ISSN 2385-2755
DiSSE Working papers
[online]

SAPIENZA
UNIVERSITÀ DI ROMA

WORKING PAPERS SERIES
DIPARTIMENTO DI SCIENZE SOCIALI ED ECONOMICHE

Diversity on the screen

Manthos D. Delis, Anastasia Litina, Skerdilajda Zanaj

N. 08/2021

SAPIENZA - UNIVERSITY OF ROME
P.le Aldo Moro n.5 – 00185 Roma T(+39) 0649910563
CF80209930587 – P.IVA 02133771002

Diversity on the Screen

Manthos D. Delis
Montpellier Business School

Anastasia Litina
University of Macedonia

Skerdilajda Zanaj
University of Luxembourg

For comments and suggestions on our paper, we are grateful to Conchita D'Ambrosio, Eric Roca Fernandez, Iftekhar Hasan, Giorgia Menta, Naci Mocan, Steven Ongena, Eleonora Patacchini, Vassilis Sarantides, and Chris Tsoumas. We would also like to thank participants at the CREA workshop on gender and economics and the webinar on gender and family economics.

Coordinates of the authors: Delis is with Montpellier Business School, 2300 Avenue des Moulins, 34185, Montpellier, France. Email: m.delis@montpellier-bs.com. Litina is with the University of Macedonia, Egnatia 156, 54636, Thessaloniki, Greece. Email: alitina.uoi@gmail.com. Zanaj is with the University of Luxembourg, Kirchberg Campus, 6 Rue Richard Coudenhove-Kalergi, L-1359 Luxembourg. Email: skerdilajda.zanaj@uni.lu.

Diversity on the Screen

Abstract

Using hand-collected data on movies from 1998 to 2008, we examine how deep-rooted population diversity in the origin countries of the cast (actors) and the production team (director, writer, and producer) affects movie performance (spectator ratings and box office revenue). We contend that distinguishing between the cast (what is visible by spectators-consumers) and the production team allows an analysis of how "visible diversity" affects performance. Once meticulously controlling for selection-endogeneity concerns, we find that the visible component has a hump-shaped effect on our movie performance measures and mostly drives our findings. We also show that the optimal level of cast diversity (the one that maximizes movie performance) is significantly higher than the sample's average value.

Keywords: Population diversity; Visible diversity; Movie industry; Movie ratings; Box office; Origin country

One of the interesting things that the most successful movies share is that they're broadly appealing to diverse audiences… People want to see a world that looks like theirs."
"Christy Haubegger, Creative Artist Agency, 2019"

1. Introduction

The *New York Times* (November 2019) list of 50 must-see movies includes *Coco* (2017), *Jackie Brown* (1997), *Moonlight* (2016), and *Philadelphia* (1993).[1] These movies are all box office bombs and present a high diversity of their casts and production teams, including the leading actors, directors, writers, and producers. *Coco*, produced in the USA with co-director Adrian Molina, who is a second-generation Mexican immigrant, cheerfully interweaves the traditions and culture of Mexico. *Jackie Brown*, written and directed by Quentin Tarantino, whose father is of Italian descent and whose mother has Irish and Cherokee ancestry, features Pam Grier, an African-American actress. *Moonlight*, written and directed by Barry Jenkins, is the story of a young gay African-American man exploring and challenging modern perceptions of masculinity and family. *Philadelphia*, directed by the Jewish-American director and writer Ron Nyswaner, features Denzel Washington and Tom Hanks as lead actors.

In this study, we examine the relationship between population diversity and performance using information from the unique setting of the movie industry. By population diversity, we refer to the variations in deep-rooted genetic and morphological features that define different populations. In turn, we consider that this deep-rooted population diversity might affect team performance. This variation relates to the key members of a movie's team, including actors, directors, writers, and producers. By movie performance, we refer to a movie's ratings via spectators and the box office.

Our research is relevant along two key dimensions. The first dimension is the importance of cinema in modern societies. The movie industry has a financial turnover of $136 billion USD (MPA Statistics, 2018), constituting approximately 2% of global GDP without considering the positive financial spillovers

[1] https://www.nytimes.com/article/best-movies-netflix.html Last accessed on November 2019.

to other industries.[2] Along the same lines, cinema is the most influential art form to reach large and heterogeneous audiences. Historically, many leaders have used the power of film to help achieve their political agendas. Moreover, movies are the main vehicle of edutainment — educating by entertaining. In this regard, multicultural societies reinforce social cohesion, tolerance, and acceptance via movies that foster population diversity.

Second and most important, a unique feature of the movie industry is that part of its diversity is visible to consumers, who directly observe the lead actors. We call this "visible diversity" or "cast diversity" and we shed light on its effect on movie performance, distinguishing it from the effect of invisible diversity (diversity of the production team, including the directors, writers, and producers). Population diversity is increasing sharply in a globalized world and has become a fact of life in Europe, the United States, and beyond (Alesina and La Ferrara, 2005; Lazear, 1999a; Ozgen et al., 2011b). In the United States, for example, minorities are expected to become the majority in 2042 (Bernstein and Edwards, 2008). Consequently, it does not come as a surprise that movie spectators have become much more diverse than decades ago, most probably leading to a rise in the demand for films with higher visible diversity. Very recently, McKinsey (2021) illustrates that improving Black representation could potentially raise movies' annual revenue by approximately 7%.[3]

A movie's invisible diversity (henceforth production team diversity) is also relevant. It represents a pool of resources for diverse screenplays, new styles, fresh insights, and original stories, partly driven by the directors', writers', and producers' own ethnic backgrounds. This team diversity feeds movies in multiple ways (i.e., via modes of cooperation, the realm of ideas, the diversity of the cast the production team chooses), which might directly or indirectly affect movie performance. In a nutshell, the movie industry offers a unique field in which to examine the separate roles of visible and production team diversity (as well as the total movie diversity obtained from both components) on movie performance.

[2] MPA: 2019 Global Box Office and Home Entertainment". Motion Picture Association. 2020.
[3] https://www.mckinsey.com/featured-insights/diversity-and-inclusion/black-representation-in-film-and-tv-the-challenges-and-impact-of-increasing-diversity, last accessed on 20/05/2021.

For our empirical analysis, we use two sources of data. First, we build a novel and mostly hand-collected dataset of more than 70,000 movies (the final sample varies across specifications due to the limited availability of key variables) produced worldwide from 1998 to 2008. Our outcome variables are movie ratings (mostly the number of scores between 8/10 and 10/10) in the International Movie Database (IMDb) and box office revenue from OPUS.

In turn, we measure total movie diversity using information about the nationalities of movie stars, directors, writers, and producers, as well as information on historical population diversity scores in these countries from Ashraf and Galor (2013). Specifically, we input the country-specific population diversity score for each movie team member and calculate the standard deviation of the scores as our movie-specific diversity score. This measure is fully consistent with the definition of population diversity in social sciences (e.g., Delis et al., 2017). Moreover, this measure of population diversity allows us to investigate the extent to which the deep-rooted historical (and thus mostly exogenous) diversity among movies' members affects movies' performance. Aside from a measure of total movie diversity, calculated from the scores for all movie members, we calculate separate scores based on visible diversity (the main three stars of the movie) and production team diversity (writer, producer, director). This distinguishes between the effects of visible and invisible diversity.

Our results show that total movie diversity has a statistically significant and inverted U-shaped effect on movie ratings. In fact, our estimates show the optimal total diversity score is 0.018, and the average movie in our sample has a diversity score of 0.0038. We thus find that a 0.0142 increase in total diversity increases ratings between 8 and 10 by approximately 207 votes, which corresponds to a 16% increase for the average movie in our sample. The baseline findings are very similar when separately considering cast diversity and team diversity. These findings survive a large battery of robustness tests, including the use of additional controls and fixed effects.

In obtaining the causal effect of population diversity on movie performance, we also consider the possibility that selection bias and omitted-variable bias drive our inferences (simultaneity/reverse causality is not an important concern due to the historical features of our diversity measure). The selection bias can

arise due to either voter self-selection on the IMDb platform or directors' selection of actors to boost the movie's diversity. Concerning the former selection problem, using the movie's budget as the response variable and finding very similar results alleviates the concerns. However, the problems related to the selection of actors or omitted variables require different approaches. A common solution is to use an instrumental variable within the Heckman model or the standard two-stage least squares (as remedies for selection and omitted variables, respectively). Our exogenous instrument is ultraviolet (UV) radiation. According to Delis et al. (2017), UV radiation can mutate genes, thereby affecting alleles, which are the underlying element in genetic diversity. Although this would have a direct impact on genetic diversity, there is no theoretical channel directly linking UV radiation in movie team members' countries to movie ratings.

Our findings based on either the Heckman or the 2SLS regressions show that the dominant humped-shaped effect comes from cast (visible) diversity; production team (producer, writer, director) diversity alone does not strongly affect movie outcomes. Economically, the results for cast diversity from the Heckman model are equivalent to our baseline and are somewhat more potent in the 2SLS regressions. This finding contributes to the extant literature, which reflects on the cooperation-versus-creativity argument; that is, the ability to collaborate and coordinate decreases in diverse teams, but these teams are also more creative. Our key finding on the dominant effect of visible diversity implies that what matters most is what individuals readily observe and identify with. For the movie industry in particular, this mechanism hints that spectators like diversity on the screen, most probably appreciating that movies represent a world that increasingly looks like theirs.

The structure of the paper is as follows. Section 2 presents the theoretical setting and discusses our contribution to prior literature. Section 3 motivates the paper by presenting facts and figures that highlight the importance of the movie industry, discussing the production function of movies, and providing anecdotes about the role of diversity. Section 4 discusses the data and empirical method of the paper, and section 5 presents the empirical results. Section 6 concludes the paper.

2. Theoretical setting and contribution

2.1. Population diversity and performance

Population diversity might be a double-edged sword for team performance. A diverse team enjoys different backgrounds, different experiences, and diverse ideas and knowledge; in turn, these attributes yield innovation and improved productivity. However, diversity might also bring disagreement, misinterpretations, segmentation into smaller exclusive groups, disruption, mistrust, conflict, and ultimately adverse effects on productivity and performance. The central consensus in the extant literature is that gains from diversity occur when team members have disjointed information sets that are relevant to one another and when this information can be communicated to (or understood by) other team members at low cost (Lazear, 1999a). If, instead, communication among team members is very costly, the adverse effects of diversity prevail.

Diversity and performance have been under scrutiny in different fields of social science research, such as economics, management science, political science, and psychology, to name a few. Our paper mainly relates to the economics and management science literatures. In economics, the literature examines how diversity affects team performance and countries' growth. Alesina, Spolaore, and Wacziarg (2000) employ a Dixit-Stiglitz production structure, where higher "inputs" variety reflects higher variety in individual skills that increase output. Lazear (1999a; 1999b) discusses how different skills in a production unit might increase overall productivity. He identifies a tradeoff between the productive benefits of diversity and the possible costs that can arise due to communication and language differences.

