Comparing Artistic Values
The Example of Movies*

Victor Ginsburgh
ECARES, Université Libre de Bruxelles
CORE, Université catholique de Louvain

and

Sheila Weyers
Université catholique de Louvain

December 2004

Monroe Beardsley, Bruce Vermazen, and George Dickie¹ suggest that artworks carry independently valuable properties that can be rated. However, they point out that since these properties are incommensurable, one cannot infer the overall value of works, and compare them. To simplify the discussion, we restrict our attention to works that share the same properties, and show, using a very simple example, that this is not a sufficient condition to rank works. Assume that we want to do so for two works \(a\) and \(b\) endowed with three identical properties. Following Dickie, we denote by \(A\), \(B\) and \(C\) the three properties, and by the numbers 1, 2 and 3 the ranks of the properties in each work. \((A_3, B_1, C_1)\) is a work with rank 3 (the highest) in property \(A\) and rank 1 (the lowest) in \(B\) and \(C\). Consider the situation in which work \(a\) is rated \((A_3, B_1, C_1)\), while work \(b\) is rated \((A_1, B_3, C_2)\). If the properties are incommensurable, the two works cannot be compared.

There are however cases in which comparisons can be made, but these are far from covering all situations. An easy case is the one in which work \(a\) is described by \((A_3, B_1, C_1)\) and work \(b\) by \((A_2, B_1, C_1)\). Then \(a\) is overall more valuable than \(b\). A second case is the one in which property \(A\) is always "more important" than property \(B\), that is itself more important than \(C\) (a lexicographic, but somewhat arbitrary, ordering of properties). Work \(a\) rated \((A_3, B_1, C_1)\) would then be overall more valuable than work \(b\) rated \((A_2, .., .)\) or \((A_1, .., .)\), where the rating of properties \(B\) and \(C\) can be 3, 2 or 1.

Dickie suggests to construct matrices that allow comparisons with respect to a given work, say \(a\), as long as the works to which \(a\) is compared have more of one property, and not less of any other, or less of one property and not more of any other. For example, work \(a\) endowed with properties \((A_1, B_2, C_3)\) is better than the five

* We are grateful to Catherine Dehon for useful discussions and suggestions.
works that are located above \( a \), and worse than those located below \( a \) in the following matrix:\(^2\)

<p>| | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>( A_1, B_1, C_1 )</td>
<td>( A_1, B_1, C_1 )</td>
<td>( A_1, B_1, C_1 )</td>
</tr>
<tr>
<td>( A_1, B_2, C_1 )</td>
<td>( A_1, B_2, C_1 )</td>
<td>( A_1, B_2, C_1 )</td>
</tr>
<tr>
<td>( A_1, B_2, C_2 )</td>
<td>( A_1, B_2, C_2 )</td>
<td>( A_1, B_2, C_2 )</td>
</tr>
<tr>
<td>( A_2, B_2, C_3 )</td>
<td>( A_2, B_2, C_3 )</td>
<td>( A_2, B_2, C_3 )</td>
</tr>
<tr>
<td>( A_3, B_2, C_3 )</td>
<td>( A_3, B_2, C_3 )</td>
<td>( A_3, B_2, C_3 )</td>
</tr>
<tr>
<td>( A_3, B_3, C_3 )</td>
<td>( A_3, B_3, C_3 )</td>
<td>( A_3, B_3, C_3 )</td>
</tr>
</tbody>
</table>

Such a matrix contains only a partial ordering of works, but not a complete ordering, since we cannot decide whether a work with rates \( (A_2, B_3, C_1) \) is overall better or worse than \( a \).

Dickie\(^3\) claims that "there is no better way or even any other way at all to arrive at reasonable specific evaluations," and recognizes that critics do indeed rely on such (partial) orderings, though their evaluation is intuitive, and made without any formal construction of the matrices that he describes.

