Competing for Prizes
Arts, Wines, Sports and Science

A flurry of prizes and awards

- There are more film prizes awarded each year than there are feature films produced (J. English)

  Gore Vidal: "The US has more prizes than authors"

  Peter Porter: There are so many prizes in Australia that "there is hardly any writer in Sidney who has not won one"

- People: From Michael Jackson ($\geq$ 240) to Einstein (1)

- Amounts: From the Nobel ($1.7 million) to the Hey Frey Memorial Award ($100)

- If some of you are interested, there exists a reference work *Awards, Honors, and Prizes* – two volumes, 3,000 pages thick, $1,000 new...

- But only $20 used

Why prizes?

- Demand from consumers to sort out and guide

- Often prizes focus attention on people whose reputations are already solidly established (not so for sports)

- Turn things into marketing gimmicks or horse races: "Musical competitions are for horses, not for musicians" (B. Bartok)

- Above all, "provide a closed elitist forum where cultural insiders – artists, critics, functionaries, sponsors, publicists, journalists, consumers, kibitzers and beggars – engage in political influence peddling and mutual back scratching" (J. English)

Why prizes?

- Do they reward excellence or mediocrity?

- This is not always contradictory

- In 1997, the movie *Diabolique* with Sharon Stone won the Award of Distinction for Feature Productions Cinema from the Australian Cinematographers Society, while Sharon Stone was nominated as Worst New Star for the same movie by the Razzie Awards

- In 1998, Kevin Costner (in *The Postman*) was nominated as Best Actor by the US Academy of Science Fiction, Fantasy and Horror Films and as Worst Actor and Worst Director by the Razzie Awards

Why prizes?

- Prizes are sometimes created to repair the failures made by existing prizes

- The US National Book Award established itself to correct the fact that the Pullitzer never recognized Faulkner

- After a couple of years (1975), both were awarding the same authors. Arrives, therefore, the new National Book Critics Circle Award

- In 1981, all three converge and give their prize to the same book, John Updike's *Rabbit is Rich*

What do winners say?

- Eddie Vedder (rock musician, Grammy Award): "I don't know what this means. I don't think it means anything"

- John Berger (Booker prize acceptance speech, 1972): "You may like to know, briefly, what (this prize) means to me. The competitiveness of prizes I find distasteful… (T)he publication of the shortlist, the deliberately publicized suspense, the speculation of writers, the whole emphasis on winners and losers is false and out of place in the context of literature"

- Tolstoy : "(T)he lack of worth, the harm in money prizes"

- Jean-Paul Sartre turned down the Nobel (though politely)

Problems and some aesthetics

- Fields in which judgments can hardly be verified by others. Can one make the process objective?

- This is one of the topics hotly discussed in aesthetics

- One view (taken up in the 20th century by analytical philosophers) is that artworks contain intrinsic properties which can be listed and "graded," and the grades can be "aggregated" to reach an overall evaluation

- Remarkable example of the *Balance des Peintres* by de Piles (1635-1709)
  - Decomposes painting into 4 properties: composition, drawing, color, expression
  - Grades each property on a scale 0 to 20 for 56 painters
  - Does not say how to aggregate, and does not do so

De Piles Balance des Peintres

Painter	Composition	Drawing	Color	Expression
Raphael	17	18	12	18
Rubens	18	13	17	17
Titian	12	15	18	6
Rembrandt	15	6	17	12
Leonardo	15	16	4	14
Durer	12	16	9	8
Veronese	15	10	16	3
...				

Problems and some aesthetics

- Empirical philosophers (Hume, 1757):
  - "aesthetic beauty does not lie in works, but in the mind of those who look at them...
  - the basics of evaluation and taste are to be found in those who judge works (true judges)"

- Test of time (Hume, 1757): Time makes it possible to reduce the "noise" due to fashion, envy and jealousy, present in evaluations made shortly after a work is produced. Time sorts out works that transcend fashion

- Sociology of art (Bourdieu)
  - Evaluation comes from social conventions, and is objective only because those who carry it out are empowered with social authority
  - Artistic genius does not exist

How do awards fare?