Ashraf and Galor (2013) study the macroeconomic effects of countries' historical population diversity on contemporary economic outcomes. Extending data on within-country genetic diversity from an initial set of indigenous populations to a large range of contemporary countries, they find an inverted U-shaped relationship between diversity and countries' comparative economic development. The nonlinear relationship reflects the interplay between the positive and the negative effects of diversity: very high

variation makes the transmission of human capital more complicated and is associated with conflict (Ashraf et al., 2020), but moderate diversity brings innovation and improved economic outcomes.

The management literature also perceives the diversity-performance nexus as nonlinear (Elron, 1997; Earley and Mosakowski, 2000; Jackson et al., 2003). O' Reilly, Williams, and Barsade (1997) analyze 32 project teams and find that more diversity leads to more conflict and less communication; however, when solving communication problems, diversity leads to higher productivity. Watson et al. (1998) find that ethnic diversity is detrimental to performance in the long run. In Pratt (2000), when jobs are complementary, homogeneity has positive effects; the opposite effect occurs when jobs are substitutes. Delis et al. (2017) use a panel of UK and U.S. public firms to link board diversity with firm performance. They document positive effects on performance, as measured by risk-adjusted returns.

2.2. Contribution

Using the movie industry as our setting is important for two interrelated reasons. First, we contribute a novel channel to the diversity literature: visible diversity (cast diversity that spectators see on the screen). Visible diversity affects demand (i.e., whether spectators prefer cast diversity). This means that besides the cooperation-versus-conflict mechanism, which is the focus of previous literature, we examine a more easily identifiable aspect of diversity in the movie industry. The role of visible diversity is extremely difficult to identify in other industries because the output is not directly related to what consumers observe. Along these lines, diversity (both total and visible) in the movie industry is straightforward to measure, and the same is true for the outcome variables (i.e., the box office and the ratings).

Second, our paper is the first on the movie industry (see McKenzie, 2012 for the most recent survey on the movie-industry literature). Prior papers study movie outcomes as a function of film-specific explanatory variables (e.g., budgets, awards, reviews, star power, genre, and ratings). Moreover, the literature on movie performance focuses on the United States or considers a relatively small number of films. For example, Elberse and Eliashberg (2003) study the weekly box office determinants for 164 movies released in 1999 in the United States, France, Germany, Spain, and the UK. They find that the number of

screens is the primary determinant of revenues and that advertising is the main predictor of opening week revenues.

A vivid debate in the movie literature is on the role of the stars. Employing a star is a signaling strategy for the quality of the film, increasing the chances of success (Ainslie et al., 2005), but it also increases costs. Other studies consider the effect of being the first movie in the opening weekend (Cabral and Natividad, 2016), the efficiency of using social media data to measure audience "sentiments" (Lehrer and Xie, 2017), when to stop screening a film at a specific theater (Chisholm and Norman, 2006), or the effects of changes in U.S. copyright law (Hui and Png, 2002). In our paper, these movie characteristics only serve as control variables.

Besides our key motivation to study the population diversity-performance nexus and distinguish the visible component, our analysis also speaks to the edutainment aspect of movies. Banerjee, Barnhardt, and Duflo (2015); Ravallion et al. (2015); Coville et al. (2014); Berg and Zia (2017); and Banerjee et al. (2019) all explore edutainment, especially how TV movies and series affect choices and behavior. For example, Banerjee et al. (2019) show how exposure to educational TV shows improves knowledge and attitudes toward HIV and risky sexual behavior. Treated subjects are twice as likely to get tested for HIV eight months after the intervention. One reason for this success is that the education message is not perceived as a top-down advice. Diversity on the screen is vital within the same line of thinking. Movies are a great place to start a diversity dialogue, and ethnically diverse movies can be instruments to normalize diversity.

Movies are also a key edutainment channel for diversity among young generations because they engage a wide audience of teenagers and young adults. Learning from movies is very direct and quick, making movies well adapted to millennials' speedy lifestyles. Furthermore, movies are vehicles of collaboration between teachers and students on diversity issues; they build connections through generations.

As the American Film Institute puts it, "The revolutionary power of visual storytelling inspires empathy and drives culture forward" (AFI 2018-2019, Impact Report).[4]

3. The movie industry

3.1. Figures and facts

The diversity of movie teams comes from the individuals involved in the production process, including the director, writer, producer, and actors. Movie outcomes are viewership (the box office) and ratings. The industry itself is very large: in 2018 alone, the global box office was worth $41.7 billion, according to UN statistics. When including box office and home entertainment (i.e., series and telemovies) revenue, the global film industry was worth $136 billion in 2018, which is almost 2% of global GDP. In 2017, China (1,620 million tickets), India (2,020 million tickets), and the United States (1,240 million tickets) were the three largest film markets in the world. In 2019, 29.6 million Americans watched the Academy Awards ceremony, whereas in 2014 the viewership was 40 million.

Besides its economic magnitude, the movie industry it is an important means of reaching wide audiences. Cinema is "the seventh art" and maybe the most influential one given its reach among large and heterogeneous audiences, as well as its large share of global GDP. Historically, many leaders used the power of film to achieve their goals (e.g., Hitler and Stalin successfully used movies as propaganda). Moreover, movies advertise products, create a demand for experiences, and affect lifestyles. In a nutshell, the movie industry is both an important and understudied industry, and it is an excellent setting to examine how diversity affects economic outcomes.

3.2. The production function of movies

Sedgwick and Pokorny (1998 and 2004) describe in detail film as a commodity and the characteristics of its production function. They also analyze film from a historical perspective, as well as the evolution of the

[4] https://www.afi.com/Assets/Impact-Report/ last accessed on November 2019.

budget and distribution roles, which generate box office hits or flops. The key feature of a film's production is its multistage technology, involving the identification of an initial story or idea, the screenwriting, the casting, the shooting, the sound recording and preproduction, editing, and finally promoting, advertisement, distribution, and screening.

A movie's team is thus a key source of its eventual success. However, it is unclear how diversity among the cast and overall team affect movie performance. Movie stars are the main attraction for all audiences, but the writers, producers, directors, and their collaboration are also instrumental for success. The writers closely collaborate with the directors and develop the initial idea. Directors, responsible for the casting and shooting, make screenplays come alive. Producers finance and coordinate the effort (Hollywood Report, 2019).[5]

Given the above, a movie's crucial "factors of production" as related to the potential role of diversity are (i) the leading actors who belong to diverse groups and who might attract diverse audiences; (ii) the director(s), who select and lead the team and can either exploit diversity positively or create disruption while making technical decisions concerning cinematographic techniques (cameras, screen types, language, locations, sound mix, etc.); (iii) the writer(s), who are the primary source of diversity in bringing new stories to the screen and defining genres; (iv) the producer(s), who sustain movie budgets and distribution; and (v) the distribution companies, which are in charge of movie promotion. To identify the effects of diversity given this production technology and to account for omitted-variable bias, we produce an extensive and novel database that includes information on all these factors of production.

4. Data and variables

4.1. Data and outcome variables

We obtain our data from two main sources: hand-collected data from the publicly available IMDb database and purchased data from OPUS. The largest available sample for which we have at least the ethnicity of

[5] https://socialsciences.ucla.edu/hollywood-diversity-report-2019/ last accessed on November 2019.

one member of the cast/team/both and the number of votes amounts to 77,316 movies and covers 1998 to 2008. However, for many movies we lack information about budgets, the ethnicity (country of origin) of movie members (cast and team), and important controls in the country of origin for the movie members. Thus, for more restrictive specifications, the number of observations falls significantly.[6] Table 1 defines the variables in our empirical analysis, and table 2 reports summary statistics.

[Insert Tables 1 & 2 about here]

We measure movie performance using viewer ratings on IMDb (*Ratings*) or total box office revenue (*Box office*). These are complementary measures, as they capture different aspects of spectator preferences. The IMDb ratings measure the overall appeal of the film to the public. We mainly use the number of IMDb users that give a rating of 8 out of 10 or higher (scaled by 1,000 for expositional brevity), as 8 is the threshold distinguishing very appealing movies from the rest.[7] Figure 1 reports the distribution of votes by each value (from 1 to 10). In robustness checks, we use additional data features such as viewer age range and gender. In turn, the total box office is the natural logarithm of the box office in current USD for the first country in which the movie was distributed. As this measure is available for fewer movies, our preferred measure is *Ratings*. *Ratings* is also more immune to the role of inputs (e.g., the production budget). For example, people may choose to see a blockbuster that has been heavily advertised, but the actual evaluation is usually quite objective.

[Insert Figure 1 about here]

[6] We view specifications with widely different observations as a robustness test showing that our subsamples do not drive our findings.

[7] The IMDb provides an overall rating, which we do not use because the exact formula is private information. On its website IMDb notes that, "We take all the individual ratings cast by IMDb registered users and use them to calculate a single rating. We don't use the arithmetic mean (i.e., the sum of all votes divided by the number of votes), although we do display the mean and average votes on the votes breakdown page; instead the rating displayed on a title's page is a weighted average." Concerning the weight, IMDb notes that: "IMDb publishes weighted vote averages rather than raw data averages. The simplest way to explain it is that although we accept and consider all votes received by users, not all votes have the same impact (or 'weight') on the final rating. Various filters are applied to the raw data in order to eliminate and reduce attempts at vote stuffing by people more interested in changing the current rating of a movie than giving their true opinion of it. In order to ensure that our rating mechanism remains effective, we do not disclose the exact method used to generate the rating. However, please rest assured that the same calculations are used to generate the rating for every title listed in the database: we don't adjust the rating for individual titles. There is no bias in how votes are weighted based on which title they have been cast for."

4.2. Diversity measures

We construct three diversity scores: *Total diversity*, *Cast diversity*, and *Team diversity*. For all measures, we rely on diversity data from Ashraf and Galor (2013a), who construct an index of country-specific genetic diversity scores based on data from the HGDP-CEPH Human Genome. The genetic diversity score of a given population or ethnic group measures the expected heterozygosity (i.e., the probability that two randomly selected individuals are genetically different from each other). As this is available for a limited number of countries, Ashraf and Galor (2013a) predict genetic diversity for all countries by estimating a model of genetic diversity on migratory distance from East Africa. The two variables are 90% correlated for the available observations and thus the predicted values from the regression serve as a genetic diversity score for all countries. Several additional papers from the same research team (Ashraf and Galor, 2013b; 2018; Arbatl et al., 2020) show that their index reflects population diversity (instead of genetic diversity).

