Journal of Cultural Economics
https://doi.org/10.1007/s10824-025-09557-x

ORIGINAL PAPER

Socioeconomic factors of national representation in the global film festival circuit: skewed toward the large and wealthy, but small countries can beat the odds

Andres Karjus[1,2,3] · **Vejune Zemaityte**[1]

Received: 17 October 2024 / Accepted: 22 August 2025

Abstract

This study analyzes how economic, demographic, and geographic factors predict the representation of different countries in the global film festival circuit. It relies on the combination of several open-access databases, including festival programming information from the Cinando platform of the Cannes Film Market. The dataset consists of over 20,000 unique films from almost 600 festivals across the world over a decade, a total of more than 30,000 film festival entries. It is shown that while films from large affluent countries indeed dominate the festival screen, the bias is nevertheless not fully proportional to the large demographic and economic worldwide disparities and that several smaller countries perform better than expected. Further computational simulations demonstrate how much including films from smaller countries contributes to cultural diversity, and how countries vary in cultural "trade balance" dynamics, revealing differences between net exporters and importers of festival films. This research underscores the importance of representation in film festivals and the public value of increasing cultural diversity. The data-driven insights and quantitative approaches to festival programming and cultural event analytics should be useful for both the academic community as well as film festival organizers and policymakers aiming to foster more inclusive and diverse cultural landscapes.

Keywords Film festivals · National representation · GDP · Cultural diversity · Gender balance · Small countries · Cultural trade balance

JEL Classification Z11 · F14 · R12 · O57

✉ Andres Karjus
akarjus@tlu.ee

[1] Tallinn University, Tallinn, Estonia

[2] Estonian Business School, Tallinn, Estonia

[3] University of Tartu, Tartu, Estonia

Published online: 16 September 2025

1 Introduction

Film festivals are a vital component of the film industry, occurring in various forms worldwide, ranging from major industry hubs like Cannes, Sundance, or the Berlinale, to regional and smaller national events. These festivals differ in their budget, audience size, specialization, acceptance rates, and diversity, but form a well-connected global network in terms of the flow of recurring films, filmmakers, and industry stakeholders (Krainhöfer, 2018; Loist, 2016; Zemaityte et al., 2024). They offer filmmakers a venue to showcase and market their productions, which can determine their further success (Stringer, 2001; Falicov, 2016). Festivals also promote cultural interactions and dialog between makers and audiences, public and industry interests, and generate public value (in the sense of Benington and Moore, 2011; O'Regan and Potter, 2022; Mazzucato et al., 2020) to societies and public spheres, by showcasing culturally varied programs that broaden public access to films and foster international cultural learning and exchange (Harbord, 2002; Rüling & Pedersen, 2010; Valck & Marijke, 2007; Roy, 2014; De Valck, 2016a; Mair & Duffy, 2020; Diestro-Dópido, 2021). Increasing diversity in cultural landscapes, including by platforming otherwise low capacity or underrepresented voices and potentially novel perspectives, can in turn foster societal resilience, innovation, and transformation (Farrell, 2016; Burgess, 2020; Hartley et al., 2020; Peralta García & Simour, 2025). Cultural exports, including films, can also act as soft power (Guan et al., 2023).

While balancing between various pressures—like marketability, artistic aspirations, activism, diversity, and audience choice—can be challenging in terms of curation (Bosma, 2015), power over programming means festivals can also bypass traditional distribution limitations and address topical issues (De Valck, 2016b; Elsaesser, 2005). The role of film festivals also extends beyond exhibition and cultural learning. Like other cultural festivals (Wilson et al., 2017), they boost local economies, create financial and symbolic value for host countries and regions via public funding and subsidies, create employment opportunities, and increase revenues via tourism (Monson et al., 2023; Locarno Film Festival, 2022; Grunwell & Ha, 2008; Harbord, 2002; Kendall et al., 2021; Kostopoulou et al., 2013). Hybrid formats have also served as artistic platforms during global crises such as pandemics (Smits, 2024; Odabasi, 2025)

But the festival circuit is not a level playing field. Countries that are smaller, less affluent or whose language is not among the few international lingua franca—naturally face various challenges when it comes to the production as well as (local and global) distribution of cultural products, including films (Hjort & Petrie, 2007; Martin-Jones & Montañez, 2013; MacPherson, 2010; Fu & Sim, 2010). The dynamics of programming content from large global producers and small countries in festivals, cinema, and television, including questions around the value of cultural diversity, are therefore an active research area (Ibrus and Ojamaa, 2014; Sand, 2019; Félix-Jäger, 2020; Navarro et al., 2022; Zemaityte et al., 2024; Ibrus et al., 2023; Zemaityte et al., to appear, see example). This contribution makes use of several sources of open data, including a refined version

of a snapshot of the Cinando database of international film festivals (Zemaityte et al., 2023, 2024), further described in Methods and materials below. A first glance at the data indicates that the global film festival circuit is rather biased toward large, wealthy nations (Fig. 1). Almost a quarter of the films in the circuit in 2012–2021 listed either France or the USA as the production or co-production country, and over a quarter of the festivals took place in these two countries (see Fig. 1c, and the blue crosses on Fig. 1a). The database includes films from 163 countries, but half the films in that period were made in or involved one of the top 8 most productive countries.

However, smaller countries likely have proportionally fewer filmmakers, and less prosperous nations have fewer resources to invest in film production. Iceland cannot be expected to match the volume of American film production; Switzerland and Latvia are physically closer to more international festival locations than Malaysia or Uganda (Fig. 1a). Films from the UK or Australia may be easier to follow culturally and language-wise in many places due to being in English (Chung, 2011). Filmmakers or stakeholders in countries distant from the current hot spots may find fewer opportunities for travel to showcase and advertise their products. India is a large film producer, but not all Bollywood films might make it to the mostly West-centric festivals nor be aimed at them (Viswamohan & Chaudhuri, 2022). First-world countries simply have more resources to drive their film industry and local cinema-going audiences with higher disposable incomes (Lorenzen, 2007). Yet if economic and demographic disparities lead to unequal representation, it can undermine the potential for

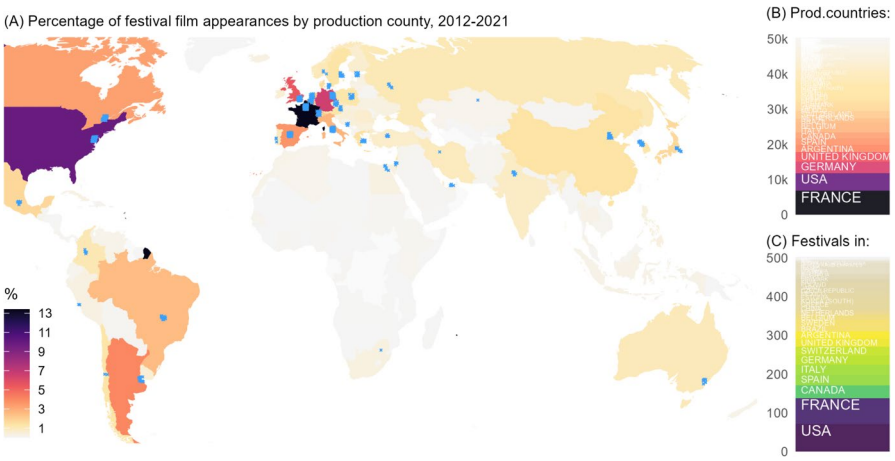

Fig. 1 Mapping the film festivals and production countries across the globe. Panel **a** shows the volume of film festival pair entries in the database (indicating one or more screenings at a given festival) in the circuit from a given country, and festival locations as blue × crosses, positioned at capital city coordinates of hosting countries (due to lack of more precise location data). Slight noise is added so that overlapping festivals are visible; tightly "packed" blue groups indicate more festivals in the given country. Panels **b** and **c** are stacked bar charts, showing the cumulative number of pairs and hosted events by country, respectively. Both stacks are dominated by the USA and France, followed by a long tail of less prominent production and host countries. The smaller countries with lesser shares are explored on subsequent graphs

film festivals to democratize access to diverse cultural expressions and public value available to global audiences.

This short contribution attempts to shed light on these dynamics and fill a gap in film festival research literature by directly measuring national representation at festivals, and comparing this to countries' economic and demographic factors, as well as providing a quantitative cultural analytics of the flows of festival films between countries of the world. Until recently, film festival research has remained mostly theoretical (e.g., de Valck, 2007; Loist, 2016; Wilson et al., 2017), qualitative or participant experience focused (Peirano, 2020; Mair & Duffy, 2020; Sand, 2019), or based on smaller data samples (Ehrich et al., 2022). Following more recent research (Zemaityte et al., 2024; to appear), this contribution study aims to leverage big data, systematic quantitative analysis, and predictive modeling. Particular attention is paid here to small countries, how they fit into the festival ecosystem, and how much diversity and therefore public value they add to the circuit. In summary, the following research questions are addressed:

- RQ1: To what extent is the festival circuit skewed or biased toward populous and wealthy countries?
- RQ2: To what extent are factors like relative wealth and population size predictive of national representation in the film festival circuit? What characterizes countries that perform better or worse than expected?
- RQ3: How and how much do countries differ in terms of the cultural "trade balance" of festival films?
- RQ4: How would platforming more films from smaller countries affect diversity of the festival circuit?

2 Methods

2.1 Film festival data

This research makes use of three large open-access datasets and two additional sources of information (see Data availability). The primary source is the Cinando dataset (Zemaityte et al., 2023), published along with Zemaityte et al. (2024). It is a cleaned version of a static snapshot of the proprietary relational database underlying the namesake platform operated by the Cannes Film Market (also known as Marché du Film - Festival de Cannes). The 2012–2021 subset used for this study is operationalized as a set of 50,408 data points, where each point is the listing of a film at a film festival (the database does not specify how many times a listed film was screened at a given festival, but at least once can be assumed). If there are multiple co-production countries, then the listing is repeated for each to include the country metadata (but this is taken into account in the analyses). Ignoring co-productions, there are 34,535 such film festival data points, in turn reflecting high levels of co-production (cf. Bondebjerg, 2016; Parc, 2020). There are 22,167 unique films in this sample (approximately; see below), 592 festivals in 129 event series, 148 primary

production countries (listed first in the ordered producer variable in the database) and 163 countries in total, i.e., 84% of the world's countries occur at least once.

