

Supplement to:

Burgdorf, Katharina, and Henning Hillmann. 2024.  
“Identity from Symbolic Networks: The Rise of New  
Hollywood.” *Sociological Science* 11: 297-339.

# Identity from Symbolic Networks: The Rise of New Hollywood

Katharina Burgdorf and Henning Hillmann

## ONLINE SUPPLEMENT: DATA APPENDIX

**Table A.1:** Film genres (1930-1999)

| Genre       | Freq   |
|-------------|--------|
| action      | 4,296  |
| adventure   | 1,788  |
| animation   | 6,501  |
| biography   | 761    |
| comedy      | 10,491 |
| crime       | 2,429  |
| documentary | 4,609  |
| drama       | 7,147  |
| family      | 182    |
| fantasy     | 214    |
| film-noir   | 36     |
| history     | 79     |
| horror      | 1,336  |
| music       | 956    |
| musical     | 777    |
| mystery     | 359    |
| romance     | 313    |
| sci-fi      | 232    |
| short       | 4,213  |
| sport       | 18     |
| thriller    | 324    |
| war         | 24     |
| western     | 1,765  |
| NA          | 1,981  |
| Total       | 50,831 |

Note: Results from a sample of 50,831 unique films.

One film may belong to up to three genres. Films must belong to at least one genre to be included in the sample. Genre information for second and/or third genre is missing for 17,685 films (33.8%).

### A.1. Robustness Checks: User Preferences and Reference Inclusion in the IMDb

In the third section of our empirical analysis, we identify a Matthew effect in the distribution of references among films. Canonical films are those that receive a substantially greater number of

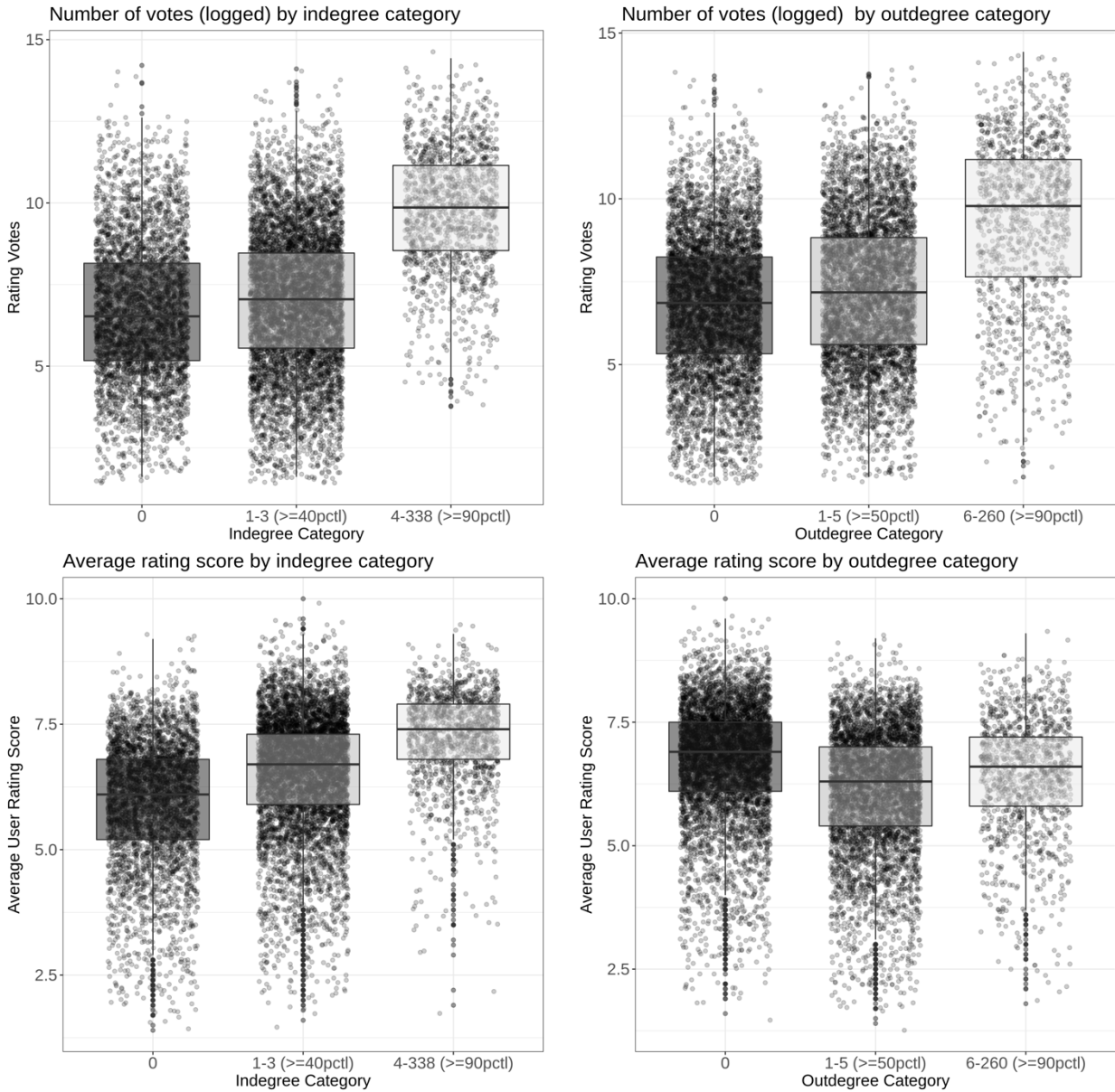
references than most other films. We may wonder, however, to what extent high degree scores are in fact driven by the preferences of IMDb users who enter information about films into the database. Because the IMDb is a user-generated dataset, we may end up with a selective sample of films and citation ties among them that merely reflect the taste of users for some films and their distaste of others. This selectivity is potentially problematic for our argument because the latter rests on the *status order* of films in the field of filmmaking, and not on the *popularity rank* of films among IMDb users. Hence, we ask: do some films score high on degree because they are influential among filmmakers, or because they are popular among IMDb users?