Delis et al. (2017) use this index as a population diversity score to construct an exogenous index of firms' board diversity. Following this paradigm, we consider the following measure:[8]

$$\sigma = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(d_i - m)^2} \quad (1)$$

In equation 1, σ is the standard deviation of the predicted and adjusted genetic diversity score d attached to each actor, director, producer, and writer (aka the movie team), based on their country of origin. The variable m is the average genetic score of the movie team.

For *Total diversity*, we attach the country-specific scores to each member of the movie team, for *Cast diversity* we use only the three leading actors, and for *Production team diversity* we use everyone except the three leading actors (i.e., the director, the writers, and the producers). *Cast diversity* and *Production team diversity* are complementary, the first relating more to how spectators engage and identify with what they see on screen (visible diversity); the second measures the role of diversity in optimizing the

[8] We extensively discuss this exogeneity assumption in the next section.

production process. Thus, importantly, how *Cast diversity* affects movie performance relates more to visible diversity; *Production team diversity* relates more to the usual production-efficiency mechanism.

We note three interrelated issues regarding our diversity measures (following Delis et al., 2017). First, we attach scores to individuals' origin countries. This implies that we do not examine how individuals' actual genes (for which we have no information) affect movie performance. Instead, we stress the importance of diversity in the movie members' countries of origin to identify and measure the all-too-many characteristics of the countries in which individuals were born and raised that shape human behavior and cannot be captured—or are very incompletely captured—by the cultural and institutional variables of existing databases. These characteristics reflect the sociological, psychological, cultural, and physiological elements that the diversity of the underlying population of the countries correlate with or shape. It is precisely in this manner that we use our population diversity index.

Second, the standard deviation of diversity disregards whether genetic diversity in the individuals' origin country is high or low. The mean score of movie diversity would capture this, which would imply that the actual diversity score of movie members' origin countries (the relative homogeneity or heterogeneity of populations in that country) affects movie performance. Instead, the deviation of movie members' diversity scores (those carried with them) considers only how individuals differ with respect to various observable and unobservable personal traits in the country-specific diversity index. In other words, what possibly matters for movie performance is the inclusion of team members with different experiences, as well as the degree of difference in these experiences. This heterogeneity can be created by adding a team member from either a country with higher or lower diversity (or even multiple directors from different countries).

Third, given missing information concerning the origin country of some movie members, we use two approaches when constructing our indices. In our benchmark approach, we keep in the sample movies for which we observe at least one member for *Cast diversity* or *Production team diversity* (this implies a least two members for *Total diversity*, with both cast and team represented). This approach naturally yields more observations; however, it also naturally yields larger measurement error originating in the missing

information. To mitigate this concern, our second approach restricts our sample to movies with data for at least the first two main actors and the director and the producer. This leads to a lower number of observations but also a considerably lower measurement error. Our results are robust to these different approaches.

4.3. Control variables

We include two sets of controls: movie-related controls and country-of-origin (of the team and the cast) controls. Specifically, we control for each movie's budget, genre, distribution company, and main language. All are important aspects that can have a profound effect on a movie's success, whether measured by ratings or the box office. We also introduce dummies for the major production companies, as they are more likely to deliver a successful movie due to better networking.[9] When using the IMDb ratings as our outcome variable, we further control for the total number of voters.

We construct the macroeconomic controls using the same method for the movie's diversity scores (i.e., we use the values for the corresponding variables based on the directors' nationalities and take the standard deviation). As with *Team diversity* and *Visible diversity*, we construct three sets of controls for the stars and team, as well as separately for each of the two groups. We control for diversity in the origin countries' interpersonal trust, the index of democracy, constitutional characteristics, and economic conditions. For trust, we use aggregate information from the World Values Survey; for the quality of democracy and constitutional characteristics, we use data from the Polity IV database and the ICRG; for economic conditions, we use GDP per capita (in real U.S. dollars) from the World Development Indicators.[10]

[9] We further control for total number of production companies, total runtime, number of actors, and number of awards (if any). Given the lower number of observations and the fact that results remain unchanged, we only use these variables in robustness checks.

[10] In the robustness section we also control for other types of cultural diversity. In particular we use the diversity measure of the principal component of Hofstede's five cultural indicators: power distance, individualism, masculinity, uncertainty avoidance, and long-term orientation. We restrict this test only to the robustness section due to a sharp decrease in the number of available observations.

5. Empirical identification and estimation results

5.1. Benchmark empirical model and estimation results

Our benchmark specification is the following:

$$Movie\ performance_{ict} = a_0 + a_1 Diversity_i + a_2 Diversity_i^2 + a_3 P_i + a_4 O_i + a_5 C_c + a_6 T_t + e_{ict}. \quad (2)$$

In equation (2), *Movie performance* is either *Box office* or *Rating* of movie *i* produced in country *c* and distributed in year *t. Diversity* is as defined in equation (1). P_i is the vector of controls that are associated with the production function of the movie and its qualitative characteristics. O_i is the vector of origin country controls, and C_c and T_t are country and year fixed effects controlling for time-invariant unobservables at the country level and common trends across countries, respectively. Finally, e_{ict} is the error term, and we cluster our standard errors at the country level to ensure that they are robust to arbitrary within-country autocorrelation.[11]

Table 3 reports our baseline results using *Ratings* as the outcome variable and *Total diversity* as the key explanatory variable. In column 1, we include only country and year fixed effects; in the rest of the specifications we sequentially add more fixed effects and control variables. Adding the controls (and especially the important macro controls) does not inflict significant changes in our inferences. Moreover, the reduction in the number of observations from sequentially adding controls provides the first evidence that our results are robust to sample size differences.

[Insert Table 3 about here]

Our results show a statistically significant and inverted U-shaped effect of *Total diversity* on *Ratings*. Based on the analysis of column (4), which includes the highest number of controls, the coefficient on *Total diversity* equals 11.34, and the coefficient on the quadratic term is -320. These estimates yield an optimal *Total diversity* of 0.018 [= 11.34 / (2 ×320)]. Importantly, for most of the movies in our dataset

[11] In the robustness section, we also conduct two-way clustering of the error term at the country and year levels.

Total diversity is significantly below this optimal value, having an average of 0.0038. Thus, a 0.0142 increase in *Total diversity* (the difference between the mean in our sample and the optimal value) increases *Ratings* by approximately 0.161 (= 0.0142 × 11.34) or 207 votes (= 0.161 × 1,287) giving ratings between 8 and 10 (*Ratings* is scaled by 1,000). This increase in votes represents a substantial 16% increase for the movie with the mean number of votes. Thus, most movies have substantial popularity gains to reap if they have higher overall diversity.[12]

Next, we examine the role of cast diversity, which relates more to the visible component. In table 4, we find that the inverted U-shaped effect of *Cast diversity* is also highly statistically and economically significant across all specifications. Based on the results in column (4), the optimal *Cast diversity* is 0.018 [= 12.35 / (2 × 345)], as is the case for *Total diversity* in table 3. Given that mean *Cast diversity* is particularly low (equals 0.002), the additional 0.016 (from 0.002 to 0.018) diversity yields an increase of 0.198 points or 156 votes, which represents a 20% increase for the movie with the average number of ratings between 8 and 10. This finding shows that visible diversity might be the most important component of diversity's effect on a movie's popularity.

[Insert Table 4 about here]

Table 5 reports the equivalent results from using *Production team diversity*, which also reflect an inverted U-shaped relation with the movie ratings. According to the results in column (4), optimal *Production team diversity* is 0.017. Working in the same way as for total and cast diversity, we find that moving from the mean value to the optimal value yields a 0.54 point increase in *Ratings*, which also represents a significant increase in the number of top ratings and shows the importance of diversity in productivity and performance.

[Insert Table 5 about here]

Figures 2a to 2c report our baseline results in graphical form and illustrate the inverted U-shaped relations in our regressions. We note that even though the average diversity values in our samples are very

[12] In appendix table A1, we report movies in our sample with diversity scores near its optimal value.

low, implying important benefits to reap from increasing diversity, moving to very high diversity (above optimal values) yields significant decreases in *Ratings* for all three measures of diversity.[13]

[Insert Figure 2 about here]

5.2. Accounting for identification problems

An important characteristic of diversity, as measured in our paper, is that the country-specific values attached to individuals reflect population diversity and associated traits determined centuries ago; thus, simultaneity / reverse causality are not of concern. Omitted-variable bias should also be limited because of the extensive set of controls and fixed effects, as well as the predetermined nature of *Diversity*. Thus, our key identification problems are selection bias and measurement error.

Focusing on selection bias, we have two main concerns. First, people who vote for a movie via the IMDb platform may self-select. For this to be a concern, we need the stretched assumption that only people who vote favorably / unfavorably systematically self-select into the platform. We argue that people who self-select are willing to vote and share their opinion, but we cannot infer any trend toward favorable or unfavorable voting. As shown above, Figure 1 illustrates the distribution of votes ranking 1-10 for the full set of movies for which we have available data. The distribution does not show evidence of such voter self-selection.

Nevertheless, we address any such bias by using *Box office* as an alternative outcome variable, which is immune to this type of selection bias. The results in table 6 are equivalent to those of the previous tables: in all three panels, reflecting *Total diversity*, *Cast diversity*, and *Production team diversity*, the results show diversity has an inverted U-shaped effect on *Box office*. These results are despite the fact that *Box office* is available for considerably fewer observations compared to *Ratings*. We also find that across all

[13] In appendix table A2, we use linear specification (i.e., without the quadratic term). The results in panel A clearly favor the nonlinear model.

specifications the optimal value of diversity is between 0.015 and 0.021, consistent with the findings in tables 3 to 5 and considerably higher than the mean values for total, cast, and production team diversity.[14]

[Insert Table 6 about here]

The second concern on selection bias relates to selecting the director based, for example, on past experience, which might in turn be correlated with diversity score. This includes the premise that directors from countries with high diversity scores also choose actors with high diversity scores to improve the movie's appeal. To safeguard our analysis against such bias, we design a two-stage Heckman model in which the first-stage selection equation is the following:

$$Director\ native_{ict} = a_0 + a_1 Diversity_i + a_2 Diversity_i^2 + a_3 P_i + a_4 O_i + a_5 C_c + a_6 T_t + e_{ict}. \quad (3)$$

Director native is a binary variable that equals 1 when the director is from the country where the movie is produced. On top of the controls in the second stage (i.e., equation 2), equation 3 includes two additional variables. The first is *Producer native*, which equals 1 if the producer is from the same country as the movie's country of origin (i.e., the country of the movie's main production company) and 0 otherwise. This variable indicates the production company's incentives to use a foreign (and diverse) producer. The second variable is *Multiple movies*, which takes values of 1 to 49, indicating the number of movies the director produced during our sample period (an indicator of experience and propensity to make more movies). Equation 2 naturally includes the Mills ratio from equation 3.