In this discussion, Dickie and Vermazen never mention that if properties could be ranked according to their order of importance, that is given weights, then an overall value could be computed by a weighted sum of the ratings. After all, if works can be rated or ranked according to each property, properties themselves can perhaps also be ranked, and their relative importance assessed as a consequence, though this may be considered a task that is even more formidable than the separate rating of each property. The lexicographic ordering that was discussed earlier, can be interpreted as an extreme weighting scheme, in which properties are ranked by order of assumed importance. The "most important" property is looked at first and works are ordered according to the rates of this first property, then, in case of ties in the first property, one orders according to the rates of the second property, etc. Another obvious scheme is to give equal weights to all properties, so that the overall value results from the mere addition of individual ratings. But one can imagine more subtle weighting schemes, and obviously, if art critics are able to decompose works into properties, to rate each property for each work, then they also weight the properties to obtain an overall value, even if this last step is informal. This is so when students are rated on the basis of several exams, in musical and other artistic competitions, including literary awards, and in some sports such as figure skating, diving, or gymnastics.

---

\(^2\) The example is borrowed from Dickie, pp. 167 and ff.

The idea of describing artworks by properties is not fully new. In his *Balance des Peintres*, the French art critic and historian Roger de Piles (1635-1709) rated Renaissance and Baroque painters according to a point system on four properties: drawing, color, composition and expression. Jonathan Richardson adds three more properties to those of de Piles: grace and greatness, invention and handling. This is very close to the idea expressed by economists, in particular, Kevin Lancaster, according to whom a commodity can be thought of as a bundle of characteristics (or properties), purchased by consumers for the value it provides by the combination of such characteristics. Since there are no markets for characteristics, these cannot be bought and combined by consumers to construct their preferred choices. Therefore producers (and artists) make these choices for them, and consumers choose the commodity (or the work) that provides the combination they find closest to their preferred combination.

We show how weights can be retrieved from observations in which art critics or experts rate properties, as well as the overall value of works. This will also make clear that experts do indeed implicitly use such weights.

Once weights are available, the overall value of a work can be obtained by a weighted sum of the values of all the properties or characteristics embodied in the work (such as actor and script for a movie, or composition and drawing for a painter). If work i can be fully described by say, three properties its total value $V_i$ is simply:

$$V_i = \beta_1 x_{1i} + \beta_2 x_{2i} + \beta_3 x_{3i},$$

where the $x_{ki}$ represent the rates of the three properties, and the $\beta_k$ are weights.

We illustrate the idea using movies for which we do observe both a rating of total value $V$ and the ratings of a certain number of properties, focusing on the 270 movies, both nominated and having been awarded the Oscar for "Best Picture" by the National Academy of Motion Picture Arts and Sciences between 1950 and 2003. In each year, five movies are nominated by the Academy and one of these wins the race ("...and the winner is..."). For each such movie, we also collected all other nominations and distinctions given by the Academy for a certain number of characteristics or properties: actor in a leading role, actress in a leading role, director, screenplay, etc. The properties that seem to matter for motion pictures may look quite different from the much more general and abstract ones such as unity, intensity.

---

5 Jonathan Richardson, *An essay on the whole art of criticism as it relates to painting*, London, 1719.
7 The data are available on http://awards.fennec.org.
8 See Appendix 1 for the full set of properties.
and complexity, postulated by Beardsley for example, but there is no reason to think that they contradict each other.

The data consist of rates of the overall value and rates for each property: V, the overall value, is dichotomous, with value 1 if the movie was awarded the Oscar, and 0 if it was nominated, but failed the Oscar. Each property (say "actor in a leading role") can fall in one of three categories: the movie was awarded the Oscar, or it was nominated, but failed getting the Oscar, or it was not nominated. A movie is thus represented by an array of numbers, the first of which is V (1 or 0) while the other are variables that also take values 1 or 0 according to whether they fall or not into one of the categories.