While the Cinando snapshot used here is currently one of if not the most extensive global dataset available on the programming of international film festivals, it comes with limitations stemming from its initial European origins, having started as a service solely for the Cannes Film Festival. Using it has arguably later become an industry standard, however, reflected by the wide range of festival events recorded there and correlating with UNESCO production estimates (see below). The potential biases include possible under-representation of non-European events and industries, and analyses based on it risk suffering from "survivorship bias," as the database or at least the version available here does not record rejections by the programming committees, only programs of successfully selected films. The initial years of the database are also sparse, reflecting its gradual uptake by the industry. Zemaityte et al. (2024) used a subset of data between 2009 and 2021 for most, and from 2012 for more conservative analyses. The latter approach is taken here for aggregated data, as including 2009–2011 risks biasing the results toward France and French films by showing larger shares and counts than would be representative of the subsequent decade.

Most of the data appears to be entered by platform users, requiring extensive data cleaning described in Zemaityte et al. (2024). One shortcoming that remains is the identification of individual films, as films have been often entered into the database multiple times using different ID values. Here, films are further identified using the title and the reported production year, but both of those values are also known to occasionally change during film production, and some films have short single-word titles that can easily be confused. Another shortcoming is festival dates, limited to years in the dataset, which in turn constrains temporal analyses. The cultural flows section in the Results only considers direct links between film production and festival host countries but not the journeys of films through the festival circuit, including getting selected for a prestigious festival that may affect subsequent selections. This would require data with more precise festival dates than what is currently available.

Another technical detail concerns co-productions. The database does include an order variable for metadata, including production countries, but this is difficult to interpret systematically and is ignored here as in Zemaityte et al. (2024). The existence of multiple producers is accounted for the analyses however, but doing so requires choices about aggregation to match the goals of a given analysis. In the study on representation below, each listed country is counted as one full entry, whereas in the analyses on cultural flows and diversity, they are weighted, so that for multiple co-producers n, the weights equal $1/n$, i.e., sum to one (see Results section).

The available data only includes programs, i.e., already selected films at festivals—there is no information on acceptance rates and rejections, which would be interesting to investigate of course if available. What can still be modeled is aggregated counts of film festival pairs per country, which is done below. The current count approach ignores individual festivals, and it is not possible to model effects such as festival types and distance of the festival from the production country, nor to control for the festival year (the period is just under 10 years however). There are of course other ways to aggregate, but each has its downsides. The percentage

of countries in a festival is one option, but that would give undue weight to small festivals. Binomial representation, i.e., just counting whether a country is present in a festival or year or not, would enable modeling more detailed metadata, but that would make the results equivalent between having a single film at a festival from a given country and films from that country filling the entire festival program (and be largely a dataset of zeroes, especially if done per year, given that many countries in the dataset only occur a few times).

In the linear regression modeling below, counts on a logarithmic scale are used as the response variable. Another option would be a Poisson regression or a negative binomial model, given the large discrepancy between a few large counts and a long tail of small values. The log-linear approach lends itself to a more convenient interpretation, especially given that all other variables are similarly power law distributed and also on the log scale. For the representation and bias analysis, the FIAPF or International Federation of Film Producers Associations accreditation list is used as additional categorization (FIAPF, 2023). This list represents the so-called "A-list" top competitive feature film festivals at the center of the festival industry, widely recognized by filmmakers, although it does not include all large and otherwise significant event series such as Sundance. Most of the festival series in the dataset, 115 out of 129, are B-listers, however.

2.2 Socioeconomic and film production data

The two other datasets are the World Bank (WB) database (World Bank Group, 2024), which is used to source information on population and gross domestic product (GDP), as in Zemaityte et al. (to appear), and the UNESCO Institute of Statistics database on the counts of feature films produced by country (UIS, 2024). Both were aligned with the Cinando dataset by country name (manually correcting names where there was a mismatch in naming formats) to support exploring RQ2. Where data for a year was missing for any country, linear interpolation was used to predict the values from adjacent present years.

The UIS production numbers are not directly comparable to the Cinando festival programming data, even if the latter were transformed into unique film counts. The data collection methods are clearly different, and UIS counts only feature films, whereas festivals feature various formats including short films. Therefore, instead of measuring over or under-representation in absolute numbers, a correlation between the magnitudes is instead provided in the Results section. This relationship could be further analyzed in future research.

These data from the additional databases were added by production country to the final dataset, but aligned by the year of the festival event, not the year of film production. This may be a debatable choice but hopefully leads to more informative results. One argument is the existence of retrospective festivals or program segments. As an extreme example, the socioeconomic situation of the United Kingdom in 1927 probably has little bearing on the screening of the film "The Ring" from that year in Cannes in 2012. The economic situation of a country could itself be a factor, however, when it comes to representation in a given year in the festival circuit.

Furthermore, films typically stay in the circuit only a few years after production anyway, including due to rules of admission in many festivals (Zemaityte et al., 2024). In the Cinando subset, the average difference between the year of production and festival occurrence is just 0.7 years (both variables are at yearly resolution, so this is a somewhat rough estimate).

These should be considered however only rough estimates of more interesting target measures, which would be the number of filmmakers in a country and its media market size, and the financial means, revenue, and subsidies of a country's film industry. Total population is a proxy, but of course, it cannot be expected that the number of filmmakers is proportional in every country. GDP is the total value of produced goods and services produced but says nothing about wealth distribution or investment in creative industries. Since GDP and population typically correlate, the measure used in the analyses here is GDP per capita, i.e., divided by the population estimate of a given year, as a rough proxy of relative prosperity. These variables are averaged for the analyses below for the 9-year observation period, weighted by the number of film festival pairs in a year: this way the aggregate for each country is most representative of when it was most present in the circuit.

While the preceding literature discussed in the Introduction has discussed and compared small and large countries, using continuous population and GDP variables here means there is no need to categorize or bin countries into any arbitrary discrete categories such as small-medium-large, which in turn could easily bias analyses down the line. Unlike Zemaityte et al. (to appear), which this analysis otherwise complements, the models here of representation, bias, and diversity all use the continuous scales of the aforementioned variables. A country as such is not necessarily an ideal unit either, as many countries may include regions with different population levels, affluence, or access to arts education. Here, country and nation are used interchangeably for simplicity, but many countries of course consist of multiple ethnic and cultural groups, which again may be associated with differential access to resources. Another simplification concerns location and distance. The Cinando dataset only includes production and festival host information at the country level, so capital city locations and inter-distances are used where needed (including the map in Fig. 1a and the regression model in the results).

2.3 Measuring diversity

The diversity section in the Results below follows the method of diversity calculation and the open source codebase of Zemaityte et al. (2024), who argue against using discrete labels for calculating diversity when the distances between categories are not equidistant (cf. Moreau and Peltier, 2004). This is relevant in data like thematic metadata tags such as drama or horror, where some genres are inherently more similar than others, or languages of the world, which also vary in similarity and relatedness. The delineation of language, variety, or dialect is often more political than anything else. For example, a festival featuring films in Dutch, Flemish, and German would be as diverse on paper as one featuring Japanese, Zulu, and Finnish, if one were to simply count the labels. Yet the first set features films in

closely related languages, and by proxy, likely produced in culturally more similar contexts. The proposed solution is to use latent spaces, induced directly from the tag co-occurrence data for genres (analogously to word embeddings), and an externally sourced language typology vectors dataset for linguistic similarity (Malaviya et al., 2017). The distances in the latter embedding space latter align well with perceptual language similarity and genealogical relatedness. The Cinando dataset includes a genre or thematic variable, but the set of 41 tags is suboptimal for the intended diversity analysis here, as it consists of a mix of genres (drama, comedy) but also medium and production-related tags. There is also little information about cultural diversity to be gleaned from it, as the majority of films are just tagged as drama and a few other common genres.

The continuous space approach enables the calculation of diversity as the average distance from a global mean vector (analogously to mean absolute deviation or MAD). The mean is normalized by the maximum possible distance in a given space, yielding a metric in the [0, 1] range. The intuition is that a distribution tight around the (multidimensional) average indicates low diversity—everything is close to the average. A high mean distance from the average indicates a wider spread and therefore higher diversity. Diversity can be conceptualized in various ways: this operationalization is meant to capture the macro-level, ecosystem-wide, or "external" diversity (following McQuail and Van Cuilenburg, 1983; Zemaityte et al., 2024). As with any metric based on means, including MAD, this naturally assumes a (here multivariate) normal distribution. Given the normalization, a value of 1 is practically only possible in a bimodal distribution, e.g., a system consisting of only two of the most different but frequent languages—which is probably no longer an intuitive representation of diversity, nor a likely policy goal. This limitation of the metric as it approaches its maximum should be kept in mind, but is not found to be an issue in the analyses below.

For gender diversity, simply the fraction of women filmmakers (directors and producers, the most commonly present crew metadata) is used here. This is based on their first names as listed in the Cinando database, matched against a binary gendered names database (Bérubé et al., 2020), as in Zemaityte et al. (2024). The binary is a database limitation and does not reflect the views of authors on gender as such.

2.4 Simulating diversity in variable festival programming scenarios

The simulation results described in the Results, corresponding to RQ4, are based on the diversity measures discussed above. The simulation is essentially based on the concept of bootstrapping, and is designed to explore different potential scenarios of film festival or circuit balancing. What follows is a description in prose, while the more mathematical description can be found in the Supplementary. A parameter grid of 14×14 points is distributed uniformly in logarithmic space across the population (1M to 330M) and GDP per capita ($1200 to 60,000) axes of the dataset (see the Supplementary for a visualization). Weighted sampling is used to construct 196 distributions of film production countries with population and relative GDP averages varying across this space, each replicated 100 times (and their results averaged)

for reliability. In short, this enables asking questions such as, what if a given festival or the entire circuit would favor and platform more films from large but low-income countries across the world, instead of the current wealthy western focus? What would a circuit platforming more small countries look like in terms of linguistic and gender diversity?

Regarding the parameter space, a significant simplification was deemed necessary in the end, namely to limit the population range so that the USA is included, but larger countries (i.e., China and India) are not, in terms of primary production status. This is simply due to their extreme outlier status in the global population–GDP space: both countries combine exceptionally large populations with comparatively low per capita GDP, which places them well outside the part of the simulation grid that is otherwise relatively dense in terms of distributional support. As discussed below, these two countries may well be also structurally underrepresented in the Cinando dataset that is used as a basis of the simulation. While the parameter space could theoretically be extended to accommodate the outliers, doing so would isolate these countries in two sparsely populated corners of the grid, with no other comparable cases and large intervening regions devoid of empirical support. Since the simulation design relies on (log) normal sampling around each grid point (target distribution means), a meaningful sampling in such isolated regions would not be feasible. In short, this outlier removal enables the rest of the design to work as intended. A more representative global simulation could be explored in future research. The current simulation and its interpretation remains necessarily limited in this aspect, but the country distributions and their resulting diversity values are still informative.