We address this caveat in two ways. First, we consider the correlation between the number of user votes for films and their average user rating scores, on the one hand, and, on the other, the indegree and outdegree in the citation network among films. The number of user votes reflects how recognized a film is among the IMDb audience, whereas the rating score tells us how valued it is. Both numbers indicate how popular a film is in the eyes of IMDb users. We assess the correlation between network degree (in and out) and user votes and ratings for the subset of 12,314 films that sent ( $n = 6,439$  films) or received ( $n = 8,273$  films) at least one reference. This may include cases where either indegree  $> 0$ , and outdegree  $= 0$ , or indegree  $= 0$ , and outdegree  $> 0$ .

Overall, we find only moderate correlations between network degree and user scores. Certainly, IMDb users are an expert group of film connoisseurs, and if a canon of influential films does exist, they should be able to identify such classics. Hence, a moderate correlation between a film's citations and users' votes should be expected. Indeed, the number of references that a film received correlates with the number of rating votes ( $r = 0.40$ ), which suggests that IMDb users recognize canonical films. When it comes to the valuation of films, the correlation

between indegree and the average rating score is unimpressive ( $r = 0.19$ ). Problematic for our argument would be a strong correlation between network outdegree and user scores because it may imply that, for films IMDb users hold in high esteem, they see references that may not even exist. This is not the case: the correlations between outdegree and the number of rating votes ( $r = 0.33$ ), and between outdegree and the average rating score ( $r = -0.01$ ) are even weaker than for indegree.

Second, beyond a simple summary statistic, we show that the status order among films (as measured by the number of references received and sent) is not determined entirely by the popularity rank of films (as measured by users' votes and ratings). In the boxplots in figure A.1, we compare three broad status groups of films (high, medium, low network degree) with respect to their average user votes and ratings. We logged the number of user votes because the underlying distribution is highly skewed (mean = 16,012.72; sd = 66,891.82). For indegree, we group films that received no references into the lowest status (indegree = 0;  $n = 3,986$  observations), films that received 1-3 references into the medium status (equal to, or above the 40<sup>th</sup> percentile in the degree distribution;  $n = 6,694$ ), and films that received 4 and up to 338 references into the highest status (equal to, or above the 90<sup>th</sup> percentile in the degree distribution;  $n = 1,433$ ). For outdegree, the low status group includes films that made no references to other films ( $n = 5,731$ ); the medium status group includes films that made 1-5 references (equal to, or above the 50<sup>th</sup> percentile in the degree distribution;  $n = 5,155$ ); and the high-status group entails films that made 6 and up to 260 references to other films (equal to, or above the 90<sup>th</sup> percentile;  $n = 1,227$ ).



