However, it is a difficult task to find such a variable, given that the probability of arrival of a given film and its expected revenue are typically affected by similar shocks. Moreover, the selection of one variable to affect the former but not the latter can only be based on speculative arguments. For example, Walls and McKenzie (2012) consider the revenue from opening week in the USA as the identification variable. Although this could

be judged as a plausible option, in principle, the decision as to which countries will show a given film has already been made by the time of the opening week in the USA in most cases. Moreover, this variable is very likely to be highly correlated with the expected revenue of the film, as the specific factors that affect film revenues in the USA are, in principle, like those affecting film revenues in other countries.

As the identification variable, we propose an indicator that takes value one if the most similar movie in terms of the regressors in model (2) 2 years before was exhibited in that country. The concept of most similar movie was applied by adopting the K-nearest neighbor approach; see, for example, Altman (1992). It can be defined as the group of films whose covariates minimize the canonical distance from the film under analysis. Thus, we can identify similar movies across the sample, given their quality and idiosyncratic features, such as genre, content, and sources.

To implement the nearest neighbor, the first 2 years of the dataset (2003 and 2004) are drawn out of the sample and used as a training population to instruct the algorithm, thus reducing the test sample to 1152 movies. We consider a dummy variable that takes the value 1 if the most similar movie was released in the same country in the previous 2 years. We also tested for broader neighbors, $k=3$ and $k=5$, and we did not find any improvement in the quality of our analysis.[4]

The time lag (2 years) is set based on the average time of production of a movie, so that production companies can react properly to the performance of previously released films and strategically choose what types of movies to promote in different markets.

Of course, it can always be argued that these film features, even in other films, could be correlated with the expected revenue of the film. However, it is not true that all the films with the same features, i.e., belonging to the same genre, rating, uniquely determine a high percentage of the variation in revenues. If that were the case, only some specific types of films would be seen in the market. Moreover, an identification variable based on features of other films that were released 2 years before is, in principle, less likely to be correlated with its expected revenue than is an identification variable obtained from the same film.

The first two columns of Table 3 show the estimation results of model (1) in a pool regression for the 45 countries in the sample for both the selection and primary equations. The statistical significance of the coefficient associated with the inverse Mills ratio suggests that it is indeed appropriate to correct for selection bias. Its negative sign is consistent with the assumption that, in some countries, the size of the demand for some specific types of American movies is constrained, which implies that a saturation of the market can occur, thus lowering the expected revenue for each movie exhibited.

To control for the potential correlation of the error term in the primary and the selection equations, we also considered the Mundlak–Chamberlain approach, as proposed by Wooldridge (2010); however, there was no qualitative change in the estimated results. In this paper, for the sake of brevity, we show our baseline

[4] The nearest neighbor was estimated with the R package FNN. Results of this experiment and all the other estimations not explicitly reported in this paper can be obtained from the authors upon request.

Table 3 Determinants of movies revenues

	Estimated parameters			
	(1)		(2)	
	Selection	Primary	Selection	Primary
Nearest neighbor	0.745*** (60.70)		0.677*** (53.64)	
HDI			1.305*** (3.86)	1.552*** (3.71)
Cultural distance			−0.128*** (−27.94)	−0.0408*** (−6.88)
Geographic distance			0.0239 (1.54)	−0.113*** (−6.18)
Budget	0.283*** (35.20)	0.560*** (39.76)	0.297*** (36.11)	0.589*** (41.55)
Nomination	0.480*** (23.12)	0.501*** (21.00)	0.495*** (23.72)	0.535*** (22.17)
Sequel	0.113*** (4.91)	0.392*** (16.96)	0.124*** (5.32)	0.408*** (17.50)
Sequel2	0.361*** (11.02)	0.463*** (17.55)	0.359*** (10.78)	0.471*** (17.84)
Action	0.198*** (11.45)	0.285*** (13.57)	0.200*** (11.43)	0.299*** (14.22)
Drama	−0.0409* (−2.30)	−0.202*** (−8.07)	−0.0418* (−2.34)	−0.205*** (−8.19)
Thriller	0.247*** (13.63)	0.244*** (10.62)	0.256*** (14.03)	0.271*** (11.70)
G	−0.105* (−2.52)	0.158*** (3.99)	−0.106* (−2.53)	0.147*** (3.74)
PG	−0.0885*** (−5.00)	0.0884*** (4.34)	−0.0918*** (−5.16)	0.0818*** (4.02)
R	−0.0370** (−2.63)	−0.0517** (−3.06)	−0.0397** (−2.80)	0.0560*** (−3.31)
Original screenplay	0.0504*** (3.95)	0.0557*** (3.79)	0.0471*** (3.68)	0.0561*** (3.82)
Inverse Mills ratio		−0.228*** (−5.80)		−0.101***(−2.32)
N	51,840	31,710	51,840	31,710
Pseudo R^2	0.147		0.158	
Adjusted R^2		0.307		0.310