For each combination of population and GDP per capita target means, a two-step sampling procedure was applied to draw film production countries from the dataset. To keep things simple, most of the simulation parameters are fixed for now, and only the means on the two axes are manipulated. The target sample size was set to 500 films (but adjusted to account for boundary effects, yielding a median of 368; see below). Sampling uses truncated log-normal distributions with means at the given parameter values, and a fixed standard deviation of 0.3 (on the log scale; see the Supplementary for exploration of lower and higher values). This was observed to yield reasonable distributions, as much lower SD values would lead to unrealistic narrow samples of only a few countries (not representative of the simulation intention of a global festival circuit), and high SD would yield uninformative samples pulled toward global averages. Both variables being limited by zero on the one end (no such thing as a country with less than zero people or GDP) and practically by the largest countries on the other end, a truncated distribution is necessary. Even so, if either the target mean is very close to the boundary, or the SD is large enough that the sample tends toward uniform, the target and real sample mean would begin to diverge and the simulation would become uninformative. As for co-productions, these are simplified by considering only the population and GDP of the firstly-listed production country, but still taking into account all languages and crew member genders attached to the film when calculating sample diversity.

The first step of the sampling procedure, on each replicate of each target combination, was to draw a near-uniform stratified subsample of up to 20 films per country was drawn from the Cinando dataset. If the entire dataset would be used for

sampling without stratification, the simulations would simply reproduced the biases therein, as films from countries such as the USA and France would continue to be most likely to be sampled. Increasing the maximum per country would also begin to bias it, as not all smaller producers have even that many films present. While each individual replicate sample is small (much smaller than an annual circuit in the Cinando dataset), the repeated sampling across replicates allows the noise to average out. This bootstrapping approach yields stable estimates of diversity that can be interpreted as circuit (or festival program) values for a combination of countries centered at each bivariate target mean.

The second step was to calculate sampling weights for each country. This was done by constructing two truncated log-normal distributions, for population sizes and GDP per capita, with a constant SD and means defined by the target parameters. The density values from these two distributions were normalized and multiplied to produce joint sampling weights, and as the third step, used in weighted sampling (with replacement) to produce the target sample of about 500 films and their production countries. When sampling, the values attached to the production year of the film for a given country are used (without adjusting for inflation, given the short time span).

The final sample size is adjusted, again due to the truncated nature of real-life distributions, as well as the fact that certain regions of the joint population–GDP space can be sparse, with fewer countries falling within some parameter ranges. For example, if sampling from a segment of small population *and* low relative GDP, there simply are not that many countries and films to sample from. Setting these values at 1M and $1200 would yields a sample of which a fourth of the 500 consists of repeatedly sampled films from Bhutan, of which there are only six to begin with. The scaling factor reduces this sample size to around 130, with around 20 unique countries, and the effect of being on the distribution boundary typically shift the empirical means to 1.8M and $3000. This small sample can be either filtered out as non-representative, or marked as such on a graph. The example values above have been approximate as they depend on the sampling; hence the bootstrapping approach of averaging over many replicates. The grid of empirical (replicate-averaged) means resulting from the full simulation, reported in Results, is slightly distorted from its original uniform shape exactly due to these sampling and boundary effects.

To illustrate with another example: a parameter intersection around 9M population and 10k per capita, a region with strong support across for both distributions, yields a sample with an (across replicates) average of 511 films (here the adjustment slightly increases it above the target) from 42 countries. For example, one linguistic simulation replicate sample (also of size 511) contained 45 countries, with the most frequent being Serbia (31 primary production entries), Bulgaria (29), Hungary (29) and Cuba (26)—all countries roughly in that size and relative GDP range. The most common languages out of the 46 present in the sample were Spanish (from various Latin-American countries, present in 153 entries), English (74), Hungarian (35; both of the latter two values likely increased by co-productions or multilingual films), and then a number of related Slavic languages including Serbian, Russian, Bulgarian, Slovakian, and Czech. The linguistic diversity value in this replicate is a moderate 0.44, as most languages are more or less closely related Indo-European ones;

Hungarian is genealogically Finno-Ugric, but naturally shares many areal European features with its neighbors, which the typological similarity measure picks up on.

The (log) normal distribution sampling intentionally creates a scenario where countries around the bivariate target means are most frequent, but the moderate SD ensures a smooth distribution that should emulate a circuit where certain socioeconomic ranges are boosted but other are not necessarily entirely excluded. In this replicate sample, there is also for example 9 films from Poland, 3 from Latvia (a much smaller country), and one from Mexico on the other end of the population scale.

3 Results

This section is organized by the research questions proposed in the Introduction, starting with the question of bias, followed by statistical modeling of representation affecting factors in the film festival circuit, modeling of cultural flows, and finally an assessment of estimated diversity contributions by countries of different socioeconomic profiles.

3.1 Representational balance and bias in the film festival circuit

Quantifying balance or bias depends on the unit of analysis, and a potentially ideological question of what counts as balance, equality, or equity. The result presented here is rather agnostic, offering two perspectives on balance across two variables. As discussed in the Methods section, the unit is the (co)production country of a film programmed at a festival; multiple appearances of a film at different festivals are counted as separate data points. Figure 2a, b shows the distribution of country appearances by their respective Population and GDP per capita (according to the WB). Both are on logarithmic scales, given widely known Pareto distributions of both variables. The same data are shown by individual countries in Fig. 3a. The means (black bars) are therefore geometric rather than arithmetic averages. The figure is further split by festival accreditation categories: all festivals, the prestigious competitive A-listers, and the rest (B-category) festivals. There do not appear to be particularly large differences between them on average, however. There is also not much difference in regions of the world, as most regions contain both small and large countries (see the complementary graph showing distributions by region in the Supplementary Materials appendix). Figure 2c shows the decadal trajectories of three A-list and two B-list festivals, demonstrating relative stability in the two variables over time. The Tallinn Black Nights International Film Festival is an A-lister, yet arguably smaller and less prestigious than Cannes or Sundance—the latter itself a B-list yet widely known and influential festival series. Black Nights, taking place in the 1.3 million country of Estonia, also appears to program films on average from smaller and less prosperous countries compared to the others.

The two models of balance are labeled “uniform” and “proportional” (Kedar et al., 2016; Kurz et al., 2017). The uniform expectation is calculated as the mean value of a simulated distribution of appearances where each country in the world

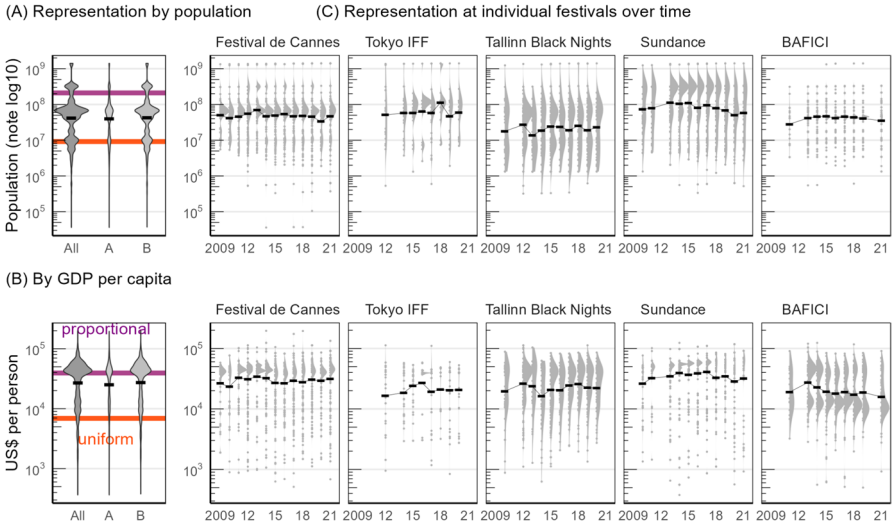

Fig. 2 Representational equality in the global film festival circuit. Panels **a** and **b** on the left depict the circuit distributions (densities) and averages (black bars) in 2012–2021 for population and GDP per capita of production countries (each underlying data point is a festival listing of a film from a given country), and the two models of balance as horizontal lines (see text for discussion). Note that the axes are logarithmic, and the bars are means of log-scaled values. The small panels in **c**, aligned with the axes of panels **a** and **b**, illustrate these values over time for a selection of individual festivals; the first three are FIAPF-accredited competitive event series, while Sundance and BACIFI (Buenos Aires International Independent Film Festival) are not. The graph shows most festivals stay relatively stable in terms of the average population size and affluence of production countries of films that get selected

(that has participated in the festival circuit according to Cinando), regardless of size and prosperity, gets to show the same number of films. These values are below the real-world means, again related to the power law: there are simply fewer very large or prosperous countries and smaller ones. This model, while not necessarily the most realistic, demonstrates what (the average of) a festival circuit, blindly balanced by countries as units, would look like (or visually: where the black mean bar would be if the representation was uniform in this manner).

The second, proportional model is operationalized as the mean of a distribution where the number of festival appearances per country is perfectly proportional to their respective Population (Fig. 2a) and GDP per capita (B). The purple line shows where the mean (black bar) would be in this scenario. This illustrates a possible world where each country trains filmmakers and hosts film production and exhibition industries commensurate with their population size, or produces films proportionally to their relative GDP. Here these factors are modeled separately for simplicity; their interaction is explored below. These values are higher than the real-world averages for both variables, demonstrating that the circuit is, from this perspective, indeed not as biased toward the larger and wealthier as it could potentially be. However, if one were to imagine a perfectly population-proportional international festival of say 100 films, then about 37 should be from China or India, 4 from the USA, and the rest of the 59 slots split between the rest of the 163 countries in the circuit.