**Figure A.1:** Distribution of IMDb user votes and average rating scores across indegree and outdegree categories.

If the popularity of films among IMDb users dictates the number of references sent and received, then we should observe little overlap in the distribution of user votes and ratings between the three status groups of films, and references should be concentrated in the most popular films. We may expect a moderate positive relationship if a film canon does exist and is recognized by an audience of connoisseurs (i.e., the films with high indegree also receive higher

user ratings). The boxplots suggest a slight tendency towards such a relationship for the number of user votes, which reflects mainly that users can recognize these films. More important, however, we find that the distributions of all three status groups overlap. Put differently, films with few, middling, or large numbers of references are all likely to receive low, middling, or high scores from IMDb users. This finding is particularly striking for the comparison of rating scores, and hence the valuation of films by users, beyond mere recognition. In sum, we find little evidence that supports the alternative argument that the recorded references among films in the IMDb are merely the revealed preferences of users for popular films at the expense of less popular ones.

## **A.2 Criteria for Selecting the Elite Group of New Hollywood Filmmakers**

To identify the leading proponents of New Hollywood, we relied primarily on the IMDb list of New Hollywood directors, which has been compiled by IMDb users.<sup>i</sup> We have cross-validated this sample with four influential film historical books on New Hollywood and two additional internet sources (Biskind 1999; Elsaesser, Horwath, and King 2004; King 2002; Thompson 1999).<sup>ii</sup> We excluded all filmmakers from our sample who are not mentioned at least once by any of the other sources. Three filmmakers had to be excluded: David Cronenberg, Randal Kleiser, and John Waters. We added filmmakers to the sample who are mentioned in at least four out of the six other sources but not in our main IMDb source of New Hollywood directors. We included three additional filmmakers: James Cameron, James Toback, and Paul Mazursky. While Michelangelo Antonioni was also mentioned in four sources, we decided to exclude him because he was more representative of European cinema, especially Italian Neorealism. Table A.2 shows our final list of 61 New Hollywood filmmakers.

---

**Table A.2:** Prominent New Hollywood film directors

---

|  |   |   |  |
|--|---|---|--|
| Abel Ferrara (SUNY Film and Media Studies)         | Jerry Schatzberg                                      | Norman Jewison  | Sam Peckinpah (USC Film)                           |
| Alan J. Pakula (Yale, Drama)                       | Jim Sharman (National Institute of Dramatic Art)      | Mike Nichols  | Roger Corman                                       |
| Arthur Penn  | Joe Dante   | Milos Forman  | Roman Polanski (National Film School Łódź, Poland) |
| Bob Rafelson                                       | John Boorman  | Monte Hellman   | Sidney Lumet (Columbia University, Theatre)        |
| Brian de Palma (Sarah Lawrence College, Theatre)   | John Carpenter (USC Film)                             | Paul Schrader (UCLA Film)                             | Stanley Kubrick                                    |
| Clint Eastwood                                     | John Cassavetes (American Academy of Dramatic Arts)   | Peter Bogdanovich (Stella Adler Conservatory, Acting) | Steven Spielberg                                   |
| Paul Mazursky (Brooklyn College, Theatre)          | John G. Avildsen (NYU)                                | Peter Yates (Royal Academy of Dramatic Art)           | Stuart Hagmann                                     |
| David Lynch (AFI Conservatory)                     | John Landis   | Philip Kaufman  | Stuart Rosenberg (NYU, editing)                    |
| Dennis Hopper (Actors Studio, Old Globe Theatre)   | John Milius (USC Film)                                | Ralph Nelson  | Sydney Pollack (Playhouse School of the Theatre)   |
| Don Siegel   | John Schlesinger (Oxford University Dramatic Society) | James Toback  | Terrence Malick (AFI Conservatory)                 |
| Francis Ford Coppola (UCLA Film)                   | James Cameron   | Richard C. Sarafian (NYU, Film)                       | Tobe Hooper (University of Texas, Drama)           |
| Franklin J. Schaffner                              | Lawrence Kasdan (UCLA, did not graduate)              | Richard Donner (NYU, did not graduate)                | Warren Beatty (Stella Adler Studio of Acting)      |
| George A. Romero (Carnegie Mellon School of Drama) | Martin Scorsese (NYU Tisch)                           | Ridley Scott  | Wes Craven   |
| George Lucas (USC Film)                            | Mel Brooks  | Robert Altman   | William Friedkin                                   |
| George Roy Hill (HB Studio (theatre))              | Michael Cimino (Actors Studio)                        | Robert Benton   | Woody Allen (NYU Tisch, Film, did not graduate)    |
| Hal Ashby  |   |   |  |