Omitted: comedy, PG-13; *t* statistics in parentheses

*$p<0.05$; **$p<0.01$, ***$p<0.001$

specification, which is based on a unique estimated inverse Mills ratio for the entire sample period.

When we look at movie-specific variables, the relevant and positive impact of covariates on quality (production budget and Academy Awards nominations) can be observed in both the probability of exhibition and the revenue. As for genres, a global preference for action and thriller movies emerges, with respect to the reference category of comedies. This result is consistent with the findings of Lee (2009) and the empirical evidence (see The Economist 2011)[5] that American humor does not sell well in markets with a broad CD from the USA. Moreover, the positive

[5] "Hollywood goes global. Bigger abroad." The Economist, Feb 17th 2011.

impact that sequels and subsequent movies in a series have on revenues confirms the validity of Hollywood's strategy of investing in existing intellectual proprieties, rather than introducing new, original products into the market. The validity of this strategy is supported by the small magnitude of the coefficient associated with original screenplays that refer to movies in which the plot and characters are not part of a previous intellectual propriety (e.g., books, comics, or video games).

We augmented the baseline model by including some typical indicators of distance that have also been considered in the literature, in addition to HDI, CD, and geographic distance in both the primary and the selection equations, with the aim of apprising the relative importance of these variables in relation to revenue and the probability of arrival of a film in a non-American country. This regression, which is used to obtain an initial appraisal on the relative importance of the different distance indicators, does not include year-time fixed effects to avoid perfect multicollinearity with HDI. The results of this estimation are shown in columns 3 and 4 of Table 3. As discussed in Sect. 2, according to Hanson and Xiang (2011), distance indicators should exert a more significant impact on the intensive margin (primary equation) than in the number of released US films, which is appraised here based on the selection equation in Table 3. Note in this respect that an important difference with Hanson and Xiang (2011) is that we consider an estimation at the film level taking into consideration the impact of different film features on the probability of theatrical release rather than estimating the total number of films only as a function of country-level variables.

GD is only significant in the primary equation, while CD is a relevant variable in both equations. This result suggests that CD represents a more important bilateral fixed shadow cost of introducing a movie in a given country than GD. In addition, the estimated model suggests that not only distance indicators but also the degree of development of a country are important variables in determining both the number of films exhibited and their expected revenues.

However, as the descriptive statistics in Sect. 3 suggest, CD and HDI alone are not sufficient to explain movie success. This is confirmed here by the associated t statistics. This issue is further explored by disentangling the R^2 of the regression by groups of variables following a Shapley value analysis; see Huettner and Sunder (2012). Results indicate that country characteristics (CD, HDI, and GD) globally account for less than 1% of the R^2, while the inverse Mills ratio, which represents the selection hazard, explains 23% of the variance in revenue and movie-related variables explain 77%. These results clearly suggest the importance of controlling by film features in the analysis.

Previous results do not exclude the possibility of different types of adjustments at the intensive and extensive margins for different clusters of countries. However, a main problem with HDI and CD (the most significant indicators of distance) is that they are highly negatively correlated.[6] To deal with this issue, we break down the total sample of countries into four groups, according to their relative HDI

[6] In a regression between these two variables, the coefficient associated to CD is −0.02 with a p value equal to 0.048. However, the coefficient of a regression between CD and GD is insignificant, $p = 0.81$.