Table 1 displays the rates that four movies (All About Eve, produced in 1950, An American in Paris, and A Streetcar Named Desire both produced in 1951, and Cabaret, produced in 1972) were given on the various properties. We omit for some time the overall value.

Is it possible to compare All About Eve, An American in Paris and Cabaret? The answer is negative if we follow Vermazen or Dickie, since, as can be checked, only a partial ordering can be constructed. Indeed, the first work does not worse than the second on 11 properties, but does worse on cinematography, art direction and score, while the second does not worse than the first on nine properties, but the first does better on five. Comparisons between the first and the third movies, and between the second and the third one lead to similar conclusions. Therefore, one has to conclude that no overall judgment is possible on the basis of individual properties. The total number of O (Oscar), N (nominated) and F (failed) ratings may help (Cabaret would be first, All About Eve second, and An American in Paris last), but this is not the result of the Academy's choices, since Cabaret was not awarded the Oscar for best picture, while the two others received it. One may argue that judges change their way of evaluating over time. Let us therefore compare two works produced during the same year, An American in Paris with A Streetcar Named Desire. Again, the partial ordering suggested above will be of no help, nor is the total number of O, N and F ratings.

---

9 We ignore here all the movies produced in any given year, that were not nominated as best pictures by the Academy.

10 There will for instance be three variables for "Actor in a Leading Role:" the first takes the value "one" if the movie was awarded the Oscar, and "zero" otherwise, the second takes the value "one" if the movie was only nominated, and "zero" otherwise, the third takes the value "one" if the movie was not selected, and "zero" otherwise.
Table 1
An Example of Rates for Four Movies

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Actor in a Leading Role</td>
<td>F</td>
<td>F</td>
<td>N</td>
<td>N</td>
</tr>
<tr>
<td>Actress in a Leading Role</td>
<td>N</td>
<td>F</td>
<td>O</td>
<td>O</td>
</tr>
<tr>
<td>Actor in a Supporting Role</td>
<td>O</td>
<td>F</td>
<td>N</td>
<td>O</td>
</tr>
<tr>
<td>Actress in a Supporting Role</td>
<td>N</td>
<td>F</td>
<td>O</td>
<td>O</td>
</tr>
<tr>
<td>Director</td>
<td>O</td>
<td>N</td>
<td>N</td>
<td>N</td>
</tr>
<tr>
<td>Screenplay</td>
<td>O</td>
<td>O</td>
<td>N</td>
<td>N</td>
</tr>
<tr>
<td>Cinematography</td>
<td>N</td>
<td>O</td>
<td>O</td>
<td>N</td>
</tr>
<tr>
<td>Art Direction/Set Decoration</td>
<td>N</td>
<td>O</td>
<td>O</td>
<td>O</td>
</tr>
<tr>
<td>Costume Design</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>N</td>
</tr>
<tr>
<td>Score</td>
<td>N</td>
<td>O</td>
<td>N</td>
<td>N</td>
</tr>
<tr>
<td>Original Song</td>
<td>F</td>
<td>F</td>
<td>F</td>
<td>F</td>
</tr>
<tr>
<td>Film Editing</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>F</td>
</tr>
<tr>
<td>Sound Mixing</td>
<td>O</td>
<td>F</td>
<td>N</td>
<td>N</td>
</tr>
<tr>
<td>Special Effects</td>
<td>F</td>
<td>F</td>
<td>F</td>
<td>F</td>
</tr>
</tbody>
</table>

No. of "O" ratings  | 5  | 5  | 5  | 4  |
No. of "N" ratings  | 6  | 2  | 7  | 7  |
No. of "F" ratings  | 3  | 7  | 2  | 3  |

O: the movie was awarded the Oscar; N: the movie was nominated, but failed to receive the Oscar; F: the movie was not nominated.