---

Note: Main source: <https://www.imdb.com/list/ls073927086/> (accessed September 24, 2023). Enrollments in film or theatre study programs in parantheses (sources: IMDb, LA Times, Washington Post, The Guardian, Wikipedia).

**Table A.3:** Network Statistics: The New Hollywood elite (incl. additional professional roles: writer, director, cinematographer, editor, producer)

|                        | 1950-1959 | 1960-1969 | 1970-1979 | 1980-1989 | 1990-1999 |
|------------------------|-----------|-----------|-----------|-----------|-----------|
| Collaboration Network  |           |           |           |           |           |
| Num Filmmakers         | 13        | 47        | 60        | 58        | 51        |
| Num Edges              | 1         | 16        | 36        | 25        | 12        |
| Mean Degree            | 0.15      | 0.68      | 1.20      | 0.86      | 0.47      |
| SD Degree              | 0.38      | 1.14      | 1.71      | 1.53      | 1.03      |
| Num. Comp.             | 12        | 34        | 32        | 40        | 42        |
| Prop. In Largest Comp. | 0.15      | 0.19      | 0.42      | 0.24      | 0.14      |
| Network Integration    | 0.09      | 0.06      | 0.19      | 0.07      | 0.04      |
| Prop. Isolates         | 0.85      | 0.66      | 0.48      | 0.59      | 0.75      |
| Co-Citation Network    |           |           |           |           |           |
| Num Filmmakers         | 7         | 36        | 57        | 55        | 49        |
| Num Edges              | 0         | 32        | 438       | 469       | 347       |
| Mean Degree            | 0.00      | 1.78      | 15.37     | 17.05     | 14.16     |
| SD Degree              | 0.00      | 2.14      | 10.57     | 12.12     | 10.48     |
| Num. Comp.             | 7         | 16        | 2         | 4         | 8         |
| Prop. In Largest Comp. | 0.14      | 0.56      | 0.98      | 0.95      | 0.86      |
| Network Integration    | 0.14      | 0.32      | 0.97      | 0.89      | 0.74      |
| Prop. Isolates         | 1.00      | 0.39      | 0.02      | 0.05      | 0.14      |



**Table A.4:** Network Statistics: All Filmmakers (incl. additional professional roles: Writer, director, cinematographer, editor, producer)