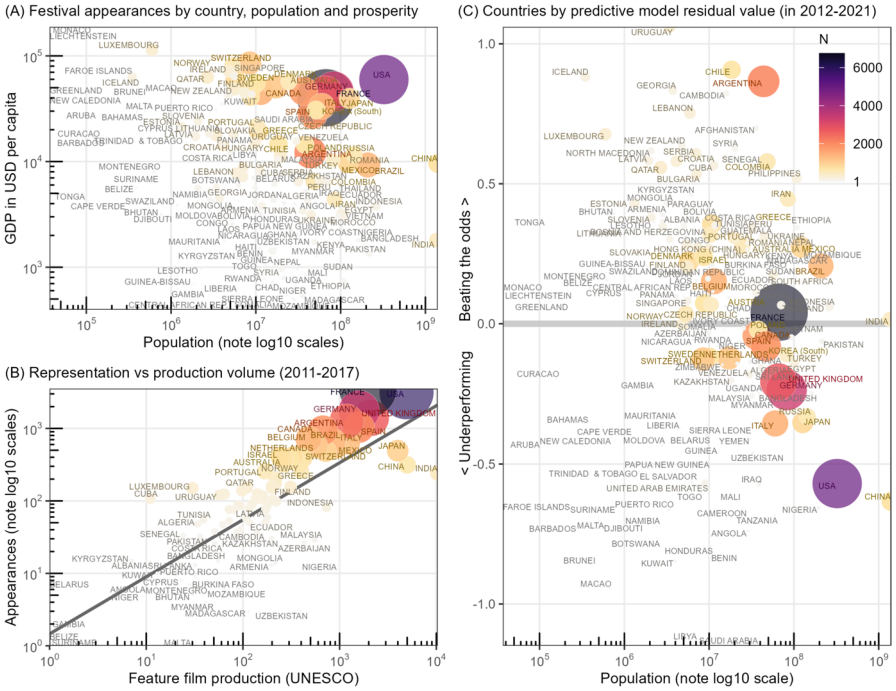

Fig. 3 Three perspectives of festival film producing countries. Panel **a** has all the countries arranged by population and GDP per capita, with size indicating the number of film festival pairs. As expected, the optimum is a wealthy economy and large population, while large African and Asian countries remain very much underrepresented. Panel B compares the number of film festival pairs in the Cinando dataset and the UIS feature film production estimates in the overlapping period, showing they roughly correlate. Panel **c** shows the population on the x-axis and the residuals from the representation prediction model on the vertical: countries above the zero line occur more often than would be expected on average, while those below do worse given their population, GDP per capita, and other factors discussed in the text. The Cinando dataset counts are based on 2012–2021 data on panels **a** and **c**, and 2011–2017 on panel **b**, to match the UIS data in the latter

Depending on one's balancing goal, proportionality, diversity, and other goals should likely be taken into account to some extent.

The answer to RQ1 is therefore: it's complicated. If one were to take proportional balance as a lens to view the world, the bias actually goes the other way, as smaller countries would appear favored. It is also only a global outlook, ignoring the functions of local or regional festivals (see also the Supplementary for region-by-region models). Given the nature of this database, it is also difficult to ascertain whether the current distribution rises from festival organizers selecting relatively more films from smaller countries and less prosperous regions, or higher-resource countries not producing or submitting as many films as they could. From yet another perspective, the film festival ecosystem as a whole could be seen as already managing to strike a middle ground between these two conceptualizations of equality—whether by default or by design—with different festivals

screening different content and catering to different international, regional and thematic audiences (cf. Zemaityte et al., 2024).

Figure 3 provides three additional views into the dataset. Each country is shown as a marker, the size of which reflects the number of produced or co-produced film listings in festivals in the circuit. Reflecting the map in Fig. 1, most of the film sources are indeed clustered on the ends of the scales in Fig. 3a, with France and the USA the largest. The question remains, how well this particular festival database reflects country-wise film production capacity, and what could be said about the relative festival success of a given country without festival acceptance and rejection data. When compared with the UIS database (see Methods and materials), the correspondence might be close enough, depicted in Fig. 3b. The range of the latter database is narrower, so the Cinando data are also cropped accordingly, but the positive correlation (on log scales) is rather strong, especially given the disparity between the two databases discussed in Methods (adjusted $R^2 = 0.57$, i.e., 57% variance in festival appearances is described by the production numbers). Most of the largest producers perform even better than expected relative to the others (points above the regression line). The largest countries in the world, China and India, appear both as relatively underrepresented, given their production volumes. As discussed in Methods, it is difficult to tell at this point on these data alone (but would be interesting to know) whether this is because their filmmakers do not apply to that many festivals, if their films are not selected, or if they do appear in festivals but just not those recorded in the Cinando dataset. Not all films in all genres are of course produced with festivals in mind, and potential regional differences in this dimension may also affect these results.

3.2 Predicting festival performance

Due to the data-specific limitations and those of aggregation discussed in Methods, it is not possible to model festival selection processes directly here, but factors of representation based on country aggregates can be quantified. Figure 3c depicts the residual values from a linear regression model operationalizing RQ2, predicting the ($\log_{10}$) number of appearances of (co)productions n in the circuit per country. The predictors in the model are the population and GDP per capita of each country (centered at 10M and 10k, respectively), the number of festival events hosted in the country (with +1 Laplace-smoothing to allow for log-scaling), and distance from France in thousands of kilometers. All but the last are modeled on the $\log_{10}$ scale, with an interaction allowed between population and prosperity. The reasoning for the event count variable is that countries may be more successful if they host their own festivals, which may be more favorable toward local films. The distance variable reflects the observation that the festival circuit, as recorded in the database, is rather Euro-centric, which may provide advantages to nearby countries in terms of ease of travel (often not required for participation but potentially supporting dissemination) or cultural proximity

Table 1 Multiple linear regression model predicting film appearance counts of countries ($N = 163$) by demographic, economic and geographic variables

	β	SE	p
Intercept	2.05	0.08	<0.001
Population	0.72	0.06	<0.001
GDP per capita	0.84	0.08	<0.001
Events	0.29	0.12	0.02
Distance from France	−0.05	0.01	<0.001
Population: GDP per capita	0.19	0.07	0.01

Response and predictor variables, except for distance, on the $\log_{10}$ scale. $F(5, 157) = 105$, adjusted $R^2 = 0.763$

(Straubhaar, 1991; Fu & Sim, 2010). The model was checked for and met the assumptions of linearity, equal variance, and homoscedasticity, and exhibited only mild multicollinearity. Table 1 shows the coefficients of the model.

All predictors are significant at the $\alpha = 0.05$ level, and the model as a whole is highly predictive of the outcome, describing 76% of the variance in country-wise festival representation. All β coefficients except distance are positive, indicating that, as already illustrated in the graphs above, larger and wealthier countries have on average more films appearing in the circuit, and also that hosting festivals indeed increases representation. The intercept of $10^{2.05} = 112.2$ corresponds to the expected n for (or a hypothetical mid-sized country at) the reference values of the variables, i.e., 10M population, $10k per capita, hosting zero events, located where France is. Since this is a log–log model, the coefficients are directly interpretable as percent changes (elasticities). For example, a 10% change in population (with everything else held constant) leads to a 7.2% increase in predicted n, e.g., for a country of 10M, 10% or 1M more people predicts $10^{\mathrm{n}} - 10^{2.05} \approx 8$ more appearances (where n is the predicted: $2.05 + 0.72 \times c$, where c is the 10M-centered target population $\log_{10}(11\mathrm{M}) - \log_{10}(10\mathrm{M})$). However, their interaction is also significant, indicating that their effect is multiplicative: having both a large population and a high GDP per capita boosts festival appearances more than being just big or just wealthy.

Distance from France is significant and negative: for each additional 1000 kms, n decreases by a multiplicative factor of $10^{-0.05} = 0.89$, or in percent by $(1 - 0.89) \times 100 = 11\%$. Notably, this factor is significant even in this model that controls for GDP per capita (so it is not just an effect of many less prosperous countries, e.g., the Global South being far from Europe). A simpler model just predicting n by distance describes about 7% in variance. The fact that even this crude proxy to geographical effects holds predictive power indicates that where a country is located may significantly affect its chances of festival representation.

In summary, most of the variation in how many films get selected from each country comes down to their size, wealth, where they are relative to Europe, and how many festivals they host themselves. While this is somewhat obvious, a statistical model like this allows for a more precise quantification of that intuition.

Coming back to Fig. 3c: the residuals of this model indicate which countries are near their predicted values (those around zero), which do better than average, and which do worse. The zero is effectively the regression line and can be thought of as the global average prediction. The residuals are on the same $\log_{10}$ scale: a value of 1 means the country appears 10× more in the circuit than would be expected by the aforementioned factors. For example, Argentina, despite its moderate average GDP per capita of $12k and about 44M population in the period, has 2041 film festival pairs listed in the database, while the global model prediction would place it at 290. It is shown below how this may be boosted by a local festival circuit (see also Isaza, 2012). The over 10 times smaller Uruguay still has 270 appearances of 142 films on record (prediction would be just 23). France in comparison performs slightly lower than predicted (has 6759, prediction is 6541), despite hosting a large array of festivals. The USA and China have considerably lower numbers than would be expected by their size and prosperity. As discussed above, this may be an effect of local markets and festivals or specifics of the dataset. The UK is shown to under-perform despite its prosperity and arguably favorable position in European audiovisual industries, especially in the pre-Brexit era that is part of the observation period here (Donders & Van den Bulck, 2016; Steemers, 2016; Ibrus et al., 2023). Afghanistan is among countries with a rather high positive residual: despite its very low GDP and other factors, it is listed on 47 appearances of 26 films in total in 2012–2021, although 23 of them appear to be co-productions with mostly European countries. Many small European countries also beat the odds, including Iceland and Luxembourg, but also Estonia, Latvia and Lithuania (all below 3 million in size), potentially reflecting recent developments of their film and television industries (cf. Hjort and Petrie, 2007; Iosifidis, 2007; Ibrus et al., 2019; Ibrus et al., 2023).

3.3 Cultural flows of festival films between countries of the world

These results raise the question (RQ3) of how films may travel between countries and festivals, how much apparent success comes down to (potentially favorable) local festivals, and how much countries may differ in terms of cultural "trade balance" (Disdier et al., 2010; Baorui et al., 2024)—in this case, how much a country's films get exported to foreign festivals, versus locally hosted festivals "importing" foreign films. These questions have been explored in the past in terms of television or cinema, but not festivals, and usually on a country rather than world scale Meloni et al. (2018); McKenzie and Smirnov (2018); Crane (2018); Coate et al. (2017); Avezzù (2022); Ibrus et al. (2023); Weber et al. (2024); Hołobut et al. (2024). While all these topics deserve deeper study, this section attempts to shed some light on these global dynamics and cultural flows (see also Disdier et al., 2010; Fu and Sim, 2010; Christensen, 2013; Kawashima and Lee, 2018; Hartley et al., 2020). This analysis is of course still limited to festivals and does not consider other venues and platforms.