We report our findings in table 7. The inverse Mills ratio is in fact positive and statistically significant at the 1% level (showing an important role for selection), but this does not affect the statistical significance of *Total diversity* and its square in the *Ratings* equation. Importantly, the optimal total diversity score equals the respective value in column 4 of table 3. We obtain similar inferences concerning cast diversity, with our findings being equivalent to those in column 4 of table 4. A wrinkle arises when we consider production team diversity in column 3 of table 9. We find that the inverted U-shaped effect does

[14] In panel B of appendix table A2, we use a linear specification for the box office model (i.e., without the quadratic term). The results clearly favor the nonlinear model.

not hold, which might imply that selection of the movie's director/writer/producer significantly influences how production team diversity affects performance.

[Insert Table 7 here]

We delve deeper into this latter finding and the role of any remainder omitted-variable bias and/or measurement error in *Diversity* using a standard 2SLS model to treat more general endogeneity (e.g., attributed to omitted-variable bias). We use the standard deviation of a measure of ultraviolet exposure and its quadratic term as instruments, both constructed as described for *Diversity*.[15] According to Delis et al. (2017), UV radiation can cause mutation of genes, thereby affecting alleles, which are the underlying element in the measure of diversity. UV radiation also has an indirect effect on landscapes and lifestyles (Sturm and Duffy, 2012; Kozmin et al., 2005). This implies that ultraviolet exposure satisfies the relevance condition for a valid instrument. Although this would have a direct effect on our diversity score, there would be no direct effect on the box office or IMDb rating (thus satisfying the exclusion condition for a valid instrument). The key reason for the validity of the exclusion condition is that UV radiation in the movie team members' countries and the standard deviation of that measure are very unlikely to directly affect a movie's performance. Moreover, it further mitigates concerns about cast selection, as UV radiation is unlikely a consideration in movie casting.

The first-stage results, reported in appendix table A6, confirm that UV radiation satisfies the relevance condition: the coefficient of *UV radiation* is positive and highly statistically significant. The 2SLS results in table 8 report findings very similar to those in the Heckman regressions. Specifically, we identify the inverted U-shaped effects in specifications 1 and 2 (including total diversity and cast diversity, respectively), but we do not identify significant effects when using production team diversity in specification 3.

[15] Given that our model includes a quadratic term of diversity, we follow Ashraf and Galor (2013) and the two-step procedure suggested by Wooldridge (2010), section 9.5. First, we regress the diversity measure on our IV and controls. This delivers predicted values for the diversity measure, which we square. Then we use that squared term as a second IV in the second-stage regressions.

Overall, our results in this section suggest that cast diversity, reflecting visible diversity, affects movie performance more than the diversity of the movie's production team (including the director, writers, and producers) does.

[Insert Table 8 about here]

5.3. Robustness checks

We provide results from several additional robustness tests. Given the discrepancy in the results between the OLS and the IV methods, we rely mostly on the latter but also report the equivalent OLS results in the appendix. In table 9, we use the additional controls *Star power*, *Very high budget*, and *Number of actors* (definitions in table 1). We do not use these controls in our baseline regressions because their inclusion significantly reduces our sample. These variables enter the regressions with positive and statistically significant coefficients (at the 1% level) but hardly affect the inferences on our main explanatory variables. Using these controls in the box office specifications (results in appendix table A3) or OLS (results in appendix table A7) also yields similar results. In table 10, we additionally control for voter characteristics, such as the share of male voters and voter age group. Again, despite the lower number of observations, our results are qualitatively the same (equivalent OLS results in appendix table A8).

[Insert Tables 9 & 10 about here]

As discussed in section 4.2, an important source of measurement error might be our benchmark choice to keep in the sample movies for which we observe at least one member for *Cast diversity* or *Production team diversity* (implying at least two members for *Total diversity*). Despite the importance of the IV approach in mitigating such measurement error concerns, in table A4 we further restrict our sample to movies that provide data for at least the first two main actors and the director and the producer. This leads to

a smaller number of observations but considerably lower measurement error. Again, our results on the optimal level of diversity are similar to our baseline.[16]

6. Conclusions

Using hand-collected data and the unique setting of the movie industry, which has a powerful economic and societal impact, we uncover a novel effect of diversity on team performance. Our diversity measure reflects population diversity in the origin country of a movie's main actors, directors, writers, and producers. The novel effect emanates from "visible" diversity, which relates to the main cast's origin country and marks what spectators observe on the screen. The effect of visible diversity on movie performance operates alongside the diversity of the movie's production team (including directors, producers, and writers). To this end, we can distinguish between the visible (cast) and invisible (production team) components of diversity and have clear-cut measures of movie performance (spectator ratings and the box office).

Our findings suggest that visible and invisible diversity have a hump-shaped effect on movie ratings from spectators and the box office. However, when more diligently controlling for endogeneity concerns, the visible component overcomes the invisible one in statistical and economic significance. Equally important, we show that the optimal level of diversity (the one that maximizes ratings and the box office) in our empirical analysis is significantly higher than the mean level of diversity in our movies sample.

Our eyesight is one component of diversity because it affects what individuals readily and directly observe and identify with. We contend that visible diversity reflects a deeper social need for people to feel that the society they live in is inclusive and representative of the real world. Although we study this in the context of the movie industry, it is essential for future research to draw the analogy with other sectors where people may feel misrepresented on boards, in public administration, or in any other professional and societal environment.

[16] In table A5, we also report very similar findings when double clustering our standard errors by country and year.

REFERENCES

Alesina, A. and La Ferrara E. (2005). "Ethnic Diversity and Economic Performance." *Journal of Economic Literature*, 43 (3), 762-800.

Alesina, A., Spolaore E, and Wacziarg R. (2000). "Economic Integration and Political Disintegration." *American Economic Review*, 90 (5), 1276-1296.

Ainslie A., Drèze X. and Zufryden F. (2005). "Modeling Movie Life Cycles and Market Share," *Marketing Science*, 24(3), pages 508-51.

Arbatlı, C.E., Ashraf, Q.H., Galor, O. and Klemp, M. (2020). "Diversity and Conflict", *Econometrica*, 88, 727-797.

Ashraf, Q. and Galor, O. (2013a). "The 'Out of Africa' Hypothesis, Human Genetic Diversity, and Comparative Economic Development", *American Economic Review* ,103, 1-46.

Ashraf, Q. and Galor O. (2013b). "Genetic Diversity and the Origins of Cultural Fragmentation," *American Economic Review*, 103, 528–33.

Ashraf, N., Bandiera O., Davenport E., and Lee S.S. (2020). "Losing Prosociality in the Quest for Talent? Sorting, Selection, and Productivity in the Delivery of Public Services." *American Economic Review*, 110 (5), 1355-94.

Banerjee, A., Barnhardt, Sh. and Duflo, E. (2017). "Movies, Margins, and Marketing: Encouraging the Adoption of Iron-Fortified Salt". *Insights in the Economics of Aging*, edited by David A. Wise, Chicago: University of Chicago Press, 285-306.

Banerjee, A. La Ferrara, E. and Orozco-Olvera, V. H. (2019). "The Entertaining Way to Behavioral Change: Fighting HIV with MTV", *Policy Research Working Paper; No. 8998. World Bank, Washington, DC.*

Berg, G., and Zia, B. (2017). "Harnessing Emotional Connections to Improve Financial Decisions: Evaluating the Impact of Financial Education in Mainstream Media", *Journal of the European Economic Association*, 15-5, 1025-55.

Bernstein, R., and Edwards, T. (2008). "An older and more diverse nation by midcentury". Washington, DC: *U.S. Census Bureau News*

Cabral L. and Natividad G. (2016). "Box-Office Demand: The Importance of Being #1," *Journal of Industrial Economics*, Wiley Blackwell, vol. 64(2), 277-294.

Chisholm, D. and Norman, G. (2006). "When to exit a product: evidence from the US motion picture industry", *American Economic Review, Papers and Proceedings,* 96(2): 57–61.

Coville, A., Di Maro, V., Zottel, S., and Dunsch, FA (2014). "Nigeria's Nollywood Nudge. An entertaining approach to savings", in *Enhancing Financial capability and behavior in low- and middle-income countries*, ch. 10, The World Bank.

Delis, M., D., Gaganis, Ch., Hasan, I. and Pasiouras, F. (2017). "The Effect of Board Directors from Countries with Different Genetic Diversity Levels on Corporate Performance", *Management Science*, 63-1, 231-249.

Earley, P.C., Mosakowski, E., 2000. "Creating hybrid team cultures: An empirical test of transnational team functioning", *Academy of Management Journal* 43, 26-49.

Elberse, A. (2007). "The power of stars: do star actors drive the success of movies?", *Journal of Marketing* 71, 102–120.

Elberse, A. and Eliashberg, J. (2003). "Demand and supply dynamics for sequentially released products in international markets: the case of motion pictures", *Marketing Science* 22(3), 329–354.

Elron, E. (1997). "Top management teams within multinational corporations: Effects of cultural heterogeneity". *The Leadership Quarterly, 8*(4), 393–412

Hui, K. and Png, I. P. L. (2002). "On the supply of creative work: evidence from the movies", *American Economic Review,* 92(2). 217–220.

Jackson, S.E., Joshi, A. and Erhardt, N.L. (2003). "Recent Research on Team and Organizational Diversity: SWOT Analysis and Implications", *Journal of Management*, 29, 801-830.

Kozmin S, Slezak G, Reynaud-Angelin A, Elie C, de Rycke Y, Boiteux S, Sage E (2005). "UVA radiation is highly mutagenic in cells that are unable to repair 7,8-dihydro-8-oxoguanine in Saccharomyces cerevisiae". *Proc Natl Acad Sci U S A* 102(38):13538-43

Lazear, E. P., (1999a). "Globalisation and the Market for Team-Mates." *Economic Journal*, 109-454, C15-40.

Lazear, E.P., (1999b). "Culture and Language." *Journal of Political Economy*, 107(6), S95-126

Lehrer S. F. and Xie T. (2017). "Box Office Buzz: Does Social Media Data Steal the Show from Model Uncertainty When Forecasting for Hollywood?", *The Review of Economics and Statistics*, 99, 749–755.

McKenzie, J., (2012). "The Economics of Movies: A Literature Survey", *Journal of Economic Surveys*, 26 (1), 42-70.

O'Reilly, C.A., III, Williams, K.Y., and Barsade, W. (1998). "Group demography and innovation: Does diversity help?" In D. Gruenfeld (Ed.), *Research on Managing Groups and Teams*, Vol. 1, pp.183-207). St. Louis, MO: Elsevier.

Pratt, M. G. (2000). "The Good, the Bad, and the Ambivalent: Managing Identification among Amway Distributors", *Administrative Science Quarterly*, *45*(3), 456–493.