- In the arts
    - Movies: Oscars, Golden Globes, and other movie awards
    - Musical interpretation: Queen Elisabeth Musical Contest
- In other fields
    - Wines
    - Ice skating
    - Scientific papers
- Two tools
    - The test of time which needs two kinds of evaluations:
        - Immediate recognition
        - Measure of long-term reputation
    - Some econometrics

Test of time. US Movies 1950-80

- Immediate recognition:
  - Oscars, Golden Globes, New York Film Critics Circle, and National Board of Reviews
  - All movies nominated and awarded from 1950 to 1980

- Survival of movies in three "100 Best Movies" lists in 2000, that is 20 to 50 years later

  - *American Film Institute America's 100 Greatest Movies*: a selection of the 100 best movies of all times, chosen by over 1,500 experts and professionals of the US movie industry

  - *Mr. Showbiz Critics Pics: 100 Best Movies of All Time*, obtained from the answers of movie critics each of whom is asked to list his 10 preferred movies of all times

  - The *100 Must-See Films of the 20th Century*, by the American movies critic Leonard Maltin

Test of time. US Movies 1950-80

- **The Oscars**
  - 20 out of 31 winners (65%) that appear on at least one of the top-100 lists, have received an Oscar, but only 8 (26%) appear in all three lists in 2000
  - In 18 years out of 31, the National Academy failed to award the best movie Oscar to the one that, with the passage of time, is considered "the best"

- **Examples**
  - In 1952, *Singing in the Rain* (cited in all three 2000 lists) was not even nominated, while *The Greatest Show on Earth*, which appears in no such list, was awarded the Oscar
  - In 1959, Hitchcock's *North by Northwest* as well as *Some Like it Hot*, appear in all three lists, while *Ben Hur*, in no list, received the Oscar (*)
  - In 1968, the Academy failed to choose as best movie Kubrick's *2001, A Space Odyssey* which, in 2000, is the *only* 1968 movie in all three lists

Test of time. US Movies 1950-80

- Golden Globes
  Only six out of the 65 winning movies make it to all three
  2000 best movie lists, while 45 are in no list

- The National Board of Reviews
  Two out of 31 winning movies in all three lists

- The New York Films Critics Circles
  Four out of 29 winning movies in all three lists

Test of time. US Movies 1950-80

	Oscars	G. G.	NBR	NYFC
No. of winning movies	31	65	31	29
No. of years with prizes	31	31	31	29
Best movie missed	18	26	26	21
No. of movies later considered to be better than the winner	29	50	57	39

Econometrics. Ranking musicians

- **The Queen Elisabeth Piano Competition**
  International competition for piano (violin and singing)
  performing, organized in Belgium and considered among
  the best and most demanding in the world

- **Problem**
  - Is success due to ranking? • Is success due to talent?
  - Is ranking due to talent?

- **Difficulty**
  Disentangle the effects of ranking and unobserved talent (if
  talent were observed, it would be easy)

- **A few names of first ranked:**
  Leonid Kogan (1951), Leon Fleisher (1952), Vladimir Ashkenazy (1956),
  Malcolm Frager (1960), Eugene Moguilevsky (1964), Valery Afanassiev
  (1972)

Ranking musicians

- **Working of the competition: Stages**
  - Starts with a random drawing of order of appearance kept in all later stages:
  - Stage 1: eliminates all but 24 semi-finalists
  - Stage 2: eliminates another 12
  - Stage 3: 12 finalists perform, 2 per evening
  - The twelve finalists are ranked 1 to 12 (6 since 1995)

- **Data**
  - 11 succcessive competitions (1952 to 1991) 132 musicians

  - Description of musicians (age gender, citizenship, order of appearance, final ranking

- **Two success indicators**
  - No. of CDs and LPs recorded and present in1995 in Belgian (2), UK, and French catalogues

  - Gradings by 11 Belgian musical critics (0 to 4)

Econometrics

- Equation
  $$s_i = \gamma_0' + \gamma_1' r_i + \gamma_2' q_i + u_i'$$

  - s_i is a variable representing the success of i
  - r_i is the rank of i after the competition (reordered 12 for the first, etc)
  - q_i is the *unobserved* talent or quality of i
  - u_i' is a nicely behaved error term
  - If $E(r_i, u_i') = E(q_i, u_i') = 0$, γ_1' and γ_2' can be estimated consistently by OLS

- Since q_i is unobserved, we estimate $s_i = \gamma_0 + \gamma_1 r_i + u_i$, where $u_i = u_i' + \gamma_2' q_i$.
  Since r_i is likely to be correlated with q_i, $E(r_i, u_i) \neq 0$ and an OLS estimated γ_1 will be a biased and inconsistent estimator of the causal effect of ranking on success.