Suppose now that one could guess the weights that critics give to the various properties, and assume that these are: 0.67 if the movie gets the Oscar for Actor in a Leading Role, 2.34 for Director, 1.16 for Screenplay, 1.06 for Costume Design, 0.61 for Film Editing, and 0 for all other properties. Then it is easy to calculate the weighted average for both movies, which is 1.16+1.06 for An American in Paris (Oscar for Direction with weight 1.16, plus Oscar for Costume Design with weight 1.06, no other Oscar with positive weight), and 0 for A Streetcar (no Oscar with positive weight), showing that the overall value is larger for An American than for A Streetcar, and hence An American has a larger overall value, and was indeed awarded the Oscar for Best Picture in 1951.

The weights that were given above are not guesses, but are obtained by a statistical technique known as regression, which can be used if the total value $V_i$ and the ratings $x_{ki}$ of each property for each work $i$ are known. The technique is...
informally dealt with in Appendix 2. Suffice it to say here that it makes it possible to "estimate" those weights, and tells us whether a weight can be considered as significantly different from zero, that is, whether the property that is associated with the weight contributes (is significantly positive) or does not contribute (is not significantly different from zero) to the overall value. It could happen that no property contributes to the overall value. This will be so if the properties do not correctly describe the works, or if judges set the overall values without taking into account any of the individual properties, thus running into contradictions, pointed out, for example, by Beardsley in the introduction of *Aesthetics*.

Before turning to the results of this procedure, it is useful to describe briefly how movies are selected by the National Academy.11 All eligible films are listed in a "reminder list of eligible releases," that is distributed to voters (active and life members of the Academy). Best picture nominees and awards are chosen by all Academy members, while nominations in other categories (acting, direction, screenplay, cinematography, art direction, costume design, music, film editing, sound mixing, special effects) are picked only by members of each discipline: only actors nominate actors, for example. Voting for nominations and awards are made by secret ballot. Five nominations at most can be made in each category, by each member who is allowed to nominate. There is thus a two-stage voting procedure. In the first stage, the five achievements in each category receiving the highest number of votes become the nominations in the second stage. This stage proceeds in the same way (by voting on the selection) and leads to the election of the winner. The two interesting characteristics are that both nominees and winners result from a voting procedure that includes a large number of movie experts (5,600 in 2003) and that only experts of the discipline choose the nominees of the discipline, while the Best Picture Oscar (overall value) is given by all the members of the Academy.

The results of the estimation procedure12 are reproduced in Table 2, where only the properties picking a positive weight that is significantly different from zero are given. These are the weights that were used in the comparison between *An American in Paris* and *A Streetcar Named Desire*. As can be seen, only five properties seem to matter, and only Oscars given for these properties matter. A nomination without Oscar has no significant impact on the overall value. Director, Screenplay and Costume Design carry the largest weights.

11 In what follows, we summarize the rules used in 2003. See http://www.oscars.org/76academyawards/rules.

12 Since the $V_j$ are variables that take only two possible values (1 for an Oscar and 0 for a nominated movie), the procedure used to estimate the weights is the so-called probit regression.
### Table 2
Properties with Positive and Significant Weights

<table>
<thead>
<tr>
<th>Property</th>
<th>Weight</th>
<th>Confidence interval (5 percent level)*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Oscar for Actor in a Leading Role</td>
<td>0.668</td>
<td>[-0.066--1.402]</td>
</tr>
<tr>
<td>Oscar for Director</td>
<td>2.342</td>
<td>[1.629--3.056]</td>
</tr>
<tr>
<td>Oscar for Screenplay</td>
<td>1.157</td>
<td>[0.521--1.793]</td>
</tr>
<tr>
<td>Oscar for Costume Design</td>
<td>1.059</td>
<td>[0.457--1.661]</td>
</tr>
<tr>
<td>Oscar for Film Editing</td>
<td>0.614</td>
<td>[-0.062--1.290]</td>
</tr>
<tr>
<td>R-square</td>
<td>0.678</td>
<td></td>
</tr>
<tr>
<td>Total Number of Movies</td>
<td>270</td>
<td></td>
</tr>
</tbody>
</table>

* See footnote 13.