Ravallion, M., van de Walle D., Dutta P. and Murgai, R., (2015). "Empowering poor people through public information? Lessons from a movie in rural India", *Journal of Public Economics*, 132, 13-22.

Sedgwick, J. and Pokorny, M. (1998) "The risk environment of film making: Warner Bros in the inter war years", *Explorations in Economic History* 35: 196–220.

Sedgwick, J., & Pokorny, M. (2004). "The characteristics of film as a commodity". In J. Sedgwick, & M. Pokorny (Eds.), *An economic history of film* (26 ed., pp. 6-23), Routledge.

Sturm, R.A., Duffy, D.L. (2012). "Human pigmentation genes under environmental selection.", *Genome Biology,* **13,** 248

Watson, W. E., Johnson, L., and Merritt, D. (1998). "Team orientation, self-orientation, and diversity in task groups", *Group and Organization Management*, 23. 161-188.

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

Table 1. Variable definitions and sources

The table defines the variables used in the empirical analysis and provides information on their sources.

Dependent variables	
Box office	The natural logarithm of the box office in current USD in the first three countries the movie was distributed.
Ratings	Number of voters that gave an 8/10 rating or higher to the movie (scaled by 1,000)
Ratings 10	Numbers of voters that gave a 10/10 rating to the movie (scaled by 1,000).
Ratings 6 to 10	Numbers of voters that gave a 6/10 rating or higher to the movie (scaled by 1,000).
Explanatory variables	
Total diversity	$\boldsymbol{\sigma = \frac{1}{n}\sum_1^n (di - m)^2}$, where σ is the standard deviation of the genetic diversity score *d* from the mean value *m* of the diversity of the movie's team (actors, director, writer, producer), according to the diversity score of the country of nationality of each member.
Cast diversity	$\boldsymbol{\sigma = \frac{1}{3}\sum_1^3 (di - m)^2}$, where σ is the standard deviation of the genetic diversity score *d* from the mean value *m* of the diversity of the three leading actors of a movie, according to the diversity score of the country of nationality of each actor.
Production team diversity	$\boldsymbol{\sigma = \frac{1}{n}\sum_1^n (di - m)^2}$, where σ is the standard deviation of the genetic diversity score *d* from the mean value *m* of the diversity of the movie's team (director(s), writer(s), producer) excluding actors, according to the diversity score of the country of nationality of each member.
Genre	Genre of the movie by order of significance.
Language	Language used for the predominant dialogue. Up to five languages are listed according to their order of prominence, with 1 being the most prominent.
Budget	Cost of production of the movie. It does not include marketing or other relevant expenditures. The budget is expressed in ranges to mitigate reporting errors.
High Budget Dummy	A dummy variable that takes the value 1 if the budget is higher than 100.000.000 $ and 0 otherwise.
Production company dummies	Dummies for the major production companies. We include dummies for Sony, Walt, Universal, Fox, Paramount, Lionsgate, Weinstein, Goldwyn, Dreamworks, Babelsberg, Gaumont, Canal, Pathe and Nordisk.
Runtime	Duration of the movie in minutes.
No of actors	Number of actors in the movie.
Star power	Star power is a dummy variable that takes the value 1 if the salary of the actor/actress is higher than 100.000$ and 0 otherwise.
% of male voters	The percentage of male voters voting in IMDB.
% of female voters	The percentage of female voters voting in IMDB.
% of voters aged 18-29	The percentage of voters aged 18-29 voting in IMDB.
% of voters aged 30-44	The percentage of voters aged 30-44 voting in IMDB.

% of voters aged 45+	The percentage of voters aged 45+ voting in IMDB.
Director native	A binary variable, taking the value 1 when the director is from the same country where the movie is produced.
Produce native	A binary variable taking the value 1 if the producer is from the same country as the movie's country of origin (i.e., the country of the movie's main production company) and 0 otherwise.
Multiple movies	An ordered variable taking values 1-49 and indicating the number of movies that the director has produced during our sample period (an indicator of directors' experience).
Year	Year when the movie was initially released (1998-2008).
Country	The country that paid for the movie (178 countries in total).
Political diversity	The standard deviation of the Polity IV index of democracy from the country of nationality of movie's team members by film and year, constructed in the same way as *Cast diversity* and *Team diversity*.
Institutional diversity	The standard deviation of the ICRG Law and Order Index from the country of nationality of movie's team members by film and year, constructed in the same way as *Cast diversity* and *Team diversity*.
Diversity in trust	The standard deviation of a trust index from the country of nationality of the movie's team by firm and year, constructed in the same way as *Cast diversity* and *Team diversity*. Average scores are calculated based on the samples of respondents from the World Values Survey.
Diversity in economic development	The standard deviation of GDP per capita (in constant USD) from the country of nationality of movie's team members by firm and year, constructed in the same way as *Cast diversity* and *Team diversity*.
Instrumental variable	
Ultraviolet exposure	The intensity of ultraviolet exposure.

,

Table 2. Summary statistics

The table reports summary statistics (number of observations, mean, standard deviation, minimum, and maximum) for the main variables used in the empirical analysis.

	N	Mean	S.d.	Min.	Max.
Total diversity					
Ratings	7,354	1.287	3.737	0	68.06
Total diversity	7,354	0.004	0.006	0	0.049
Diversity in trust	7,354	0.033	0.047	0	0.344
Political diversity	7,354	0.457	1.446	0	11.31
Institutional diversity	7,354	0.263	0.375	0	2.320
Diversity in econ. development	7,354	0.304	0.504	0	2.985
Genre	7,354	6.595	4.628	1	25
Language	7,354	60.10	31.25	2	186
Production company	7,354	0.072	0.260	0	1
Budget order	7,354	1.201	0.606	1	4
Director native	7,354	0.365	0.481	0	1
Produce native	7,354	0.671	0.469	0	1
Multiple movies	7,308	2.492	3.534	0	47
Cast diversity					
Ratings	13,020	0.791	2.917	0	68.06
Cast diversity	13,020	0.002	0.005	0	0.057
Diversity in trust	13,020	0.020	0.042	0	0.338
Political diversity	13,020	0.288	1.264	0	12.02
Institutional diversity	13,020	0.171	0.356	0	2.681
Diversity in econ. development	13,020	0.203	0.484	0	3.026
Genre	13,020	7.112	4.975	1	25
Language	13,020	62.45	34.93	2	188
Production company	13,020	0.047	0.212	0	1
Budget order	13,020	1.120	0.477	1	4
Director native	13,020	0.648	0.473	0	1
Produce native	13,020	0.822	0.381	0	1
Multiple movies	12,755	0.725	1.444	0	14
Production team diversity					
Ratings	4,553	1.494	4.132	0	68.06
Production team diversity	4,553	0.001	0.003	0	0.047
Diversity in trust	4,553	0.011	0.035	0	0.414
Political diversity	4,553	0.119	0.873	0	12.02
Institutional diversity	4,553	0.081	0.255	0	2.807
Diversity in econ. development	4,553	0.090	0.328	0	2.985
Genre	4,553	6.376	4.315	1	25
Language	4,553	59.57	30.49	2	186
Production company	4,553	0.076	0.266	0	1
Budget order	4,553	1.233	0.654	1	4
Director native	4,553	0.231	0.421	0	1
Produce native	4,553	0.593	0.491	0	1
Multiple movies	4,547	0.872	1.400	0	9

Table 3. Movie ratings and total movie diversity

The table reports coefficient estimates and standard errors (in parentheses) from the estimation of equation (2). The dependent variable is *Ratings*. All variables are defined in Table 1. Estimation method is OLS. The lower part of the table reports the number of observations, the Adjusted R-squared, and the type of fixed effects included in each specification. Optimal diversity is the turning point (maximum) at which the effect of diversity turns negative (obtained from the derivative of each specification with respect to the diversity measure). The ***, **, and * marks denote statistical significance at the 1%, 5%, and 10% levels, respectively.

	(1)	(2)	(3)	(4)
Total diversity	7.636***	9.515***	7.328***	11.34***
	(2.484)	(2.337)	(1.963)	(2.083)
Total diversity squared	-245.7**	-248.5***	-163.2**	-320.0***
	(94.22)	(68.10)	(71.89)	(106.2)
Diversity in econ. development		-0.0552***	-0.0528***	-0.0699***
		(0.0113)	(0.00785)	(0.0200)
Diversity in trust				-0.406***
				(0.151)
Political diversity				-0.00822***
				(0.00292)
Institutional diversity				0.0746**
				(0.0367)
Optimal diversity	0.015	0.019	0.022	0.018
Observations	38,470	8,409	8,223	7,354
Adj. R-squared	0.073	0.436	0.492	0.488
Country FE	YES	YES	YES	YES
Year FE	YES	YES	YES	YES
Budget order FE	NO	YES	YES	YES
Genre FE	NO	NO	YES	YES
Language FE	NO	NO	YES	YES
Production company FE	NO	NO	YES	YES

Table 4. Movie ratings and cast diversity

The table reports coefficient estimates and standard errors (in parentheses) from the estimation of equation (2). The dependent variable is *Ratings*. All variables are defined in Table 1. Estimation method is OLS. The lower part of the table reports the number of observations, the Adjusted R-squared, and the type of fixed effects included in each specification. Optimal diversity is the turning point (maximum) at which the effect of diversity turns negative (obtained from the derivative of each specification with respect to the diversity measure). The ***, **, and * marks denote statistical significance at the 1%, 5%, and 10% levels, respectively.

	(1)	(2)	(3)	(4)
Cast diversity	5.785**	12.75***	10.04***	12.35***
	(2.318)	(2.477)	(1.998)	(2.144)
Cast diversity squared	-169.5**	-338.3***	-244.0***	-345.0***
	(75.62)	(89.43)	(71.24)	(97.49)
Diversity in econ. development		-0.0190**	-0.0197***	-0.0400***
		(0.00766)	(0.00552)	(0.0131)
Diversity in trust				-0.360**
				(0.164)
Political diversity				-0.0116***
				(0.00301)
Institutional diversity				0.0932***
				(0.0299)
Optimal diversity	0.017	0.018	0.020	0.018
Observations	77,316	15,000	14,584	13,020
Adj. R-squared	0.043	0.433	0.487	0.489
Country FE	YES	YES	YES	YES
Year FE	YES	YES	YES	YES
Budget order FE	NO	YES	YES	YES
Genre FE	NO	NO	YES	YES
Language FE	NO	NO	YES	YES
Production company FE	NO	NO	YES	YES

Table 5. Movie ratings and production team diversity

The table reports coefficient estimates and standard errors (in parentheses) from the estimation of equation (2). The dependent variable is *Ratings*. All variables are defined in Table 1. Estimation method is OLS. The lower part of the table reports the number of observations, the Adjusted R-squared, and the type of fixed effects included in each specification. Optimal diversity is the turning point (maximum) at which the effect of diversity turns negative (obtained from the derivative of each specification with respect to the diversity measure). The ***, **, and * marks denote statistical significance at the 1%, 5%, and 10% levels, respectively.