This requires once again choosing between potential aggregation approaches. In the analyses above, every film festival entry of a film produced or co-produced by a country was counted equally for simplicity. This is fitting if one is interested in

whether a country is represented in a festival or the circuit as such. In this section, a slightly more nuanced approach is taken, and the counts and shares are weighted by the number of co-producers, for a hopefully more accurate picture of cultural flows. The unit is still one film entry at a festival. For example, a film produced by country *A* and shown in a festival in country *X* adds a count of 1 in the direction from *A* to *X*; while a film co-produced by countries *A* and *B* adds 0.5 for each instead.

Figure 4a illustrates the flow of films between the largest production and festival host countries. Not all those listed are large producers, for example, Estonia once again appears because of its Tallinn Black Nights film festival, which screens a fairly large number of films yearly. Normalized by rows (producers), the diagonal of the matrix effectively displays the share of a given country's productions that are screened locally. For some countries this is quite high: 61% for Greece (at two festivals, the Thessaloniki Documentary Festival and the Thessaloniki IFF). For Argentina and France, this is 42 and 41%. For others like China and Japan (and on average other countries) domestic festival selections of films are at or below 10%. Some countries export a considerable share of their films to a single destination. A fourth of Belgian film entries are at French festivals—not surprising given the geographic proximity and one shared national language). Some relationships are reciprocal: the UK screens almost a tenth of their films in the USA and vice versa, although the primary non-domestic destination for US films is France (18%). The latter is also an important destination for the UK at 14% (or 311 entries), but interestingly not so much the other way around, with only 3% or 143 French film listings in UK festivals (see also Mazdon and Wheatley, 2010). A larger version of this graph may be found in the Supplementary.

Not all countries in the database are recorded as hosting any festivals, which means their incoming number of films here is zero. Figure 4b shows a

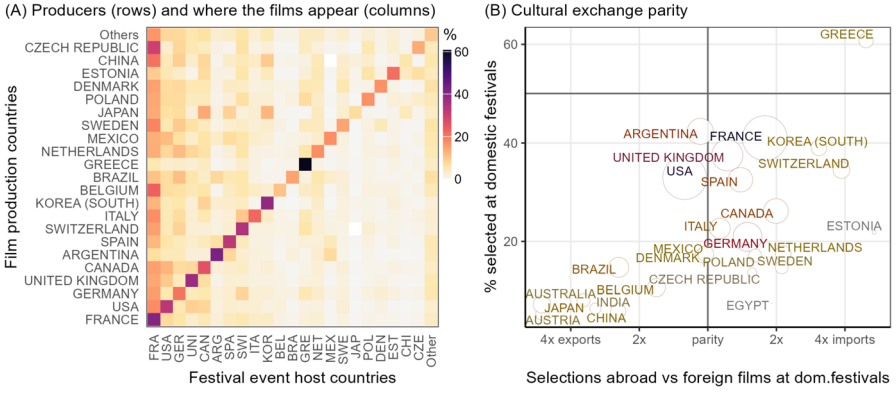

Fig. 4 Flow of films between production countries and festival-hosting countries. Panel **a** depicts the export share of festival films normalized by production countries (rows). Panel **b** displays the export–import balance on the horizontal axis on a $\log_2$ or multiplicative scale (not taking into account domestic screenings of domestic films), against the share of domestic film selections at domestic festivals on the vertical axis, for all countries with at least 5 hosted festival events in the database. The largest producer, France, also imports more than it exports: more foreign films occur in its festivals than its own films are selected at festivals abroad

sample of countries that have hosted (at least 5) festival events, arranged across an export–import parity axis. This is calculated simply as a multiplicative difference or $\log_2$(imports/exports) (where a value of 1 means 2× more imports than exports, 2 is 4× more, etc.). Here, imports are defined for each country as the sum of (weighted) film festival pairs where the producer is not the given country, and exports as the sum of instances where a given country is a producer but not the host. The vertical axis shows how many entries of a country's films occur at their own festivals (effectively the diagonal of the left side panel). Smaller production countries like Estonia and Sweden import more than they export. Despite its large production volume, 1.8× more foreign films are listed at French festivals, compared to how many French films are selected at festivals abroad. A reason may be its relatively large internal festival circuit (76 events in 19 series in 2012–2021). Along with other large production countries like the UK, USA, and Argentina, a fair share of French films are screened within the country. The USA is a net exporter not only of cinema but also festival films (cf. Meloni et al., 2018; Tunstall, 2008). The largest countries China and India are in the bottom left, exporting 2.5 and 3.1× more, and screening their productions largely in festivals abroad. This is however according to the somewhat Euro-centric Cinando database, which for example lists only the Shanghai and Beijing IFF as Chinese festivals. It is entirely possible that while some films from these countries occur in Western festivals and get listed in the database, the share of local festivals may simply not be represented here (cf. Ma, 2014; Dastidar and Elliott, 2020).

Figure 5 further illustrates the flow of festival films for a selection of countries, ranging again from large producers like France and the USA to small players like Estonia and Portugal, which also rely more on international co-productions. The data are aggregated as above. These networks only show the inflow and outflow for each target country at a time. With the arrow size indicating relative counts, one thing that is immediately striking is the difference in the sizes of the local circuits, visible as the looping edge. As also shown in Fig. 4a, countries also differ in their cultural trade partner distributions, with some sending and receiving films from a variety of nations while others are somewhat more focused on certain destinations like the UK to France. Some links here are also due to co-productions, e.g., a film jointly produced by Portugal and Spain, and screened at the San Sebastian Film Festival in the latter, would still count (at a 0.5 weight) toward a Portugal-to-Spain link.

It should also be noted that this snapshot of the Cinando database ends in 2021, therefore not reflecting the state of the festival industry since then, including the banning of Russia from major film festivals following its military invasion into Ukraine in 2022 (Vivarelli, 2023), or the defunding of the Argentinian national film body in 2024 (Geisinger, 2024).

3.4 How would platforming smaller (or larger) countries more affect diversity of the festival circuit

As illustrated in the beginning (Fig. 1), a handful of the largest production countries make up the lion's share of the film entries across festivals (although not as much as

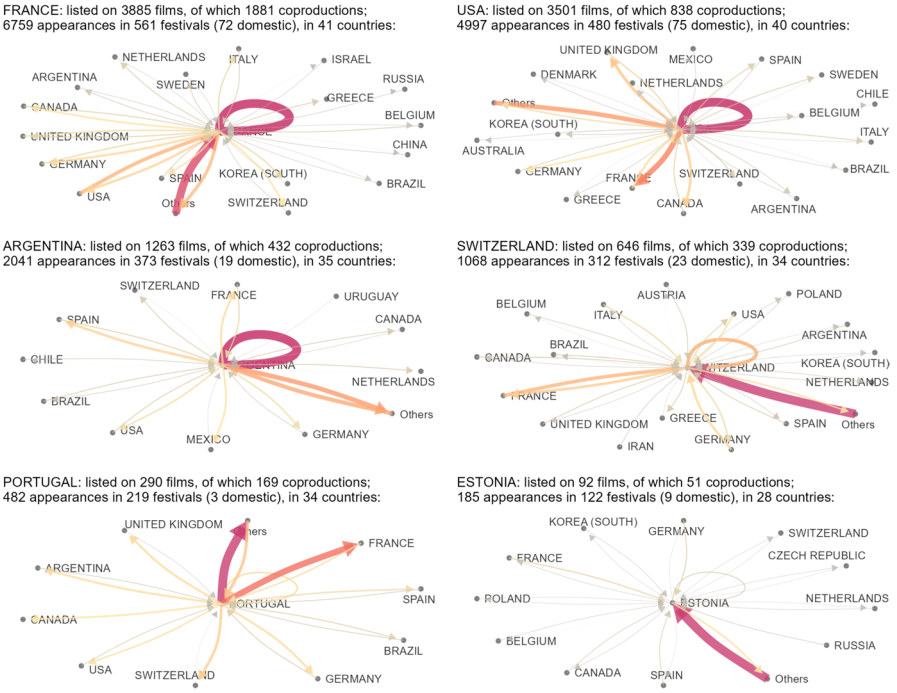

Fig. 5 Film exchange flows of selected countries. These "star networks" only show links between a given central country and countries where its films are showcased at festivals (outgoing arrows) or whose films are shown in its festivals. The line color and size are relative, with the thickest and darkest being the largest link in terms of either in- or outgoing film appearances. To maintain legibility, only those countries accounting for about a fifth of the imports–exports are shown for each country, the rest grouped as "Others." Some countries like France have strong internal festival circuits, visible as the loops in the network, while others export either in a focused or spread-out manner

entirely proportional representation would lead to; Fig. 2). This can be argued to be detrimental to diversity in festival programming. Festivals vary significantly along multiple dimensions, including thematic, linguistic, and geographic diversity, gender balance, and in terms of local, regional or international programming (Zemaityte et al., 2024). There are thematically concentrated festivals like Frightfest London, Fantasia, and DocCorner; festivals that screen films predominantly in one language like Spanish in Latin America (Guadalajara, Ventana Sur) or Sundance in English, yet others where no single language takes up much more than 10% (e.g., Tallinn Black Nights in 2019). And when it comes to gender, some festivals like Göteborg or Hot Docs often have an almost perfect balance, while others have been be shown to be heavily male-dominated in some years (see Zemaityte et al., 2024; Verhoeven et al., 2019; Ehrich et al., 2022).

The question been raised how much diversity (and potential public value) is added to cultural and audiovisual spheres, including the film festival circuit, by including products of smaller or peripheral countries, and conversely, how much is lost by excluding them (Sand, 2019; Zemaityte et al., 2024; Ibrus et al., 2023; Félix-Jäger,

2020; Zemaityte et al., to appear). While it would be difficult to quantify the exact contribution of individual countries here, it is possible to simulate what the circuit (or an individual festival) would look like in terms of diversity, if films from certain countries were more (or less) frequently platformed, and extrapolate the value of balancing representation in one direction or another from that. Quantifying a single general cultural diversity metric directly would also be difficult (and out of scope here), but we can use present film metadata such as languages spoken in the films and film crew gender (here directors and producers) as approximate proxies to cultural differences and diversity. As discussed in Methods, the simulation constructed to explore RQ4 works by generating distributions of films from the Cinando dataset, representing hypothetical festival circuits, with various distributions of production country population and relative GDP averages, and then measuring the sample diversity (averaged over 100 replications). The simulation is necessarily simplified, treating the circuit as a global pool of films, and without modeling individual festival programming. This global view may still be informative, as the circuit does form a highly inter-connected network (Loist, 2016; Zemaityte et al., 2024).