Econometrics

- Solution: Find instruments z_i, uncorrelated with unobserved q_i (and hence with u_i), but correlated with r_i

- This adds a second equation $r_i = \beta_0 + \beta_1 z_i + v_i$

- Instruments: Order of appearance, gender:

$$r_i = \beta_0 - 2.96\, first - 1.13\, last - 1.86\, female$$
$$(3.1) \qquad (2.9) \qquad (1.9)$$

where $first$ is the very first who plays on the first evening

$last$ is the second who plays every evening

Econometric results

Effect of Ranking on Success

	Presence in catalogues	Ratings by critics
Simultaneous estimation		
TSLS	0.188	2.350
	(2.5)	(3.5)
LIML	0.189	2.416
	(2.5)	(3.1)
OLS	0.092	1.475
	(3.2)	(5.4)

Conclusions

- Those who are better ranked are more successful: The final ranking does definitely affect later success

- Is this due to the fact that the competition correctly selects those who are more talented?

- Since talent is not observed, there is need to use instrumental variables estimation. Here we use the fact that *the final ranking by experts is correlated with the order in which the performers appear in the competition*

- But order of appearance is *randomly* selected before the competition starts

- Therefore the ranking is random
  Those who are more successful are probably not those who are more talented

Wines

- 4,167 wines entered in 13 wine competitions in the U.S. awarding Gold (96), Silver (90), Bronze (84) or nothing (80) Distribution of the 78 (13*12/2) correlation coefficients:

Less than 0.10	46
Between 0.11 and 0.20	25
Larger than 0.20	7

- Consensus among competitions: 375 wines in 5 competitions. Is Gold a sign of quality?

Gold in 0/5	243
Gold in 1/5	106
Gold in 2/5	20
Gold in 3/5	6
Gold in 4/5	0
Gold in 5/5	0

  These numbers are compatible with:

  Chance alone seems to account for the number of Gold medals that a wine receives

Figure Skating

- When evaluators are appraised through comparison with their peers, they have an incentive to distort their assessments and to be biased towards general consensus, that is often derived from assumptions based on pre-performance public information

- To show this, the author (Lee) exploits two kinds of variations in skating competitions 2001-2003:

  - the variation of an individual judge's aversion to outlying scores over the course of a performance, according to the degree of his previous scorings within the game
  - across the board changes in all the judges' incentives for outlier aversion due to the judging system reform after the 2002 Olympic scandal

  (The 2002 scandal: a French judge was convinced by the head of the French team leader to vote for the Russians instead of the Canadians, as part of a deal that the Russians would vote for the French in the ice dance competition)

Why is this so?

- In figure skating, the panel of judges and the referee are chosen by the International Skating Union (ISU) from a pool of international judges recommended by national federations

- Within 30 seconds after each performance, they submit two scores: a technical score and an artistic score, that are combined to form the total. Each score is then displayed on a big scoreboard

- Judges are monitored and assessed by the ISU, and critiqued by referees. Any disagreement (deviation) w.r.t. the other judges has to be explained by the "faulty" judge. Those who do not attend or cannot answer are penalized, and may not be invited again, a lost honor. Hence their outlier aversion...

Scientific Papers

- This is a combination of the "test of time" and econometrics

- The new fashion among scientist is to measure the quality of a paper by the number of times it gets cited by other scientists

- It is reasonable to assume that one gets tired after having looked at the three or four first papers of an issue. Therefore papers that appear first in an issue are probably more read and if they are good, cited by others

- Editors of journals who try to maximize the number of times their own journal is cited seem to have noticed that papers that appear first in an issue get more citations

- Therefore in order to maximize even more, they carefully judge the quality of papers and put as first those they think are best

Scientific Papers

- And it works!

- *American Economic Review* 1986-2000

Order in issue	Mean no. of citations
First	45.4
Second	49.0
Third	37.8
Fourth	42.5
Fifth	30.3

Scientific Papers

- Yes it works, but does that mean that editors are good judges?

- Between 1975 and 1997, the *European Economic Review* ran an "experiment," probably not on purpose. This is called a "natural experiment"

- In some issues, papers were published in alphabetical order, which could be taken as a random order, while in some other issues the order was "clever," that is taking into account the judgement of the editor on the quality of papers

- Therefore, those in non-alphabetically ordered issues should, on average, collect more citations than those in randomly ordered issues. Do the two orders generate different numbers of citations of the first paper?