	(1)	(2)	(3)	(4)
Production team diversity	26.25**	33.53***	27.22***	33.92***
	(10.73)	(9.555)	(9.435)	(6.808)
Production team diversity squared	-560.9**	-726.6***	-602.1**	-983.8***
	(260.5)	(264.0)	(253.8)	(198.5)
Diversity in econ. development		-0.0625***	-0.0527***	-0.112***
		(0.0217)	(0.0170)	(0.0346)
Diversity in trust				0.0377
				(0.378)
Political diversity				-0.00253
				(0.0136)
Institutional diversity				0.118**
				(0.0490)
Optimal diversity	0.023	0.023	0.022	0.017
Observations	20,135	5,300	5,189	4,553
Adj. R-squared	0.103	0.450	0.498	0.484
Country FE	YES	YES	YES	YES
Year FE	YES	YES	YES	YES
Budget order FE	NO	YES	YES	YES
Genre FE	NO	NO	YES	YES
Language FE	NO	NO	YES	YES
Production company FE	NO	NO	YES	YES

Table 6. Box office and diversity

The table reports coefficient estimates and standard errors (in parentheses) from the estimation of equation (2). Dependent variable is *Box office* and all variables are defined in Table 1. Estimation method is OLS and all regressions include the control variables of column 4 Table 3. The lower part of the table reports the number of observations, the Adjusted R-squared, and the type of fixed effects included in each specification. Optimal diversity is the turning point (maximum) at which the effect of diversity turns negative (obtained from the derivative of each specification with respect to the diversity measure). The ***, **, and * marks denote statistical significance at the 1%, 5%, and 10% level.

	(1)	(2)	(3)	(4)
Panel A: Total diversity				
Total diversity	110.4***	196.9***	113.6***	93.94***
	(13.38)	(27.83)	(24.22)	(17.25)
Total diversity squared	-3,633***	-5,846***	-3,518***	-2,658***
	(733.8)	(1,160)	(962.3)	(548.8)
Optimal diversity	0.015	0.016	0.017	0.017
Adj. R-squared	0.174	0.162	0.306	0.457
Observations	3,958	3,673	3,659	2,457
Panel B: Cast diversity				
Cast diversity	99.77***	154.1***	91.23***	56.03***
	(14.95)	(24.32)	(20.65)	(13.91)
Cast diversity squared	-2,893***	-4,352***	-2,620***	-1,735***
	(681.0)	(1,033)	(788.9)	(418.4)
Optimal diversity	0.017	0.017	0.017	0.016
Adj. R-squared	0.117	0.111	0.278	0.438
Observations	4,773	4,526	4,495	2,873
Panel C: Production team diversity				
Production team diversity	156.9***	183.8***	105.0***	70.11***
	(17.91)	(14.36)	(17.54)	(25.05)
Production team diversity squared	-3,702***	-4,216***	-2,696***	-1,884***
	(701.0)	(851.6)	(691.8)	(513.2)
Optimal diversity	0.021	0.021	0.019	0.018
Adj. R-squared	0.215	0.201	0.343	0.475
Observations	2,808	2,506	2,498	1,713
Country FE	YES	YES	YES	YES
Year FE	YES	YES	YES	YES
Budget order FE	NO	YES	YES	YES
Genre FE	NO	NO	YES	YES
Language FE	NO	NO	YES	YES
Production company FE	NO	NO	YES	YES

Table 7. Heckman selection

The table reports coefficient estimates and standard errors (in parentheses) from the estimation of a Heckman Two-Step Selection Model. (2). Dependent variable is *Ratings* and all variables are defined in Table 1. The outcome regression equation includes the control variables of column 4 Table 3. The lower part of the table reports the number of observations, the Adjusted R-squared, and the type of fixed effects included in each specification. We also report the inverse Mills Ratio, as well as the coefficient of the controls from the selection equation. Standard errors are obtained from the two-step correction model. Optimal diversity is the turning point (maximum) at which the effect of diversity turns negative (obtained from the derivative of each specification with respect to the diversity measure). The ***, **, and * marks denote statistical significance at the 1%, 5%, and 10% level.

	(1)	(2)	(3)
Total diversity	36.38***		
	(7.847)		
Total diversity squared	-899.2***		
	(248.6)		
Cast diversity		11.88***	
		(2.319)	
Cast diversity squared		-318.3***	
		(69.73)	
Production team diversity			-1.052
			(28.21)
Production team diversity squared			-114.5
			(764.9)
Optimal diversity	0.018	0.018	-
Inverse Mills Ratio	0.890***	0.124***	0.657*
Observations	7,308	12,763	4,547
Country FE	YES	YES	YES
Other FE	YES	YES	YES
Controls	YES	YES	YES

Table 8. IV estimations

The table reports coefficient estimates and standard errors (in parentheses) from the estimation of equation (2). The dependent variable is *Ratings.* All variables are defined in Table 1. The estimation method is the three-stage least squares (3SLS) regression. All specifications include the benchmark set of controls of column 4 Table 3. The instrumental variable is the standard deviation of a measure of *ultraviolet exposure* and its quadratic term in origin countries. In specification 1 the instrument is built using the origin countries of the overall team and cast; in specification 2 using the origin countries of the team; and in specification 3 using the origin countries of the cast. The lower part of the table reports the number of observations, the Adjusted R-squared, and the type of fixed effects included in each specification. Optimal diversity is the turning point (maximum) at which the effect of diversity turns negative (obtained from the derivative of each specification with respect to the diversity measure). The ***, **, and * marks denote statistical significance at the 1%, 5%, and 10% level.

	(1)	(2)	(3)
Total diversity	52.82***		
	(18.29)		
Total diversity squared	-837.3***		
	(264.4)		
Cast diversity		58.78***	
		(12.14)	
Cast diversity squared		-1,208***	
		(163.5)	
Production team diversity			59.73
			(239.4)
Production team diversity squared			-19,980
			(92,608)
F-Statistic	107.409	215.616	0.189
Optimal diversity	0.031	0.024	-
Observations	7,354	13,020	4,287
Country FE	YES	YES	YES
Year FE	YES	YES	YES
Budget order FE	YES	YES	YES
Genre FE	YES	YES	YES
Language FE	YES	YES	YES
Production company FE	YES	YES	YES

Table 9. Additional movie controls (IV)

The table reports coefficient estimates and standard errors (in parentheses) from the IV estimation of equation (2) after having added additional movie controls. Dependent variable is *Ratings*. All variables are defined in Table 1. Estimation method is OLS. All specifications additionally include the benchmark set of controls of column 4 Table 3. The lower part of the table reports the number of observations, the Adjusted R-squared, and the type of fixed effects included in each specification. Optimal diversity is the turning point (maximum) at which the effect of diversity turns negative (obtained from the derivative of each specification with respect to the diversity measure). The ***, **, and * marks denote statistical significance at the 1%, 5%, and 10% level.

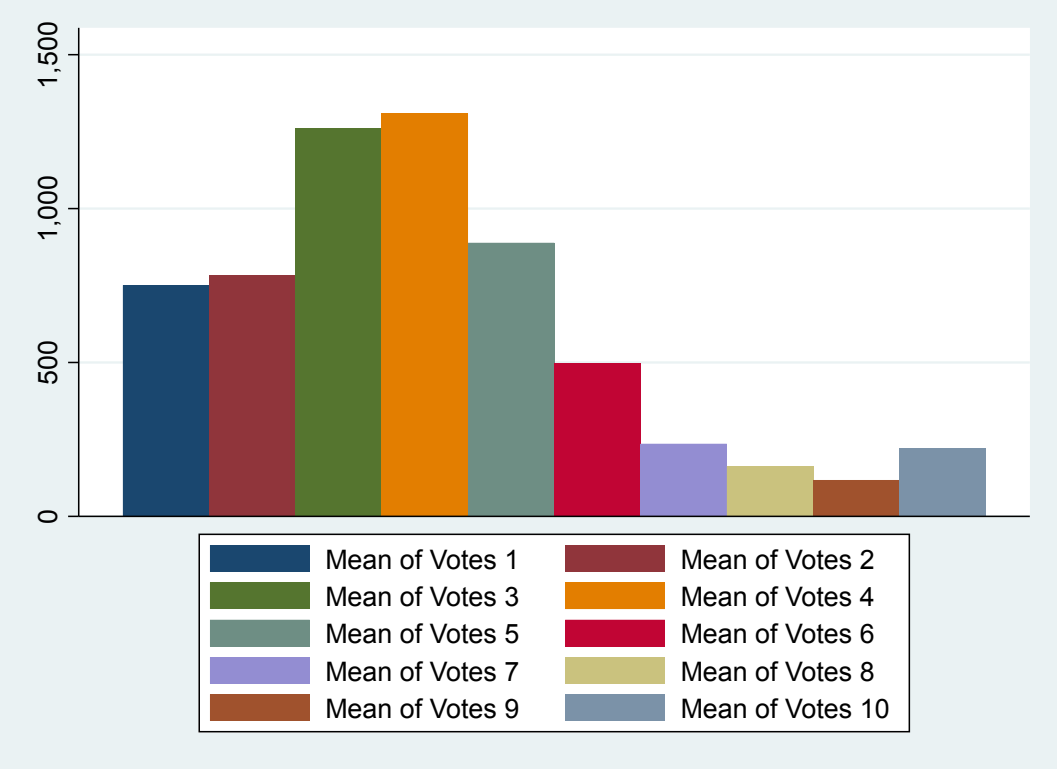
	(1)	(2)	(3)
Total diversity	51.33***		
	(17.67)		
Total diversity squared	-820.7***		
	(249.2)		
Cast diversity		54.22***	
		(11.05)	
Cast diversity squared		-1,112***	
		(117.9)	
Production team diversity			63.43
			(192.4)
Production team diversity squared			15,465
			(74,866)
Star power	0.313***	0.363***	0.391**
	(0.041)	(0.045)	(0.193)
Very high budget	0.306***	0.224***	-0.054
	(0.061)	(0.078)	(2.400)
Number of actors	0.0015***	0.001***	0.001
	(0.0001)	(0.0002)	(0.003)
F-Statistic	100.38	218.18	0.201
Optimal diversity	0.031	0.024	0.002
Observations	7,354	13,009	4,287
Country FE	YES	YES	YES
Year FE	YES	YES	YES
Budget order FE	YES	YES	YES
Genre FE	YES	YES	YES
Language FE	YES	YES	YES
Production company FE	YES	YES	YES

Table 10. Additional voters' controls (IV)