|                         | 1930-<br>1939         | 1940-<br>1949 | 1950-<br>1959 | 1960-<br>1969 | 1970-<br>1979 | 1980-<br>1989 | 1990-<br>1999 |
|-------------------------|-----------------------|---------------|---------------|---------------|---------------|---------------|---------------|
|                         | Collaboration Network |               |               |               |               |               |               |
| Num Filmmakers          | 4,445                 | 4,348         | 4,411         | 5,271         | 7,132         | 10,493        | 17,338        |
| Num Edges               | 85,029                | 67,923        | 42,912        | 28,931        | 37,269        | 65,883        | 132,720       |
| Mean Degree             | 38.26                 | 31.24         | 19.46         | 10.98         | 10.45         | 12.56         | 15.31         |
| SD Degree               | 48.31                 | 36.01         | 23.67         | 12.84         | 12.61         | 15.43         | 19.94         |
| Modularity/Log Nodesize | 0.047                 | 0.054         | 0.057         | 0.069         | 0.068         | 0.062         | 0.055         |
|                         | Co-Citation Network   |               |               |               |               |               |               |
| Num Filmmakers          | 1,089                 | 1,392         | 1,219         | 1,541         | 2,684         | 5,296         | 7,966         |
| Num Edges               | 15,681                | 28,570        | 12,598        | 27,462        | 168,772       | 1,337,871     | 4,077,973     |
| Mean Degree             | 28.80                 | 41.05         | 20.67         | 35.64         | 125.76        | 505.24        | 1023.84       |
| SD Degree               | 39.45                 | 54.02         | 30.10         | 49.54         | 153.85        | 531.14        | 1024.28       |
| Modularity/Log Nodesize | 0.067                 | 0.061         | 0.074         | 0.059         | 0.037         | 0.023         | 0.016         |

### A.3. Robustness Checks: Validity of Citations in the IMDb

Another caveat of our analysis may be that the set of citable films we have identified as the canon was not intended as such by New Hollywood directors. Ideally, we would assess the match between IMDb user's registered references and directors' intentions. Unfortunately, filmmakers rarely disclose their references in public. Professional film critics, however, play an important role in conveying a film's cultural value and meaning to a broader audience. Precisely because New Hollywood was inspired by film criticism, the interpretations by professional critics come close to the artistic intentions of directors. To address the caveat, we therefore assess to what extent the references identified by IMDb users match those mentioned in critical reviews. Close matches, especially on canonical films, would indicate that citations recorded in the IMDb are not mere artefacts of users' imagined references.

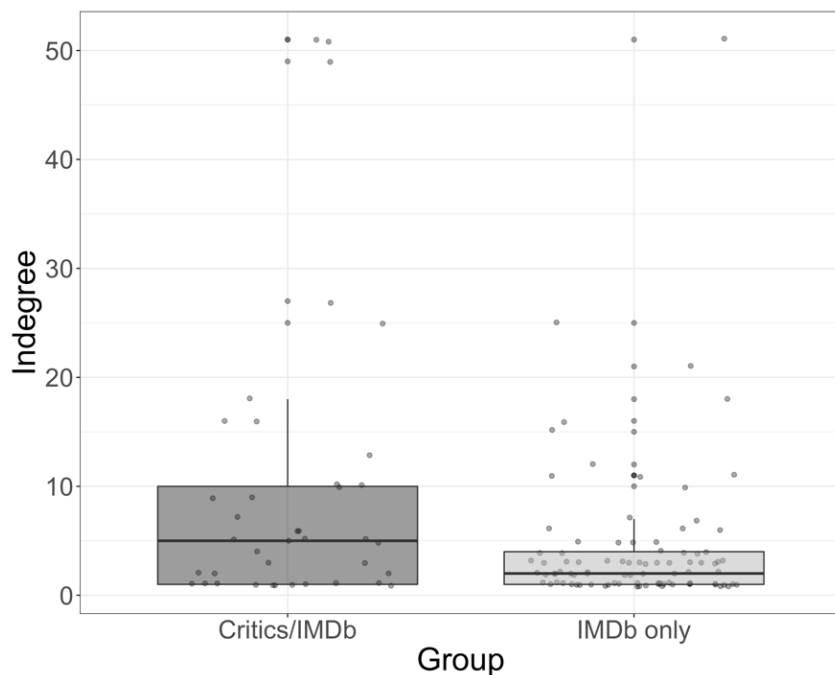
**Table A.5:** References to Films identified in the IMDb and in Critical Reviews