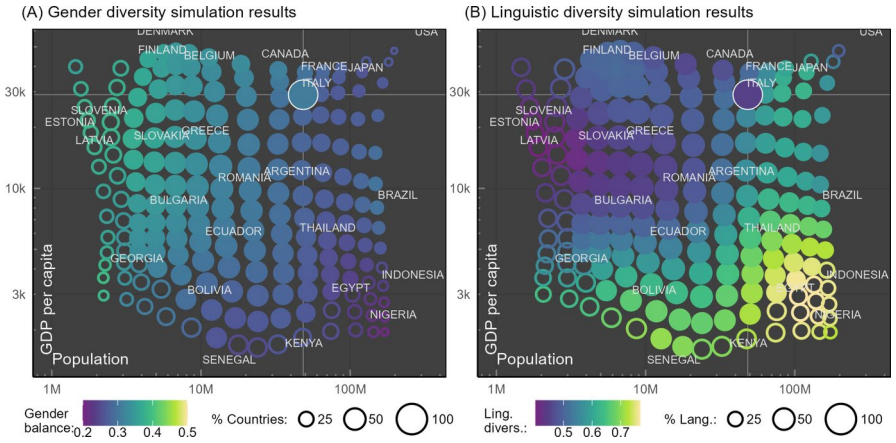

Fig. 6 Results of diversity simulations, showing how gender (**a**) and linguistic diversity (**b**) change depending on how the representation of different countries. Each circle is a sample of films or one simulated festival circuit. Its coordinates on the ↔population and ↕GDP per capita axes indicate the (geometric) means of these production country variables in that hypothetical circuit. For example, distributions in the top left consist mostly of small but relatively wealthier countries (countries similar to Finland, Slovenia). The bottom right has large but least affluent countries such as Nigeria or Indonesia, while the USA is uniquely wealthy and large in the top right corner. Circle color is diversity, from dark purple low to yellow high. For gender (**a**), this is just the fraction of women directors and producers; 0.5 would be a perfectly gender balanced circuit, but this is not achieved in any present country distribution. Circle size in A indicates the percentage of countries in the sample (of the 161 total). Panel **b** shows typological diversity in languages spoken in films; here size additionally indicates the percentage of languages, compared to the dataset total of 114 (simulation sample size in terms of countries is consistent with A, as the simulations are quite stable). These are all below 100% due to the (intentionally) small sample size and standard deviation, as discussed in Methods. Empty circles indicate samples below 60% of intended sample size (boundary effects). The white highlighted circle in the top right quadrant shows the value for the full dataset of the current festival circuit (hence 100% size), which is skewed toward larger and wealthier countries, and relatively low in diversity on both metrics

The result depicted in Fig. 6 illustrates 196 possible film samples (the colored circles) with a given population and GDP per capita geometric mean in terms of production countries, and a standard deviation of 0.3 (on the $\log_{10}$ scale; see the SI for exploration of the latter parameter). The graph background is dark to differentiate it from the real-world data-driven graphs above. The simulated samples and their diversity values can be interpreted in two ways, either as potential festival circuits or as individual festival programs: what a circuit (or a festival) would look like if films from predominantly these or those kinds of countries would be platformed. The axes and therefore coordinates of example country labels are the same as in Fig. 3a.

For gender balance (Fig. 6a), the results could be seen as either cautiously optimistic from one, yet dire from another perspective. The real dataset mean (white circle) is at 31% women, which is still far from 50–50 representation (the yellow which does not exist on panel a). Yet the results show that even less gender diverse circuits (darker purple circles) could potentially exist, if even more films from larger countries are platformed, especially from large but less affluent countries (bottom right quadrant). Slightly more balanced circuits could be created by including more films from smaller countries (left edge of the panel a). However, no matter where the bivariate distributional peak is shifted, gender balance is not achieved, if sampling from the set of current festival films. In other words, festival circuit gender balance can more likely be achieved by creating more opportunities for women filmmakers, rather than by adjusting country-wise programming, at least along these two socio-economic axes.

The linguistic diversity simulation (Fig. 6b) provides some interesting insights. The current circuit, as represented by the Cinando dataset, is very low in diversity, compared to other potential combinations, as it consists predominantly of films from USA, France, and other European countries with closely related languages (recall Fig. 1). Only shifting the distributions to small, moderately wealthy countries would make this lower yet (dark purple circles, middle left), as these are mostly small countries speaking related Indo-European languages, e.g., Slovak and Slovenian (recall Methods: this is based on a relative typological similarity metric, not mere counts of languages). Diversity is also slightly higher than the current value in samples toward the top right corner where Japan and the USA are, although the absolute number of languages (circle size) there is smaller. This is not unexpected, as there are numerous international US-led co-productions in the dataset, covering a variety of 37 languages of the 114 present; and Japanese being an language isolate increases typological diversity on its own.

Unlike gender balance, linguistic diversity is highest in samples with the distributional peak around less affluent countries, especially the larger, in the bottom right quadrant. This makes sense, as most of them are African and Asian countries speaking numerous indigenous languages from several unrelated language families which also make it to the films. What these results indicate is that linguistic (and by proxy, possibly also cultural) diversity of the international festival circuit could in theory be increased by shifting the currently dominant Western focus also toward less affluent countries around the globe. Naturally, such a simplified global aggregation neither is nor should be the only perspective to diversity, and festivals often serve many

(if not more) local cultural, artistic economic, and other functions discussed in the Introduction.

4 Discussion

This contribution has attempted to provide a global-scale quantification of national representation in the film festival circuit from an (admittedly simplified) demographic and economic perspective and to explore the dynamics and cultural flows shaping the festival network, which in turn impacts the economies and cultural landscapes of the production and event host countries. Rather than conclusive outcomes, these results and methodological proposals are meant more as a stepping stone toward deeper and more contextualized research of the festival ecosystem and cultural value generation, building on and complementing recent investigations of the festival circuit via the Cinando database (Zemaityte et al., 2024). The findings resonate with broader theoretical accounts of cultural policy and public value. As discussed above, festivals operate as arenas where cultural, symbolic, and economic value is produced and contested (Benington & Moore, 2011; Mazzucato et al., 2020; O'Regan & Potter, 2022). The observed imbalances in national representation reflect structural constraints that limit the democratizing potential often ascribed to the global festival circuit (Valck & Marijke, 2007; Loist, 2016). This study contributes to ongoing debates about equity, diversity, and soft power in global cultural flows (Hartley et al., 2020; Guan et al., 2023) by illustrating how wealth, geography, and linguistic reach shape cultural circulation.

Several simplifications were necessary due to the nature of the data, and the subsequent limitations should be kept in mind when interpreting the results. These aspects can also hopefully be improved upon in future research. For example, the diversity simulation in the last Results section is a coarse first analysis, as cultural diversity, public value and "social profitability" (Peralta García & Simour, 2025) have many more dimensions. The global averaging also simplifies and overlooks the value that regional festivals can provide by platforming local, regional or diasporic films, as well as the (both symbolic and economic) value that emerging industries and filmmakers can gain from being showcased at prestigious international film festivals (Hadida, 2009). It is hoped however that this methodological proposal paves the way for a more systematic and data-driven approach to cultural diversity and relatedly, public value, a domain which has until recently remained rather theoretical in the context of film festival and audiovisual industries research (but see Faulkner and Kaufman; 2018; Ibrus et al., 2023; Zemaityte et al., 2024).

The simulation results point to a tricky optimization problem, reflecting other cultural event and festival programming research (Yan et al., 2012; Hagan, 2018; Allemana, 2020). It is clearly not trivial to balance a festival program or the circuit as a whole, as different aspects of diversity require platforming films from different kinds of countries. Yet there is only so much a given festival (and by extension, an annual circuit) can accommodate, time being a limited resource. Consequently, increasing representation in one category necessarily means some other category will receive less exposure. As shown here with gender and

languages, balancing or prioritizing one aspect can easily create inequality in another. Additionally, diversity in country profiles, having the small and the large represented, may be a target in itself, but, again, may not always align with other aspects. Quantitative analyses and simulations like this enable testing scenarios with multiple moving parts and therefore support data-driven decision-making. This includes applications in festival curation (e.g., navigating trade-offs between diversity goals) and in film funding or subsidy policy (Rimscha & von, 2022; Freudendal, 2024), where targeted support for underrepresented groups, countries or genres could improve equity in global circulation.

An aspect touched upon here but deserving of further study concerns the patterns and dynamics of co-productions (Bondebjerg, 2016; Parc, 2020). Many of the country examples brought here exhibit high rates of international production, with some co-producing half or more of their festival films (see Fig. 5). The production patterns are also likely predictive of cultural flows, as films produced jointly may well subsequently enjoy reception at festivals in all involved co-producer countries.

Beyond the bias analysis, there was limited attention here on the categorization of festivals as the accredited A-list and the rest, which of course comes into play in representation and success, given the known variance in competitiveness and functions of these events (Zemaityte et al., 2024). A related major limitation imposed by the nature of this particular dataset is the lack of information on rejection. While the Cinando database provided details on festival programming, there was no way to tell how many films a given country actually sends to a given festival, nor the percentage that gets selected, and whether there is a bias toward or against some countries. Obtaining this information would greatly support the understanding of both the dynamics of individual festivals and the circuit as a whole. It would also enable a more informative modeling of representation and its various potential predictors, which was done here in a coarse manner, simply counting film festival pairs per country. Future research could also incorporate more precise information on film industry revenues instead of country GDP, or factors of cultural proximity between producer and host countries (Fu & Sim, 2010).

The latent embedding approach proposed in Zemaityte et al. (2024) is suitable for computing both diversity and similarity, in a manner that goes beyond discrete counts of cultural markers or categories and allows for taking into account their intra-similarity. The diversity analysis above considers gender and film language, but further cultural factors, some operationalization of content, narrative, and visual similarity, as well as the perceived quality and novelty of a film could be used to compare festivals, measure additional aspects of diversity, and predict success at festivals. Interpersonal, professional network, and personal prestige-based factors and biases are likely factors as well (De Valck, 2016a; Coate et al., 2017; Mair & Duffy, 2020; Ehrich et al., 2022).