Scientific Papers

	EER alpha		EER non-alpha		AER	
	Marg. eff.	St. err.	Marg. eff.	St. err.	Marg. eff.	St. err.
First in issue	1.93	0.48	2.78	0.32	4.90	0.61
Second in issue	0.19	0.42	1.21	0.28	4.81	0.59
Third in issue	-0.08	0.41	-0.66	0.25	0.08	0.56
Fourth in issue	0.75	0.46	-1.27	0.24	2.21	0.57
Fifth in issue	0.92	0.47	-1.53	0.24	-2.01	0.56
Other control variables						
Mean	5.12		5.43		23.24	
No. of obs.	303		760		1198	

1.93/2.78= 69 % of the marginal effect of citations to the first paper are due to the "first paper effect," and 31 % to judgement!
Or 2/3 of the papers that appear in the first position are more cited without being any better

General conclusions. Other cases

- Statistical procedures, based on the grading of explicit characteristics, perform much better in diagnosing health conditions than clinical methods, which rest on implicit mental processes

- Data from a test used to diagnose brain damage were given to a group of clinical psychologists and their secretaries. The psychologists' diagnoses were no better than those made by their secretaries

General conclusions. Differences

- Prized movies are often also box office successes

- This is not so with prize-winning books: Between 1980 and 1989 for example, no Pulitzer, winning or nominated title made it to the list of the 15 bestsellers in the United States

- Movies and books are not necessarily produced with a competition in mind and the artist is not present during the evaluation stage

- In musical competitions (as well as in some sports, such as figure skating, diving, horse dressage), the work is unwinding, judged in several stages, and in the presence of the "artist" who is producing it, which means that the "art" may reflect other qualities than talent

General conclusions. Similarities

- Awards, prizes, and critics may have an impact on success: Clear for the music example, certainly so for movies but less so in the case of books

- However
  - Movies: award-winners are often not the best ones 20 to 50 years later
  - Books: no significant difference between prize-winning and shortlisted authors – and more importantly, no Nobel for Tolstoy, Kafka, Proust, Joyce, Zweig, Calvino, Max Frisch, but one for Sienkiewicz and one for Claude Simon, the "super-boring" French writer
  - Music: interpreters are not ranked according to talent

General conclusions. Problems

- Who votes or makes the selection?
  - In many cases, several rounds of judgments, with *different* judges who pass their results to the next round, so that the members of the jury who make the final choice have not seen or read all the works
  - In some cases, the jury passes its choice(s) to the Board of the awarding foundation, and Board members make the final selection

General conclusions. Problems

- Turning in the grades
  - In some competitions the grades must be entered after each performance (figure skating) or evening (Queen Elisabeth competition) and cannot be changed later

- Unfair or cheating judges
  - Patriotic favoritism in sports competitions is often an issue (including manipulating injury time in team sports such as soccer, to promote the victory of the home team)

  - Buying judges: case of the French judge for figure skating in the 2002 Olympics

General conclusions. Problems

- Wrong incentives given to judges
  - Usually not well paid, but find it "morally" rewarding

- Centrality bias: Converging to the mean

  - Skating: Judges are monitored and questioned about "outlying" grades with respect to their peers. They may be penalized, or banned from future prestigious competitions, which induces them not to submit such grades

  - Queen Elisabeth Competition: Excluding extreme grades

General conclusions. Suggestions

- Why is it necessary to provide a winner?

  People like blood. They need winners and losers

- Choosing a winner from outside the set of candidates
  - Happened twice at the Yale Younger Poets series, under Auden's tenure as the *unique* judge. He awarded the prize to *non-submitted* manuscripts

- Not awarding the first prize
  - Yale Younger Poets Prize: Auden refused to award the prize in 1950 and 1955
  - Tchaikovsky piano competition: No first prize in 1982, 1994 and 2007
  - Chopin piano competition: No first prize in 1990 and 1995; no 4th prize in 1980
  - Van Cliburn competition: Two first prizes in 2001

- Since there is often little difference in quality or talent between winners and nominees, experts might become more reputable if they just picked say, five nominees, without ranking them

References

Coupé, T., V. Ginsburgh and A. Noury (2009), Are leading papers of better quality? *Oxford Economic Papers* forthcoming.

English, J. (2005), *The Economy of Prestige*, Cambridge, MA: Harvard University Press.

Ginsburgh, V. (2003), Awards, success and aesthetic quality in the arts, *Journal of Economic Perspectives* 17, 99-111.

Ginsburgh, V. and J. Van Ours (2003), Expert opinion and compensation: Evidence from a musical competition, *American Economic Review* 93, 73-90.

Ginsburgh, V. and S. Weyers (2008), Quantitative approaches to valuation in the arts, with an application to movies, in M. Hutter and D. Throsby, eds., *Beyond Price. Value in Culture, Economics, and the Arts*, New York: Cambridge University Press, 179-199.

Hume, D. (1965 (1757)), Of the standard of taste. In *On the Standard of Taste and Other Essays*, Indianapolis: The Bobbs-Merrill Company.

Lee, J. (2004), Outlier aversion in evaluating performance: Evidence from figure skating, IZA Discussion Paper 1257.

Tetlock, P. (2005), *Expert Political Judgment: How Good Is It?*, Princeton: Princeton University Press, 2005.