| Citing film                            | Cites listed by IMDb and critics (n = 39, with 36 unique cited films)                                  | Cites listed only by IMDb (n = 90, with 88 unique cited films)  | Source                                    |
|--|--|---|---|
| Alice Doesn't Live Here Anymore (1974) | The Wizard of Oz; Coney Island   | The Postman Always Rings Twice  | NY Times (1975); Chicago Sun-Times (1974) |
| All That Jazz (1979)                   | 8 ½; Lenny; 42nd street  | A Streetcar Named Desire; Cabaret; National Lampoon's Animal House; The Wiz   | NY Times (1979); New Yorker (1980)        |
| Annie Hall (1977)                      | Scenes from a marriage; The Misfits; The Godfather; La Strada; Fellini's Satyricon; Fellini's Casanova | La Grande Illusion; Snow White and the Seven Dwarfs; The Wizard of Oz; Children of Paradise; Singin' in the Rain; Juliet of the Spirits; House of Evil; Messiah of Evil; The House of Exorcism; Face to Face                                    | NY Times (1977); Washington Post (1977)   |
| Apocalypse Now (1979)                  | The Sweet Life; The Bridge on the River Kwai   | Citizen Kane; Lawrence of Arabia; Dr. Strangelove or: How I Learned to Stop Worrying and Love the Bomb; The Birth of a Nation; 2001: A Space Odyssey; Kelly's Heroes; Deliverance; Nashville; Stachka; Aguirre, the Wrath of God; The Passenger | NY Times (1979)                           |
| Bananas (1971)                         | Battleship Potemkin  | Duck Soup; Modern Times; Wild Strawberries; One, Two, Three; Casino Royale; The Incident  | NY Times (1979)                           |
| Bonnie and Clyde (1967)                | Battleship Potemkin  | -   | New Yorker (1967)                         |
| Carrie (1976)                          | Psycho; Sisters  | Battleship Potemkin; Lord of the Flies; A Man and a Woman; Deliverance; A Brief Vacation  | NY Times (1976)                           |
| Love and Death (1975)                  | The Seventh Seal; Persona  | The Hunchback of Notre Dame; Casino Royale; Crime and Punishment; A Night at the Opera; The Czar Wants to Sleep   | NY Times (1977)                           |
| New York, New York (1977)              | The Clock; Singin' in the Rain; A Star Is Born; On the Town  | 42nd Street; Paris Underground; Summer Stock; Guys and Dolls; West Side Story   | NY Times (1977); Chicago Tribune (1977)   |
| Obsession (1976)                       | Vertigo  | Dial M for Murder; High and Low; Marnie   | NY Times (1976; 1977)                     |

|   |   |  |  |
|---|---|--|--|
| Play It Again, Sam (1972)                 | Casablanca; The Big Sleep   | The Barefoot Contessa; The Treasure of the Sierra Madre; A Star Is Born; The Maltese Falcon; Across the Pacific; They Drive by Night; The Jungle Princess; The Petrified Forest; Marked Woman; To Have and Have Not; Dead Reckoning; Key Largo; The African Queen; Sirocco; The Unfaithful; Safari; Le coppie; All Through the Night | NY Times (1977); Chicago Sun-Times (1972)                              |
| Star Wars: Episode IV - A New Hope (1977) | Metropolis; 2001: A Space Odyssey; The Searchers; Seven Samurai; The Hidden Fortress; Yojimbo; Sanjuro; The Wizard of Oz; Triumph of the Will; Flash Gordon | Alexander Nevsky; The Adventures of Tartu; Twelve O'Clock High; The Dam Busters; The 7th Voyage of Sinbad; Prince of Space; Battle in Outer Space; 633 Squadron; 21-87; The Magic Serpent; The Good, the Bad and the Ugly; Five Easy Pieces; THX 1138; A Touch of Zen; The Getaway; One of Our Dinosaurs Is Missing; Dersu Uzala;    | NY Times (1977, 2015); Chicago Reader (1985); Chicago Sun-Times (1977) |
| The Last Picture Show (1971)              | Red River; Wagon Master; Father of the Bride;   | Winchester '73; Sands of Iwo Jima; White Heat; The Steel Helmet  | Chicago Sun-Times (1971); NY Times (1977)                              |
| Three Women (1977)                        | Persona   | Thoroughly Modern Millie; The Stepford Wives   | NY Times (1977)  |