Table 4 Movies revenues—CD and HDI clusters

	Estimated parameters			
	Low CD		High CD	
	Low HDI	High HDI	Low HDI	High HDI
Inverse Mills ratio	−0.320*** (−4.42)	0.162 (1.81)	−0.242*** (−4.32)	0.995*** (4.49)
Budget	0.525*** (18.97)	0.660*** (32.29)	0.526*** (30.19)	0.890*** (13.83)
Nomination	0.321*** (4.93)	0.867*** (20.38)	0.247*** (6.94)	0.760*** (7.31)
Sequel	0.395*** (5.67)	0.478*** (11.65)	0.373*** (10.25)	0.273*** (3.51)
Sequel2	0.382*** (4.41)	0.598*** (11.44)	0.358*** (7.90)	0.765*** (7.52)
Action	0.171** (3.14)	0.299*** (9.67)	0.315*** (10.70)	0.534*** (6.79)
Drama	−0.330*** (−5.05)	−0.143*** (−4.07)	−0.283*** (−8.05)	−0.0586 (−0.78)
Thriller	−0.0561 (−0.89)	0.280*** (7.93)	0.310*** (8.81)	0.599*** (6.34)
G	0.161 (1.18)	0.329*** (4.58)	−0.0378 (−0.56)	−0.0194 (−0.14)
PG	0.0739 (1.30)	0.0530 (1.63)	0.120*** (4.11)	−0.00,807 (−0.12)
R	−0.0603 (−1.24)	0.0322 (1.20)	−0.145*** (−5.74)	−0.0630 (−1.18)
Original screenplay	0.0696 (1.65)	0.0431 (1.82)	0.0730*** (3.31)	0.0617 (1.31)
N	2974	14,448	10,992	3296
Adjusted R^2	0.346	0.294	0.354	0.290

Countries in each cluster see Table 2

Omitted: comedy, PG-13; *t* statistics in parentheses

*$p < 0.05$; **$p < 0.01$; ***$p < 0.001$

and CD values:(1) low CD and low HDI, (2) low CD and high HDI, (3) high CD and low HDI, and (4) high CD and high HDI. A complete list of the countries belonging to each of these groups, together with a separate estimation of the baseline model for each case, can be found in Table 4. Note that this equation includes country-year fixed effect to control for unobserved country-level fluctuations such as tariffs or any other type of macroeconomic policy. Although the econometric specification follows an empirical approach, the theoretical insight is that consumer preferences of American films in a given country are affected by its cultural distance with respect to the USA as well as its degree of human and economic development.

The results from the probit model estimation are not reported in this case, for the sake of brevity, but the results show that the qualitative impact of all the variables is like their estimated effect in the primary equation.

Interestingly, the negative sign in the inverse Mills ratio observed at the aggregate level can only be found for countries either with low CD and low HDI or high CD and low HDI. A plausible explanation is that the saturation in the number of films is especially relevant in relatively poor countries with a low

development index. For the groups of countries with high HDI, there is either a positive inverse Mills ratio or no sample selection.

Results concerning the importance of film features by groups of countries are also very helpful in enabling decision makers determine which countries to target for each type of film. Just paying attention to the size of the estimated coefficients, variables related to film quality such as budget and nomination exert the highest impact on revenues. Turning the attention to the statistical significance of the parameter estimates, film quality indicators (budget and nomination) always have a positive and significant impact on revenues for the four groups; however, the magnitude of the impact of these two variables is bigger for countries with high HDI. Sequels are well received in countries that are close to the American culture (low CD), while a second or subsequent sequel is especially successful in high HDI countries. The reason for this could be that only people who are familiar with American culture are interested in the first sequel of a film, while a second or subsequent sequel of a given film means that it is sufficiently famous everywhere, and its success is mainly determined by the degree of development of the country. Conversely, original screenplay movies have a more positive impact in countries with a high CD.

In relation to genres, low HDI countries tend to show a strong preference for thrillers and a dislike for drama while high CD tends to countries like action. This result seems plausible, given that people with a very different culture likely find action movies easier to understand than other genres. Finally, the mildest movies (G rating) are preferred by countries with low CD and high HDI.