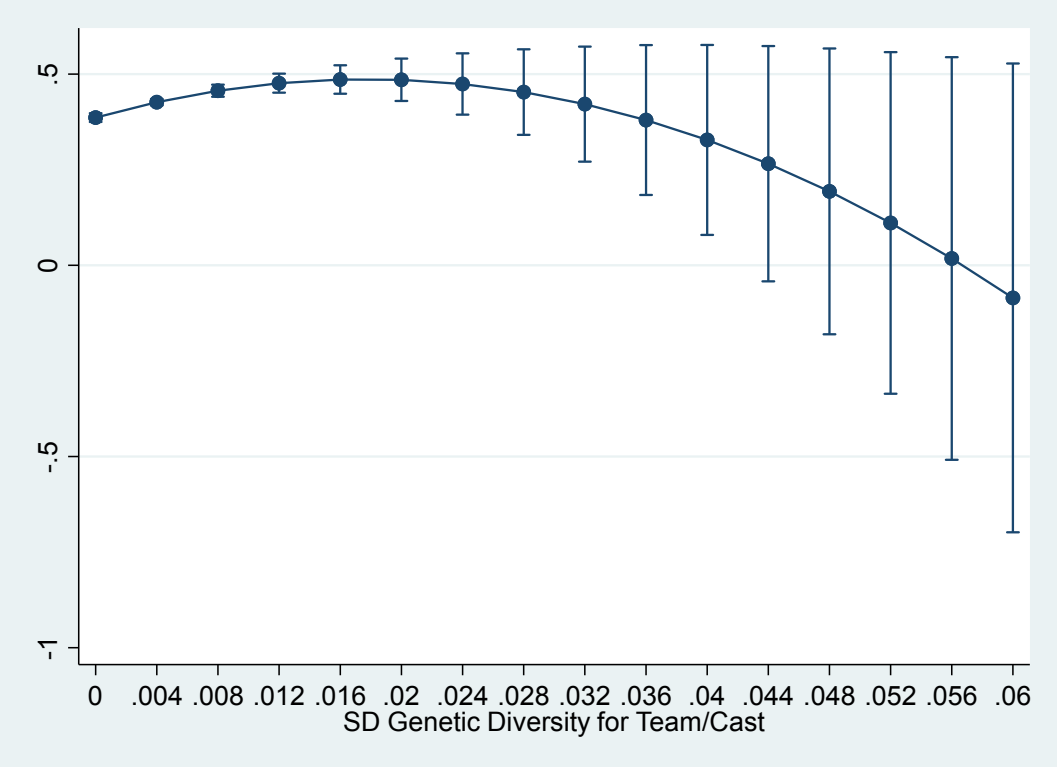
The table reports coefficient estimates and standard errors (in parentheses) from the from the IV estimation of equation (2) after having added additional voter controls. Dependent variable is *Ratings*. All variables are defined in Table 1. Estimation method is OLS. All specifications include the benchmark set of controls of column 4 Table 3. The lower part of the table reports the number of observations, the Adjusted R-squared, and the type of fixed effects included in each specification. Optimal diversity is the turning point (maximum) at which the effect of diversity turns negative (obtained from the derivative of each specification with respect to the diversity measure). The ***, **, and * marks denote statistical significance at the 1%, 5%, and 10% level.

	(1)	(2)	(3)
Total diversity	53.22***		
	(19.19)		
Total diversity squared	-908.3***	IV	
	(206.4)		
Cast diversity		53.99***	
		(10.05)	
Cast diversity squared		-1,191***	
		(139.0)	
Production team diversity			141.5
			(559.4)
Production team diversity squared			16,663
			(80,785)
% of male voters	0.141***	0.0867***	0.180*
	(0.0177)	(0.00882)	(0.109)
% of voters aged 18-29	-0.0274***	-0.0185***	-0.0562
	(0.00393)	(0.00129)	(0.107)
% of voters aged 30-44	-0.0591***	-0.0334***	-0.0354
	(0.0116)	(0.00407)	(0.107)
% of voters aged 45+	-0.0268***	-0.0160**	-0.0805
	(0.00869)	(0.00752)	(0.242)
F-Statistic	112.89	201.42	0.162
Optimal diversity	0.029	0.022	0.002
Observations	6,073	9,054	3,703
Country FE	YES	YES	YES
Year FE	YES	YES	YES
Budget order FE	YES	YES	YES
Genre FE	YES	YES	YES
Language FE	YES	YES	YES
Production company FE	YES	YES	YES

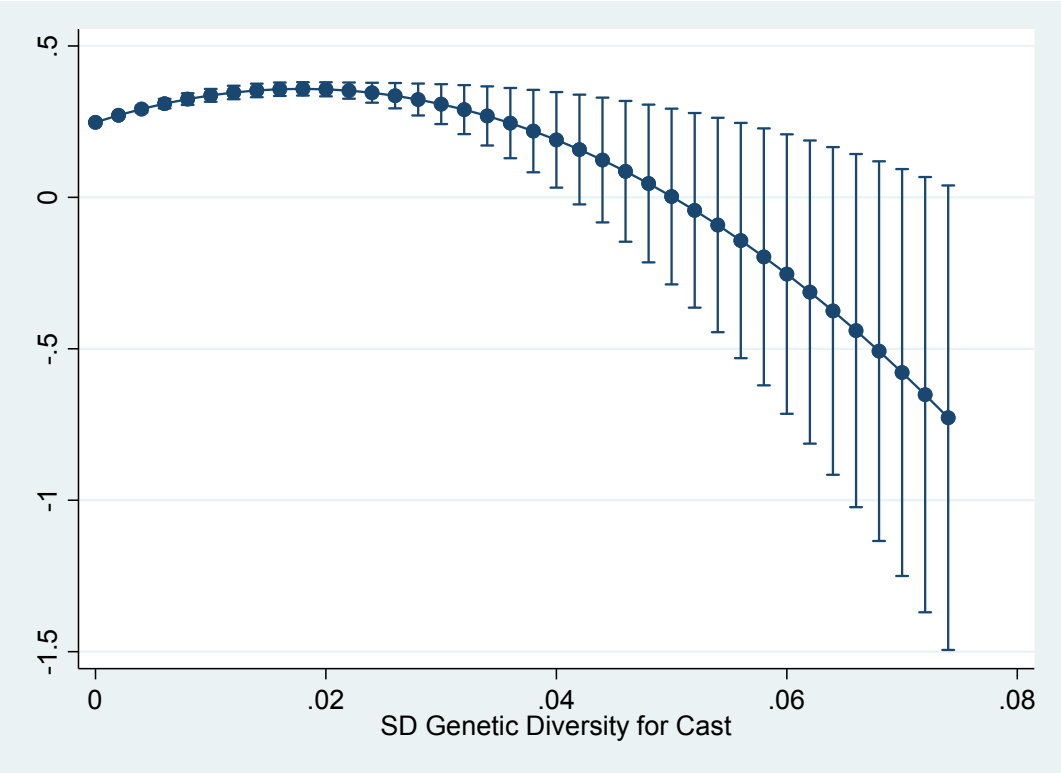
Figure 1. Distribution of votes



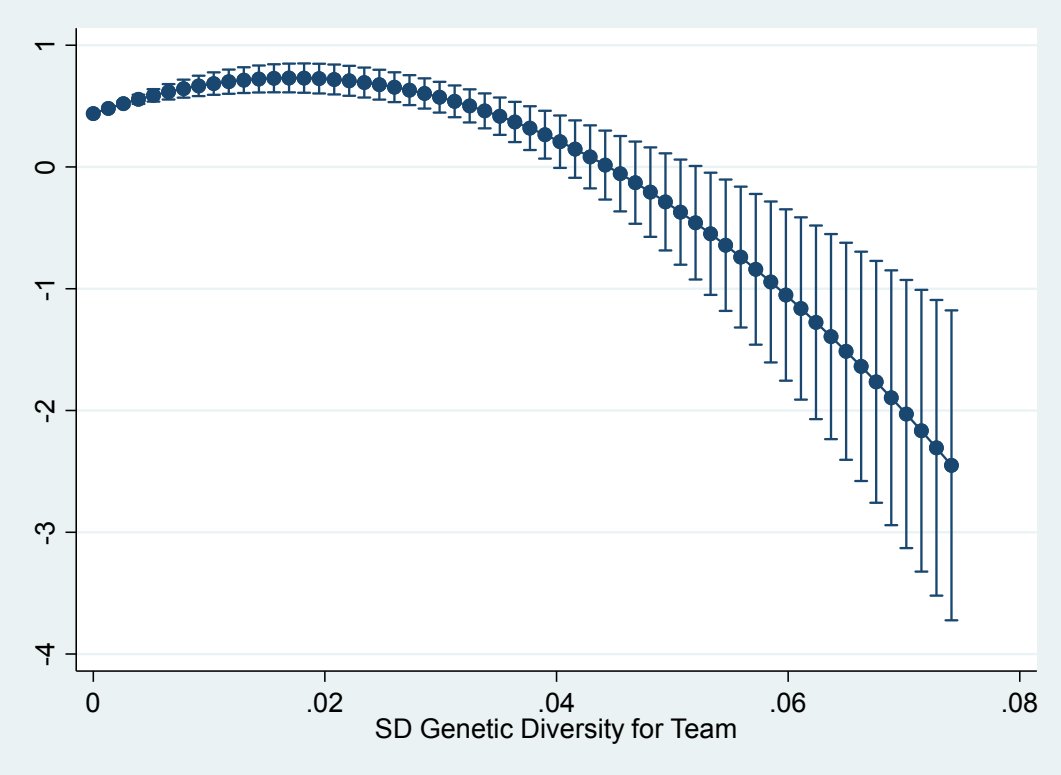
Figures 2. Graphical illustration of the benchmark results



2a. Cast and team diversity



2b. Cast diversity



2c. Team diversity

Appendix

In this appendix, intended for online use only, we provide further information and robustness tests on our baseline results.

Table A1. Movies with scores approximately equal to optimal total diversity

The table reports the movies with a diversity score approximately equal to the optimal value of Total diversity. The optimal value is derived from column 1 of Table 2.