Table A.5 lists the matchings for 14 selected New Hollywood films, produced between 1967 and 1979. Within this selective sample, 39 references to other films listed by IMDb users have also been identified by critical reviews in leading newspapers and magazines. While this is a considerable overlap, IMDb users also coded 90 references that have not been mentioned in critical reviews. Film critics, however, do not necessarily focus on references to earlier films in their writing. More important for our purposes, both IMDb users and critics should recognize references to canonical works that are widely cited and pivotal for network cohesion. This is precisely what figure A.2 shows: the average number of a film's received citations (indegree) is

significantly greater for references that have been identified by both IMDb users and critics than for references that are listed exclusively by IMDb users. Likewise, the upper limit of the interquartile range is higher for references detected by both IMDb users and critics than it is for references that are listed only by IMDb users. We conclude that references to canonical films (i.e., with high indegree) are the ones that IMDb users and professional film critics agree on, and there is little evidence to suggest that the prominence of films in the citation network is merely the outcome of choices made by IMDb users.



**Figure A.2:** Distribution of the number of citations received (network indegree): comparison of referenced films identified by both IMDb users and film critics (left;  $n = 39$ ) and identified exclusively by IMDb users (right;  $n = 90$ ). Means comparison: 9.87 citations for the left group vs. 4.31 citations for the right group ( $t = 2.44$ ;  $p = 0.02$ ).

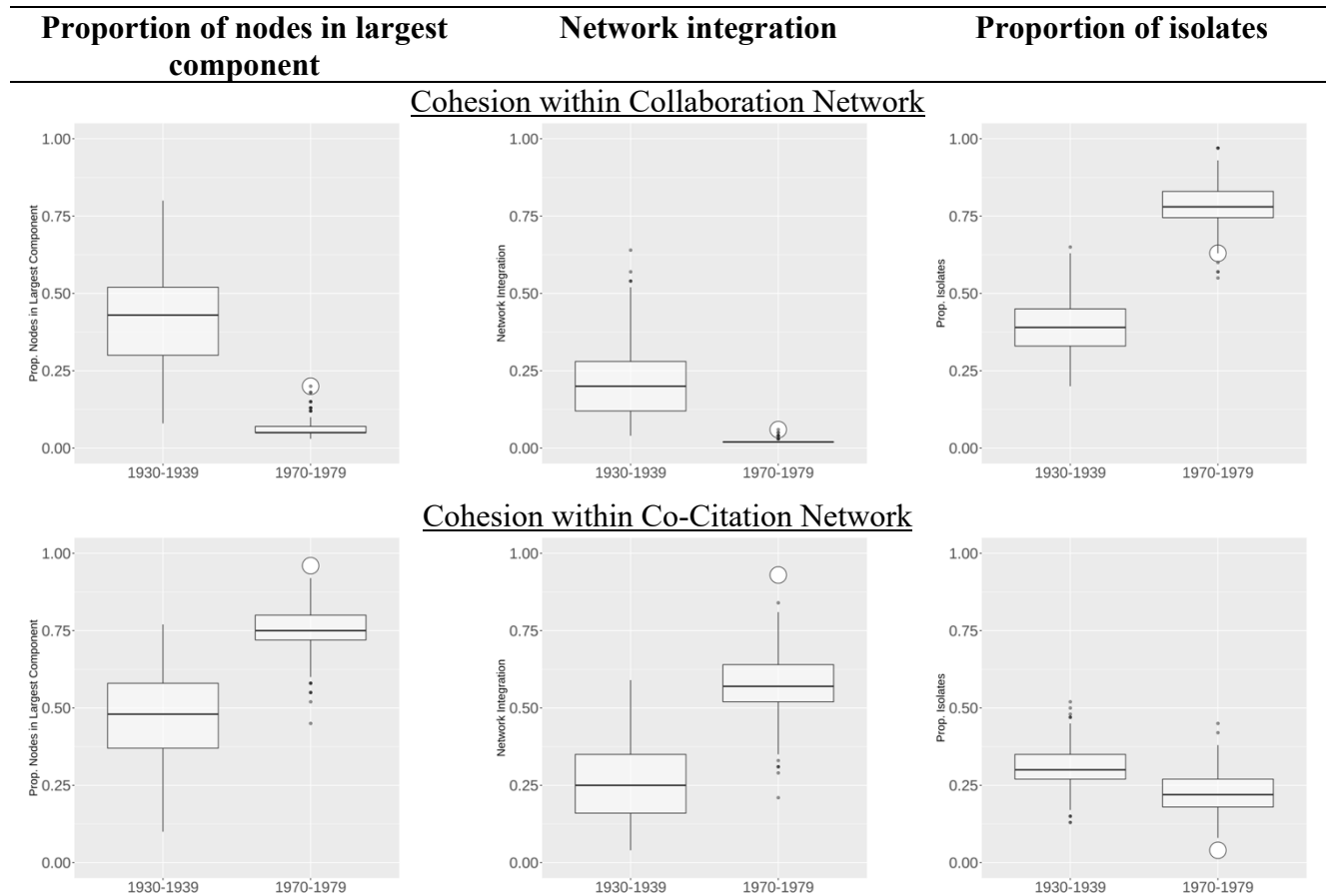
#### **A.4. Robustness Checks: How Comparable is the Extent of Cohesion within the New Hollywood Elite Network to Golden Age Filmmakers?**