The confidence intervals clearly show that three weights (Director, Screenplay, Costume) are bounded away from zero and are therefore significantly positive, while the two other (Actor and Film Editing) contain zero, but would not at a slightly larger level of probability than the one reported in Table 2.13 The R-square coefficient (whose value is zero if the adjustment is very poor as in Figure 1c of Appendix 2, and one if the adjustment is perfect as in Figure 1b) is equal to 0.678.

Using the weights that appear in Table 2, it is easy to calculate the overall value of a movie of which the properties have been valued. In Table 3, we do this for the four movies that were considered in Table 1. Their overall value is obtained by

\[ V = \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_4 x_4 + \beta_5 x_5, \]

where the \( \beta \) are the weights and the various \( x_k \) take the value 1 when the Oscar was attributed to the property, and the value 0 otherwise. All About Eve, for example, was awarded Oscars for Direction (weight: 2.342), Screenplay (1.157) and Costume Design (1.059). Its overall value is equal to 4.558. Similar calculations can be made for the three other movies, leading to the overall values in the last line of the table.

---

13 This 5 percent probability level can be explained intuitively as follows. If we had the possibility to carry out this calculation using 100 different samples of the same number of movies, then in 95 cases, the weight that is obtained would be contained in the (5 percent confidence) interval. One can also construct larger confidence intervals using larger probability levels, such as 10 percent.
The values have not much meaning in absolute terms, but allow to infer a complete ordering (ties are of course possible).

### Table 3
Rating the Overall Value for Four Movies

<table>
<thead>
<tr>
<th></th>
<th>Weights</th>
<th>All About Eve</th>
<th>An American in Paris</th>
<th>Cabaret</th>
<th>A Streetcar Named Desire</th>
</tr>
</thead>
<tbody>
<tr>
<td>Actor in a Leading Role</td>
<td>0.668</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Director</td>
<td>2.342</td>
<td>O</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Screenplay</td>
<td>1.157</td>
<td>O</td>
<td>O</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Costume Design</td>
<td>1.059</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>-</td>
</tr>
<tr>
<td>Film Editing</td>
<td>0.614</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Overall value</td>
<td>4.558</td>
<td>2.216</td>
<td>1.059</td>
<td>0.000</td>
<td></td>
</tr>
</tbody>
</table>

An O indicates that the movie was awarded the Oscar for this property.

Given the special configuration, it is also possible to use the ordering method suggested by Dickie, which turns out to give the same result as the weighting method. This is due to the very special structure of the case. To see that this would not always be possible, assume that *All About Eve* did not win the Oscar for Screenplay. Our method would still rank it first, since it would obtain an overall value of 3.401 (= 2.342+1.059), but, as is easy to check, Dickie’s method would make a complete ordering impossible, since Director and Costume Design for *All About Eve* are incommensurable with Screenplay and Costume Design for *An American in Paris*. Likewise, it would impossible to compare *An American in Paris* and *Cabaret*, if *An American* had not been awarded the Oscar for Costume Design. Our method would still rank *An American* first (value 1.157) and *Cabaret* second (value 1.059).

It is interesting to look at how this scheme does in predicting whether a movie with specific properties is correctly rated. The number of correct predictions, using the weights displayed in Table 2, is 252, that is the weighting scheme predicts whether a movie will be awarded the Oscar or not, by just looking at the rating of the properties, in more than 93 percent (252/270) of the cases. A more careful look at the wrong predictions is needed here. Indeed, two types of errors are possible.