Despite limitation, it is nevertheless hoped that these analyses and methodological discussions support further quantitative festival research. The preliminary insights could be useful for researchers of the topic but perhaps also for festival organizers and policymakers, providing a data-driven perspective and complementary ways of thinking about analyzing cultural festival programming issues, their socioeconomic aspects, as well as diversity and balance.

5 Conclusions

Film festivals boost economies, provide a platform for filmmakers to showcase their work, enhance cultural spheres and promote international interactions, provide a spotlight on underrepresented voices and narratives through selective programming, and directly or indirectly create public value. At the same time, the global festival circuit, as represented in the Cinando dataset, is dominated by a few successful production countries. This contribution has endeavored to measure this apparent bias, discuss quantifiable models of representation equality and diversity, model festival performance, and explore the global cultural flows of festival films. The observed festival circuit is saturated by films from large affluent Western nations, and representation remains highly correlated to socioeconomic and geographic factors. Yet some smaller countries appear to be beating these odds, and their inclusion can be argued to benefit aspects of cultural diversity.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s10824-025-09557-x.

Author Contributions AK designed the research, prepared the data, designed and performed the data analysis, wrote the text, and created the figures. VZ suggested the initial research direction of sections 3.1—3.2, including the WB and FIAPF data, and provided comments on the final draft.

Funding Andres Karjus was supported by the CUDAN ERA Chair project for Cultural Data Analytics, funded through the European Union Horizon 2020 research and innovation program (Project No. 810961). Vejune Zemaityte was supported by the Public Value of Open Cultural Data project, funded through the Estonian Research Council (Project No. PRG 1641), and by project CresCine—Increasing the International Competitiveness of the Film Industry in Small European Markets, funded through the European Union Horizon Europe funding scheme (Project No. 101094988). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Data Availability All data used in this study are openly accessible. The Cinando database snapshot is available via https://doi.org/10.6084/m9.figshare.22682794.v1 (Cinando technical ID values are anonymized). The data are also available to explore via an interactive dashboard that accompanied the first Cinando-related publication: https://andreskarjus.github.io/cinandofestivals The FIAPF accreditation list is available via https://fiapf.org/festivals/accredited-festivals. The population and GDP data were retrieved from the World Bank database available via https://data.worldbank.org/indicator. The UIS film production statistics are available via https://data.uis.unesco.org.

Declarations

Conflict of interest The authors have no relevant financial or non-financial interests to disclose. The research did not involve any human subjects or other topics which would require ethics approval or consent.

References

Allemana, M. (2020). The Gatekeepers' puzzle: Programming diversity and inclusion in a Jazz festival. *Jazz Research Journal, 14*(2), 107–137. https://doi.org/10.1558/jazz.42412

Avezzù, G. (2022). The market for foreign cinema in contemporary Italy: A geography of film consumption. *GeoJournal, 87*(1), 73–84. https://doi.org/10.1007/s10708-022-10639-y

Baorui, S., Zhiyuan, Z., & Zhao, L. (2024). Currencies and culture: An amusing journey into the impacts of exchange rates on global creative industries. *The American Journal of Economics and Sociology, 83*(3), 647–672. https://doi.org/10.1111/ajes.12567

Benington, J., & Moore, M. (Eds.). (2011). *Public value: Theory and practice*. Basingstoke (GB): Bloomsbury Publishing.

Bérubé, N., Ghiasi, G., Sainte-Marie, M., & Larivière, V. (2020). *Wiki-Gendersort: Automatic gender detection using first names in Wikipedia*. SocArXiv. https://doi.org/10.31235/osf.io/ezw7p.

Bondebjerg, I. (2016). Transnational Europe: TV-drama, co-production networks and mediated cultural encounters. In *Palgrave communications* 2, Article 16034. https://doi.org/10.1057/palcomms.2016.34.

Bosma, P. (2015). *Film programming: Curating for cinemas, festivals, archives*. Chichester (NY): Columbia University Press. https://doi.org/10.7312/bosm17459

Burgess, D. (2020). Capturing film festival Buzz: The methodological the methodological dilemma of measuring symbolic value. NECSUS. https://doi.org/10.25969/mediarep/15318

Christensen, M. (2013). TransNational media flows: Some key questions and debates. *International Journal of Communication, 7*, 2400–2418.

Chung, J. E. (2011). Mapping international film trade: Network analysis of international film trade between 1996 and 2004. *Journal of Communication, 61*(4), 618–640. https://doi.org/10.1111/j.1460-2466.2011.01567.x

Coate, B., Verhoeven, D., Arrowsmith, C., & Zemaityte, V. (2017). Feature film diversity on Australian Cinema Screens: Implications for cultural diversity studies using big data. In M. David (Ed.), *Australian screen in the 2000s* (pp. 341–360). Cham: Palgrave Macmillan, Ryan and Ben Goldsmith.

Crane, D. (2018). Cultural flows and the global film industry: A comparison of Asia and Europe as Regional Cultures. In N. Kawashima & H.-K. Lee (Eds.), *Asian cultural flows: Cultural policies, creative industries, and media consumers* (pp. 113–126). Singapore: Springer. https://doi.org/10.1007/978-981-10-0147-5_7

Dastidar, S. G., & Elliott, C. (2020). The Indian film industry in a changing international market. *Journal of Cultural Economics, 44*(1), 97–116. https://doi.org/10.1007/s10824-019-09351-6

Diestro-Dópido, M. (2021). *Film festivals: Cinema and cultural exchange*. NED - New edition, vol. 18. Modern Humanities Research Association.

Disdier, A.-C., Tai, S. H. T., Fontagné, L., & Mayer, T. (2010). Bilateral trade of cultural goods. *Review of World Economics, 145*(4), 575–595. https://doi.org/10.1007/s10290-009-0030-5

Donders, K., & Bulck, H. (2016). Decline and fall of public service media values in the international content acquisition market: An analysis of small public broadcasters acquiring BBC worldwide content. *European Journal of Communication, 31*, 299–316.

Ehrich, M. E., Burgdorf, K., Samoilova, Z., & Loist, S. (2022). The film festival sector and its networked structures of gender inequality. *Applied Network Science, 7*(1), 1–38. https://doi.org/10.1007/s41109-022-00457-z

Elsaesser, T. (2005). *European Cinema: Face to face with Hollywood*. Amsterdam (NL): Amsterdam University Press.

Falicov, T. L. (2016). The "Festival film": Film festival funds as cultural intermediaries. In *Film festivals: History, theory, method, practice*. Milton Park: Routledge.

Farrell, B. (2016). Understanding the value of arts & culture: The AHRC cultural value project (2016) by Geoffrey Crossick and Patrycja Kaszynska. *Cultural Trends, 25*(4), 273–276. https://doi.org/10.1080/09548963.2016.1241382

Faulkner, N., & Kaufman, S. (2018). Avoiding theoretical stagnation: A systematic review and framework for measuring public value. *Australian Journal of Public Administration, 77*(1), 69–86. https://doi.org/10.1111/1467-8500.12251

Félix-Jäger, S. (2020). The effects of globalization on art and aesthetics. In S. Félix-Jäger (Ed.), *Art theory for a global pluralistic age: The glocal artist* (pp. 39–63). Cham: Springer. https://doi.org/10.1007/978-3-030-29706-0_3

FIAPF. (2023). Accredited Festivals.
Freudendal, J. (2024). Towards audience-centric film funding: New strategies for commissioning youth films in Denmark. *Journal of Scandinavian Cinema, 14*(1), 31–39. https://doi.org/10.1386/jsca_00108_1
Fu, W. W., & Sim, C. (2010). Examining international country-to-country flow of theatrical films. *Journal of Communication, 60*(1), 120–143. https://doi.org/10.1111/j.1460-2466.2009.01455.x
Geisinger, G. (2024). *Argentinian government suspends state funding to national film body INCAA*. Screen Daily. https://www.screendaily.com/news/argentinian-government-suspends-state-funding-to-national-film-body-incaa/5191550.article Visited on 07/09/2024.
Grunwell, S., & Ha, I. (2008). Film festivals: An empirical study of factors for success. *Event Management, 11*(4), 201–210. https://doi.org/10.3727/152599508785899893
Guan, M., Chagas-Bastos, F. H., & Nishijima, M. (2023). Winning hearts and minds: soft power, cinema, and public perceptions of the united states and china in brazil. *Global Studies Quarterly*. https://doi.org/10.1093/isagsq/ksad029
Hadida, A. L. (2009). Commercial success and artistic recognition of motion picture projects. *Journal of Cultural Economics, 34*(1), 45–80. https://doi.org/10.1007/s10824-009-9109-z
Hagan, C. (2018). Curatorial practices for intersectional programming. In *The International Journal of Screendance*. https://doi.org/10.18061/ijsd.v9i0.6046.
Harbord, J. (2002). *Film cultures*. London (GB): SAGE.
Hartley, J., Ibrus, I., & Ojamaa, M. (2020). *On the Digital Semiosphere: Culture, Media and Science for the Anthropocene*. New York (NY): Bloomsbury Publishing.
Hjort, M., & Petrie, D. (Eds.). (2007). *The cinema of small nations*. Edinburgh: Edinburgh University Press.
Hołobut, A., Rybicki, J., & Stelmach, M. (2024). A statistical approach to Hollywood remake and sequel metadata. *Digital Scholarship in the Humanities, 39*(2), 556–574. https://doi.org/10.1093/llc/fqae012
Ibrus, I., Karjus, A., Zemaityte, V., Rohn, U., & Schich, M. (2023). Quantifying public value creation by public service media using big programming data. *International Journal of Communication, 17*, 24.
Ibrus, I., & Ojamaa, M. (2014). What is the cultural function and value of European transmedia independents? *International Journal of Communication, 8*, 18.
Ibrus, I., Rohn, U., & Nanì, A. (2019). Searching for public value in innovation coordination: How the Eurovision song contest was used to innovate the public service media model in Estonia. *International Journal of Cultural Studies, 22*(3), 367–382. https://doi.org/10.1177/1367877918757513
Iosifidis, P. (2007). Public television in small European countries: Challenges and strategies. *International Journal of Media & Cultural Politics, 3*(1), 65–87. https://doi.org/10.1386/macp.3.1.65_1
Isaza, L. R. (2012). *Branding Latin America: Film festivals and the international circulation of Latin American Films*. Leeds: University of Leeds.
Kawashima, N., & Lee, H.-K. (2018). *Asian cultural flows*. Berlin: Springer.
Kedar, O., Harsgor, L., & Sheinerman, R. A. (2016). Are voters equal under proportional representation? *American Journal of Political Science, 60*(3), 676–691. https://doi.org/10.1111/ajps.12225
Kendall, G., Chan, J. H., Yeung, M. C., & Law, K. K. (2021). Do film festivals attract tourists? *Current Issues in Tourism, 24*(11), 1482–1486.
Kostopoulou, S., Vagionis, N., & Kourkouridis, D. (2013). Cultural festivals and regional economic development: Perceptions of key interest groups. In Á. Matias, P. Nijkamp, & M. Sarmento (Eds.), *Quantitative Methods in Tourism Economics* (pp. 175–194). Heidelberg: Physica-Verlag HD. https://doi.org/10.1007/978-3-7908-2879-5_10
Krainhöfer, T. C. (2018). *Mapping of collaboration models among film festivals: A qualitative analysis to identify and assess collaboration models in the context of the multiple functions and objectives of film festivals*. Brussels: European Commission.
Kurz, S., Maaser, N., & Napel, S. (2017). On the democratic weights of nations. *Journal of Political Economy, 125*(5), 1599–1634. https://doi.org/10.1086/693039
Locarno Film Festival. (2022). *Annual Report 2022*.
Loist, S. (2016). The film festival circuit: Networks, hierarchies, and circulation. In M. de Valck, B. Kredell, & S. Loist (Eds.), *Film festivals: History, theory, method, practice* (pp. 49–64). London (GB): Routledge.
Lorenzen, M. (2007). Internationalization vs. globalization of the film industry. *Industry and Innovation, 14*(4), 349–357. https://doi.org/10.1080/13662710701543650