Our results indicate that the collaboration networks of elite New Hollywood filmmakers displayed little cohesion. In contrast, their co-citation networks showed high levels of cohesion. These findings suggest that filmmakers cohered symbolically through shared references to revered films and less through relationships of direct cooperation in film projects. We may wonder, however, how comparable these findings are to other groups of filmmakers: do we find similar levels of cohesion among New Hollywood's contemporaries and among filmmakers active in earlier years? Meaningful comparisons in the American film industry are between Golden Age of Hollywood cinema (1930-39) and New Hollywood filmmakers (1970-79). Because it is difficult to delineate clear boundaries around alternative groups of filmmakers in both periods, we draw 500 repeated random samples of 60 filmmakers and their ties, corresponding to the size of the elite New Hollywood network at that time we reported earlier. To be included, filmmakers must have participated in at least five film projects, corresponding to the median number of film projects that New Hollywood filmmakers were involved in between 1970 and 1979.

The box-plots in figure A.3 compare differences in cohesion in the collaboration and co-citation networks between the group of New Hollywood directors (observed values indicated by the white dot in the 1970-79 period), the random selections of Golden Age filmmakers (left-hand boxes), and the random selections of contemporaries of the New Hollywood movement (right-hand boxes). As before, we use the proportion of nodes in the largest component, network integration and the proportion of isolates as our indicators of cohesion. Considering ties of collaboration, the results show that levels of network cohesion were consistently lower and the number of isolates greater among elite New Hollywood filmmakers and their contemporaries in

the 1970s than among filmmakers at the height of Hollywood's Golden Age in the 1930s. The lack of cohesion through collaboration may have been a general trend in the 1970s. Still, in line with our argument, we do find that elite New Hollywood filmmakers, differed clearly from their predecessors in the 1930s in their reliance on direct collaborations in film production.

Turning to co-citation networks, the results show that cohesion through joint citations of other films was consistently stronger among elite New Hollywood filmmakers than among either their contemporaries in the 1970s or Golden Age filmmakers. The evidence thus supports our argument that the proponents of New Hollywood relied much less on direct collaborations to cohere as a movement than filmmakers had done in the past. Instead, cohesion among New Hollywood filmmakers as a movement emerged primarily from densely connected symbolic networks of joint references.



**Figure A.3:** Comparison of cohesion in Golden Age filmmakers' randomly drawn collaboration and co-citation network (1930-1939) and New Hollywood filmmakers' randomly drawn and observed (white dot) collaboration and co-citation network (1970-1979).

<sup>i</sup> <https://www.imdb.com/list/ls073927086/> (accessed June 7, 2023).

<sup>ii</sup> [https://en.wikipedia.org/wiki/New\\_Hollywood](https://en.wikipedia.org/wiki/New_Hollywood); <https://www.imdb.com/list/ls079993239/> (accessed June 7, 2023).