5 Concluding remarks

We have estimated how American film revenues, in 45 countries, are affected by both film features and variables, such as distance indicators and the HDI, at the country level. This analysis has three important results: (1) both film features and country-level variables significantly affect revenues; (2) patterns of film consumptions are different for countries, depending on their CD from the USA and their HDI; and (3) different film features are appreciated in different ways in different countries. It is also very remarkable that, in general, distance indicators influence both adjustment in revenue and in the number of films that arrive in different countries. Hanson and Xiang (2011) find that revenue adjustments are more important when using aggregate data at the country level; this fact warrants further investigation in future research. Overall, our results suggest that trade analysis at the country level allows us to control for the different sources of heterogeneity that are potentially present in the model.

Three potential extensions for future works are: (1) to extend this type of research to other cultural products to see whether results are consistent in other context; (2) to develop theoretical foundations to explain the different trade adjustments in different clusters of countries; and (3) to explore further the reasons why different genres could be subject to different cultural discounting.

Compliance with ethical standards

Conflict of interest The author Dimitri Paolini has received research grant from PRIN (Research projects of national interest) by Minister of Education and from Fondazione Banco di Sardegna. The authors G. Meloni and JD. Tena declare that they have no conflict of interest.

References

Altman, N. S. (1992). An introduction to kernel and nearest-neighbor nonparametric regression. *The American Statistician, 46*(3), 175–185.

Bernard, A. B., & Jensen, J. B. (1999). Exceptional exporter performance: Cause, effect, or both? *Journal of International Economics, 47*(1), 1–25.

Bernard, A. B., Jensen, J. B., Redding, S. J., & Schott, P. K. (2007). Firms in international trade. *Journal of Economic Perspective, 21*(3), 105–130.

Fu, W. W., & Lee, T. (2008). Economic and cultural influences on the theatrical consumption of foreign films in Singapore. *Journal of Media Economics, 21,* 1–27.

Hanson, G., & Xiang, C. (2011). Trade barriers and trade ows with product heterogeneity: An application to US motion picture exports. *Journal of International Economics, 83,* 14–26.

Heckman, J. J. (1979). Sample selection bias as a specification error. *Econometrica, 47,* 153–161.

Hellmanzik, Ch., & Schmitz, M. (2015). Virtual proximity and audiovisual services trade. *European Economic Review, 77,* 82–101.

Helpman, E., Melitz, M. J., & Rubinstein, Y. (2008). Estimating trade flows: Trading partners and trading volumes. *Quarterly Journal of Economics, 123*(2), 441–487.

Hofstede, G. (1980). *Culture's consequences: International differences in work-related values*. Beverly Hills, CA: Sage.

Hofstede, G. (2001). *Culture's consequences, comparing values, behaviors, institutions, and organizations across nations*. Thousand Oaks, CA: Sage.

Holloway, I. R. (2014). Foreign entry, quality, and cultural distance: Product-level evidence from US movie exports. *Review of World Economics, 150*(2), 371–392.

Huettner, F., & Sunder, M. (2012). Axiomatic arguments for decomposing goodness of fit according to Shapley and Owen values. *Electronic Journal of Statistics, 6,* 1239–1250.

Kogut, B., & Singh, H. (1988). The effect of national culture on the choice of entry mode. *Journal of International Business Studies, 19*(3), 411.

Lee, F. L. F. (2009). Cultural discount of cinematic achievement: The Academy Awards and U.S. movies' East Asian box office. *Journal of Cultural Economics, 33,* 239–263.

Marvasti, A., & Canterbery, E. R. (2005). Cultural and other barriers to motion pictures trade. *Economic Inquiry, 43*(1), 39–54.

Mayer, T., & Zignano, S. (2011). Notes on CEPII's distance measures: the GeoDist database. CEPII working paper, 2011-25.

Melitz, M. J. (2003). The impact of trade on intra-industry reallocations and aggregate industry productivity. *Econometrica, 71*(6), 1695–1725.

Palsson, C., Price, J., & Shores, J. (2013). Ratings and revenues: Evidence from movie ratings. *Contemporary Economic Policy, 31*(1), 13–21.

Wagner, C. H. (1982). The Simpson's paradox in real live. *The American Statistician, 36*(1), 46–48.

Walls, W. D., & McKenzie, J. (2012). The changing role of hollywood in the global movie market. *Journal of Media Economics, 25*(4), 198–219.

Wooldridge, J. M. (2010). *Econometric analysis of cross section and panel data*. Cambridge: MIT Press.