Title of the movie	Total diversity	No. of votes 8-10
Jadesoturi	.0140408	403
Gosford Park	.0140883	2,921
Saw	.0141075	11,365
Grindhouse	.0141147	9,746
Sin City	.0141234	23,790
Ezra	.0143005	17
I Capture the Castle	.0143221	155
Hannibal Lecter - Le origini del male	.0144603	7,044
Cradle Will Rock	.0144862	602
Domino	.0145368	6,534
Casa de Areia	.0146081	115
Felicia's Journey	.014724	439
Mobsters and Mormons	.0150961	79
Disaster Movie	.0150961	60,375
The Day the Earth Stood Still	.0152341	19,324
Hard Ball	.0152341	1,152
Babylon A.D.	.0154522	9,893
Spun	.0155006	1,837
Clockstoppers	.0156171	1,839
Summer of Sam	.0156171	2,346
Hong Kong colpo su colpo	.0156739	3,156
I, Robot	.0157057	8,123
Dragon Wars	.0157299	11,215
The Triumph of Love	.0158151	170

Table A2. Linear specification

The table reports coefficient estimates and standard errors (in parentheses) of the estimation of a linear variant of equation (2). The dependent variable is *Ratings* in Panel A and *Box office* in Panel B. Estimation method is OLS. All specifications include the benchmark set of controls of column 4 Table 3. The lower part of the table reports the number of observations, the Adjusted R-squared, and the type of fixed effects included in each specification. Optimal diversity is the turning point (maximum) at which the effect of diversity turns negative (obtained from the derivative of each specification with respect to the diversity measure). The ***, **, and * marks denote statistical significance at the 1%, 5%, and 10% level.

Panel A: Ratings

	(1)	(2)	(3)
Total diversity	2.899		
	(1.815)		
Cast diversity		1.821*	
		(0.955)	
Production team diversity			3.540
			(2.648)
Observations	7,354	13,020	4,553
R-squared	0.487	0.488	0.481
Country / year FE	YES	YES	YES
Controls	YES	YES	YES

Panel B: Box office

	(1)	(2)	(3)
Total diversity	25.28***		
	(8.523)		
Cast diversity		3.340	
		(11.41)	
Production team diversity			14.47
			(19.29)
Observations	2,457	2,873	1,713
R-squared	0.455	0.437	0.474
Country / year FE	YES	YES	YES
Controls	YES	YES	YES

Table A3. Additional controls in the box office specifications

The table reports coefficient estimates and standard errors (in parentheses) from the estimation of equation (2). Dependent variable is *Box office*. All variables are defined in Table 1. Estimation method is OLS. All specifications additionally include the benchmark set of controls of column 4 Table 3. The lower part of the table reports the number of observations, the Adjusted R-squared, and the type of fixed effects included in each specification. Optimal diversity is the turning point (maximum) at which the effect of diversity turns negative (obtained from the derivative of each specification with respect to the diversity measure). The ***, **, and * marks denote statistical significance at the 1%, 5%, and 10% level.

	(1)	(2)	(3)
Total diversity	91.25***		
	(17.14)		
Total diversity squared	-2,763***		
	(499.3)		
Cast diversity		49.05***	
		(15.16)	
Cast diversity squared		-1,654***	
		(430.1)	
Production team diversity			68.79***
			(20.03)
Production team diversity squared			-1,996***
			(529.1)
Star power	0.787***	0.792***	0.833***
	(0.082)	(0.0728)	(0.112)
Very high budget	0.372***	0.295**	0.367***
	(0.116)	(0.124)	(0.118)
Number of actors	0.010***	0.012***	0.010***
	(0.001)	(0.000508)	(0.00113)
Optimal diversity	0.016	0.014	0.017
Observations	2,457	2,872	1,713
R-squared	0.480	0.462	0.498
Country FE	YES	YES	YES
Year FE	YES	YES	YES
Controls	YES	YES	YES

Table A4. Restricted dataset

The table reports coefficient estimates and standard errors (in parentheses) of the estimation using the more restrictive assumptions when constructing the dataset. Specifically, we restrict our sample to observations with at least two members of the cast for constructing *Cast diversity*; observing both the producer and the director when constructing *Production team diversity*; and observing all these four team members when constructing *Total diversity*. Dependent variable is *Ratings*. All specifications include the benchmark set of controls of column 4 Table 3. The lower part of the table reports the number of observations, the Adjusted R-squared, and the type of fixed effects included in each specification. Optimal diversity is the turning point (maximum) at which the effect of diversity turns negative (obtained from the derivative of each specification with respect to the diversity measure). The ***, **, and * marks denote statistical significance at the 1%, 5%, and 10% level.

	(1)	(2)	(3)
Total diversity	16.62*		
	(8.930)		
Total diversity squared	-585.1*		
	(301.7)		
Cast diversity		13.74***	
		(2.313)	
Cast diversity squared		-413.5***	
		(107.9)	
Production team diversity			12.38**
			(5.294)
Production team diversity squared			-402.7*
			(234.3)
Optimal diversity	0.014	0.017	0.015
Observations	1,717	12,373	1,940
R-squared	0.495	0.483	0.505
Country / year FE	YES	YES	YES
Controls	YES	YES	YES

Table A5. Double clustering of standard errors

The table reports coefficient estimates and standard errors (in parentheses) from the estimation of equation (2). Dependent variable is *Ratings* and all variables are defined in Table 1. Estimation method is OLS and all regressions include the control variables of column 4 Table 3. The lower part of the table reports the number of observations, the Adjusted R-squared, and the type of fixed effects included in each specification. Standard errors are clustered *at the country and year level.* Optimal diversity is the turning point (maximum) at which the effect of diversity turns negative (obtained from the derivative of each specification with respect to the diversity measure). The ***, **, and * marks denote statistical significance at the 1%, 5%, and 10% level.

	(1)	(2)	(3)
Total diversity	11.34***		
	(1.753)		
Total diversity squared	-320.0***		
	(121.5)		
Cast diversity		12.35***	
		(1.461)	
Cast diversity squared		-345.0***	
		(81.68)	
Production team diversity			33.92***
			(7.753)
Production team diversity squared			-983.8***
			(230.4)
Optimal diversity	0.018	0.018	0.017
Observations	7,354	13,020	4,553
R-squared	0.488	0.489	0.484
Country / year FE	YES	YES	YES
Controls	YES	YES	YES

Table A6. First stage of the IV estimations

The table reports coefficient estimates and standard errors (in parentheses) from the first stage estimations (equation xx). The dependent variable is noted on the top of each specification. The instruments are *Ultraviolet exposure* and its squared term. In specifications 1 and 2, the instruments are constructed using the origin countries of the overall team and cast; in specifications 3 and 4 using the origin countries of the team; and in specifications 5 and 6, using the origin countries of the cast. All specifications include the benchmark set of controls of column 4 Table 3. The lower part of the table reports the number of observations, the Adjusted R-squared, and the type of fixed effects included in each specification. The ***, **, and * marks denote statistical significance at the 1%, 5%, and 10% level.

	(1)	(2)	(3)	(4)	(5)	(6)
	Cast and Team diversity	Cast and Team diversity squared	Cast diversity	Cast diversity squared	Team diversity	Team diversity squared
Total diversity	45.88***	0.0767				
	(11.38)	(0.400)				
Total diversity squared	11.61	1.092***				
	(13.99)	(0.299)				
Cast diversity			43.61***	0.124		
			(7.816)	(0.271)		
Cast diversity squared			13.56	1.268***		
			(10.49)	(0.243)		
Production team diversity					52.62***	-0.00154
					(12.96)	(0.590)
Production team diversity squared					-0.0801	0.0141
					(1.371)	(0.0603)
Observations	7,354	7,354	13,020	13,020	4,287	4,287
R-squared	0.475	0.365	0.516	0.387	0.589	0.375
Country FE	YES	YES	YES	YES	YES	YES
Year FE	YES	YES	YES	YES	YES	YES
Budget order FE	YES	YES	YES	YES	YES	YES
Genre FE	YES	YES	YES	YES	YES	YES
Language FE	YES	YES	YES	YES	YES	YES
Production company FE	YES	YES	YES	YES	YES	YES

Table A7. Additional movie controls (OLS)

The table reports coefficient estimates and standard errors (in parentheses) from the estimation of equation (2). Dependent variable is *Ratings*. All variables are defined in Table 1. Estimation method is OLS. All specifications additionally include the benchmark set of controls of column 4 Table 3. The lower part of the table reports the number of observations, the Adjusted R-squared, and the type of fixed effects included in each specification. Optimal diversity is the turning point (maximum) at which the effect of diversity turns negative (obtained from the derivative of each specification with respect to the diversity measure). The ***, **, and * marks denote statistical significance at the 1%, 5%, and 10% level.

	(1)	(2)	(3)
Total diversity	10.79***		
	(1.971)		
Total diversity squared	-309.0***		
	(99.08)		
Cast diversity		10.73***	
		(1.924)	
Cast diversity squared		-301.4***	
		(88.01)	
Production team diversity			32.28***
			(8.150)
Production team diversity squared			-969.8***
			(228.5)
Star power	0.324***	0.376***	0.330***
	(0.0387)	(0.0438)	(0.0423)
Very high budget	0.319***	0.221***	0.531***
	(0.069)	(0.0823)	(0.0435)
Number of actors	0.0015***	0.00153***	0.00200***
	(0.00012)	(0.00024)	(0.00019)
Optimal diversity	0.017	0.017	0.016
Observations	7,354	13,009	4,553
R-squared	0.517	0.520	0.520
Country FE	YES	YES	YES
Year FE	YES	YES	YES
Budget order FE	YES	YES	YES
Genre FE	YES	YES	YES
Language FE	YES	YES	YES
Production company FE	YES	YES	YES

Table A8. Additional voters' controls (OLS)

The table reports coefficient estimates and standard errors (in parentheses) from the estimation of equation (2). Dependent variable is *Ratings*. All variables are defined in Table 1. Estimation method is OLS. All specifications include the benchmark set of controls of column 4 Table 3. The lower part of the table reports the number of observations, the Adjusted R-squared, and the type of fixed effects included in each specification. Optimal diversity is the turning point (maximum) at which the effect of diversity turns negative (obtained from the derivative of each specification with respect to the diversity measure). The ***, **, and * marks denote statistical significance at the 1%, 5%, and 10% level.

	(1)	(2)	(3)
Total diversity	13.17***		
	(2.303)		
Total diversity squared	-354.6***		
	(108.1)		
Cast diversity		12.68***	
		(2.041)	
Cast diversity squared		-312.7***	
		(95.21)	
Production team diversity			31.98***
			(7.144)
Production team diversity squared			-968.4***
			(214.0)
% of male voters	0.140***	0.089***	0.16***
	(0.018)	(0.010)	(0.016)
% of voters aged 18-29	-0.028***	-0.019***	-0.029***
	(0.004)	(0.001)	(0.004)
% of voters aged 30-44	-0.058***	-0.034***	-0.064***
	(0.010)	(0.004)	(0.008)
% of voters aged 45+	-0.023***	-0.016**	-0.017**
	(0.007)	(0.008)	(0.008)
Optimal diversity	0.018	0.019	0.016
Observations	6,073	9,054	3,841
R-squared	0.483	0.483	0.482
Country FE	YES	YES	YES
Year FE	YES	YES	YES
Budget order FE	YES	YES	YES
Genre FE	YES	YES	YES
Language FE	YES	YES	YES
Production company FE	YES	YES	YES