First, in a given year, the movie that is given the Best Picture Oscar may not be the one that is rated highest using our technique. For example, in 1951, *A Place in the Sun* and *An American in Paris* are rated 5.17 and 2.22, respectively; according to our method, the first movie should have been given the Oscar, which the Academy awarded to the second one. Since in each of the 54 years (1950-2003) that we cover
only one movie in each year is awarded the Oscar for Best Picture, there are such 54 movies. Using our weighting scheme, we can rank these and check where the errors are located. Table 4 shows that our method is in disagreement with the Academy decision in nine out of the 54 years. It may also be interesting to point out that the American movie critic Leonard Maltin\textsuperscript{14} agrees with the choice made by the Academy for 1951 and 1981, agrees with us (that is, ranks our movie higher than the one which was awarded the Oscar) in 1952, 1956, 1989 and 2002, and ranks equally the two movies in 1972 and 1998 and 2000.

\textbf{Table 4}
\textit{Overall Rating Differences Between the Academy and the Weighting Method Year by Year}

\begin{tabular}{lll}
\hline
\textbf{Year} & \textbf{Received the Oscar} & \textbf{Should have received the Oscar} \\
\hline
1951 & An American in Paris (2.22) & A Place in the Sun (5.17) \\
1952 & The Greatest Show on Earth (1.16) & The Quiet Man (2.34) \\
1956 & Around the World in 80 Days (1.77) & Giant (2.34) \\
1972 & The Godfather (1.82) & Cabaret (2.96) \\
1981 & Chariots of Fire (2.22) & Reds (2.34) \\
1989 & Driving Miss Daisy (1.16) & Born on the Fourth of July (2.96) \\
1998 & Shakespeare in Love (2.22) & Saving Private Ryan (2.96) \\
2000 & Gladiator (1.73) & Traffic (4.11) \\
2002 & Chicago (1.67) & The Pianist (4.17) \\
\hline
\end{tabular}

Rates obtained by the weighted sum of properties, using the weights in Table 2, are given between brackets.

Second, the Best Picture Oscar is awarded every year. It may happen that all movies are poorly rated in some years, and that an Oscar is given to a movie that would not have deserved it had it been produced in an other year. It is difficult to predict whether \textit{Gandhi} (Best Picture Oscar in 1982) or \textit{Around the World in 80 Days} (Best Picture Oscar in 1952) would have obtained the Oscar had they been produced in the same year. Therefore, an indirect way must be found to check whether those movies that were awarded the Oscar are also those that are ranked highest, irrespective of time, and of the judging that is made only on movies that are produced during one year. This can be looked at by ranking the movies in decreasing order of overall value. Table 5 shows that Best Picture Oscars were awarded to nine highest ranked movies, to 19 out of the first 20, to 27 out of the first 30, etc. Thus, if the

choices according to the weighting method are right, the Academy missed one movie ranked among the top ten (and 20), three movies among the top 30 (and 40), etc. Only ten movies that are ranked worse than 50 and 6 movies that are ranked worse than 60 received the Best Picture Oscar. This illustrates that, even over time, there are few contradictions between what the Academy did, and what the overall rating based on properties does. But it also illustrates that in years in which all productions were poor, there was no reason to award the Best Picture Oscar.

**Table 5**

*Overall Rating Differences Between the Academy and the Weighting Method over Time*

<table>
<thead>
<tr>
<th>Ranks</th>
<th>Number of Oscars</th>
</tr>
</thead>
<tbody>
<tr>
<td>First 10</td>
<td>9</td>
</tr>
<tr>
<td>First 20</td>
<td>19</td>
</tr>
<tr>
<td>First 30</td>
<td>27</td>
</tr>
<tr>
<td>First 40</td>
<td>37</td>
</tr>
<tr>
<td>First 50</td>
<td>44</td>
</tr>
<tr>
<td>First 60</td>
<td>48</td>
</tr>
<tr>
<td>First 70</td>
<td>49</td>
</tr>
<tr>
<td>First 80</td>
<td>52</td>
</tr>
<tr>
<td>First 83</td>
<td>54</td>
</tr>
</tbody>
</table>