Ma, R. (2014). Regarding the grassroots Chinese independent film festivals: Modes of multiplicity and abnormal film networking. In *China's iGeneration: Cinema and moving image culture for the twenty-first century*, pp. 235–253.

MacPherson, R. (2010). Is bigger better? Film success in small countries—the case of Scotland, Ireland and Denmark. In *16th International conference of the association for cultural economics international*.

Mair, J., & Duffy, M. (2020). The role of festival networks in regional community building. In A. Campbell, M. Duffy, & B. Edmondson (Eds.), *Located research: Regional places, transitions and challenges* (pp. 89–116). Singapore: Springer. https://doi.org/10.1007/978-981-32-9694-7_7

Malaviya, C., Neubig, G., & Littell, P. (2017). *Learning language representations for typology prediction*. arXiv:1707.09569. https://doi.org/10.48550/arXiv.1707.09569.

Martin-Jones, D., & Montañez, M. S. (2013). Uruguay disappears: Small cinemas, Control Z Films, and the aesthetics and politics of auto-erasure. In: *Cinema Journal* (Vol. 53(1), pp. 26–51).

Mazdon, L., & Wheatley, C. (2010). *Je T'Aime... Moi Non Plus: Franco-British Cinematic Relations*. New York: Berghahn Books.

Mazzucato, M., Conway, R., Mazzoli, E. M., Knoll, E., & Albala, S. (2020). *Creating and measuring dynamic public value at the BBC*. Policy Report IIPP 2020-19. UCL Institute for Innovation and Public Purpose, London.

McKenzie, J., & Smirnov, V. (2018). Blockbusters and market expansion: Evidence from the motion picture industry. *Journal of Cultural Economics, 42*(2), 341–352. https://doi.org/10.1007/s10824-017-9308-y

McQuail, D., & Cuilenburg, J. J. (1983). Diversity as a media policy goal: A strategy for evaluative research and a Netherlands Case Study. *International Communication Gazette, 31*(3), 145–162. https://doi.org/10.1177/001654928303100301

Meloni, G., Paolini, D., & Tena, J. D. (2018). American beauty: Trade flows and export costs of US movies. *Journal of Cultural Economics, 42*(4), 701–716. https://doi.org/10.1007/s10824-018-9324-6

Monson, Q., Wilson, S., & Goodliffe, J. (2023). *Economic Impact: 2023 Sundance Film Festival*.

Moreau, F., & Peltier, S. (2004). Cultural diversity in the movie industry: A cross-national study. *Journal of Media Economics, 17*(2), 123–143.

Navarro, C., García-Muñoz, N., & Delgado, M. (2022). Local fiction series: The value of European public service media (1990–2020). In *Media, Culture & Society*. https://doi.org/10.1177/01634437211069972..

Odabasi, E. (2025). Value creation and the international film festival circuit: The Cannes 2020 label. *Transnational Screens, 16*(1), 50–66. https://doi.org/10.1080/25785273.2025.2476239

O'Regan, T., & Potter, A. (2022). The past and future of public value: The end of an illustrious career or its reinvention? In M. Hjort & T. Nannicelli (Eds.), *A companion to motion pictures and public value* (pp. 427–444). Hoboken (NJ): Wiley.

Parc, J. (2020). Understanding film co-production in the era of globalization: A value chain approach. *Global Policy, 11*(4), 458–465. https://doi.org/10.1111/1758-5899.12830

Peirano, M. P. (2020). Mapping histories and archiving ephemeral landscapes: Strategies and challenges for researching small film festivals. *Studies in European Cinema, 17*(2), 170–184. https://doi.org/10.1080/17411548.2020.1765630

Peralta García, L., & Simour, L. (2025). The conceptual debate over migration film festivals. In L. P. García & L. Simour (Eds.), *Migration Film Festivals: Social Functions, Expectations and Challenges* (pp. 79–107). Cham: Springer. https://doi.org/10.1007/978-3-031-63693-6_4

Rimscha, M. B. (2022). How subsidies promote a uniform film supply. *International Journal of Media & Cultural Politics, 18*(1), 3–20. https://doi.org/10.1386/macp_00055_1

Roy, C. (2014). Telling stories of resistance and change: Organizers of film festivals contribute to media literacy. *Canadian Journal for the Study of Adult Education, 26*(3), 1–16.

Rüling, C.-C., & Pedersen, J. (2010). Film festival research from an organizational studies perspective. *Scandinavian Journal of Management, 26*, 318–323. https://doi.org/10.1016/j.scaman.2010.06.006

Sand, S. A. (2019). Small places, universal stories. Diversity, film policy and the geographical dimension of filmmaking. *Nordisk Kulturpolitisk Tidsskrift, 22*(1), 8–25. https://doi.org/10.18261/ISSN2000-8325-2019-01-02

Smits, R. (2024). Disruption in times of COVID-19? The hybrid film festival format. *Cultural Trends, 33*(3), 309–323. https://doi.org/10.1080/09548963.2023.2193872

Steemers, J. (2016). International sales of U.K. Television content: Change and continuity in "the space in between" production and consumption. *Television & New Media, 17*(8), 734–753.

Straubhaar, J. D. (1991). Beyond media imperialism: Assymetrical interdependence and cultural proximity. *Critical Studies in Mass Communication, 8*(1), 39–59. https://doi.org/10.1080/15295039109366779

Stringer, J. (2001). Global cities and the international film festival economy. In M. Shiel & T. Fitzmaurice (Eds.), *Cinema and the City: Film and urban societies in a global context* (pp. 134–144). New York: Blackwell Publishers. https://doi.org/10.1002/9780470712948.ch11

Tunstall, J. (2008). *The media were American: U.S. Mass media in decline*. Oxford: Oxford University Press.

UIS. (2024). *UNESCO Institute for Statistics (UIS) Database*.

Valck, M. (2016). Fostering art, adding value, cultivating taste: Film festivals as sites of cultural legitimization. In M. de Valck, B. Kredell, & S. Loist (Eds.), *Film festivals: History, theory, method, practice* (pp. 100–116). London (GB): Routledge.

Valck, M. (2016). Introduction: What is a film festival? How to study festivals and why you should. In M. de Valck, B. Kredell, & S. Loist (Eds.), *Film festivals: History, theory, method, practice* (pp. 1–11). London (GB): Routledge.

Valck, D., & Marijke. (2007). *Film festivals: From European geopolitics to global Cinephilia*. Amsterdam (NL): Amsterdam University Press.

Verhoeven, D., Coate, B., & Zemaityte, V. (2019). Re-distributing gender in the global film industry: Beyond #metoo and #methree. *Media Industries, 6*(1), 135–155. https://doi.org/10.3998/mij.15031809.0006.108

Viswamohan, A. I., & Chaudhuri, S. B. (2022). Traversing boundaries: Contemporary Hindi cinema at international film festivals. *South Asian Popular Culture, 21*(1), 89–103. https://doi.org/10.1080/14746689.2022.2115736

Vivarelli, N. (2023). *Berlin film festival and EFM Won't accredit companies or media outlets with ties to Russian or Iranian Governments*. Variety. https://variety.com/2023/film/global/berlin-film-festival-russia-iran-ban-1235493196/ (visited on 07/09/2024).

Weber, N. R., Marchand, A., & Kunz, R. E. (2024). The global impact of public and private funding on cultural and economic movie success: Evidence from German film funding. *Journal of Cultural Economics, 48*(2), 259–283. https://doi.org/10.1007/s10824-023-09486-7

Wilson, J., Arshed, N., Shaw, E., & Pret, T. (2017). Expanding the domain of festival research: A review and research agenda. *International Journal of Management Reviews, 19*(2), 195–213. https://doi.org/10.1111/ijmr.12093

World Bank Group. (2024). *World Development Indicators*.

Yan, Q., Zhang, H., & Li, M. (2012). Programming quality of festivals: Conceptualization, measurement, and relation to consequences. *International Journal of Contemporary Hospitality Management, 24*(4), 653–673. https://doi.org/10.1108/09596111211217932

Zemaityte, V., Rohn, U., & Ibrus, I. (to appear). Peripheral or not? Cinema of small countries in the programming of international film festivals.

Zemaityte, V., Karjus, A., Rohn, U., Schich, M., & Ibrus, I. (2023). Cinando film festival programming dataset. *Figshare*. https://doi.org/10.6084/m9.figshare.22682794.v1

Zemaityte, V., Karjus, A., Rohn, U., Schich, M., & Ibrus, I. (2024). Quantifying the global film festival circuit: Networks, diversity, and public value creation. *PLoS ONE, 19*(3), Article e0297404. https://doi.org/10.1371/journal.pone.0297404