The method that was used makes it necessary to observe both the rating of properties and of the overall value. This may suggest that we are facing a circular reasoning: We need the overall value to predict it, and indeed, the "predictions" discussed above are "past predictions," made on movies that were used to estimate the weights. It would be more interesting to predict what will happen in the next years, and compare what the Academy did, and how the weighting method performs. This is of course possible once the weighting scheme has been made explicit. But more importantly, what we tried to show is that experts, here the members of the National Academy of Motion Picture Arts and Sciences, use an implicit weighting scheme of individual properties to determine the overall value, that such implicit weights can be recovered, and used to "predict" the overall value of movies, and thus their ranks, once properties have been rated. The partial ordering that obtains using Dickie's suggestion would make such an overall ranking impossible.
Appendix 1. Properties of Movies

The National Academy of Motion Picture Arts and Sciences attributes Oscars for a certain number of categories, characteristics or properties. Their number and denominations have undergone some (relatively minor) changes between 1950 and 2003. The Academy did, for instance, award Oscars for "cinematography-black and white" and "Cinematography-color" until 1966. Since 1967, there is only one such award, called "Cinematography." Sometimes the denominations and/or the number of categories for similar properties were changed. Until 1955, for example, there were three awards for screenplay: "Screenplay," "Story and screenplay," and "Motion picture story." Since 1956, only two categories remain: "Adapted screenplay," and "Original screenplay."

The following features were considered in 2003 (dates and possible changes are mentioned between brackets).

(1) (Best) Picture (1950-2003);
(2) Actor in a Leading Role (1950-2003);
(3) Actress in a Leading Role (1950-2003);
(4) Actor in a Supporting Role (1950-2003);
(5) Actress in a Supporting Role (1950-2003);
(6) Director (1950-2003, sometimes denominated "Art Director");
(7) Adapted Screenplay (1956-2003, sometimes denominated "Screenplay Based on Material from Another Medium");
(8) Original Screenplay (1956-2003, sometimes denominated "Story and Screenplay Written Directly for the Screen");
(9) Cinematography (2 Oscars until 1966, "black and white" and "color," except in 1957 where there was a unique Oscar for both);
(10) Art Direction/Set Decoration (2 Oscars until 1966, "black and white" and "color," except in 1957 and 1958 where there was a unique Oscar for both);
(11) Costume Design (2 Oscars until 1966, "black and white" and "color," except in 1957 and 1958 where there was a unique Oscar for both);
(12) Original Score (several categories appearing and disappearing according to circumstances, for dramatic score, comedy or musical);
(13) Original Song (1950-2003);
(14) Film Editing (1950-2003);
(15) Sound Mixing (denominated "Sound Recording" until 1957, "Sound" between 1957 and 2002);

---

15 Three Oscars until 1955: "Screenplay," Story and Screenplay," and "Motion Picture Story." There was also an Oscar for "Motion Picture Story" in 1956.
16 See footnote 11.
(16) Visual Effects (since 1963);\textsuperscript{17}
(17) Sound Editing (denominated "Sound Effects" between 1963 and 1967, dropped between 1968 and 1974, reinstalled as such in 1975 and 1977, denominated "Sound Effects Editing" or "Sound Editing" since 1981; both "Sound Effects" and "Sound Effect Editing" in 1977);\textsuperscript{18}
(18) Make-Up (in 1964, 1968, 1981, 1982 and every year since 1984);
(19) In 1961 and 1968, Oscars for "Dance Direction" was awarded.

Given the changes, the fact that some Oscars are recent, and some disappeared, we consider the following categories which are present in one form or another between 1950 and 2003: (1), (2), (3), (4), (5), (6), (7)+(8), (9), (10), (11), (12), (13), (14), (15), and (16)+(17). "Black and White" and "Color," present until 1966, were put in a unique category; so were "Adapted" and "Original Screenplays," all possible denominations for "Scores," as well as "Visual" and "Sound Effects or Editing." "Make-up" and "Dance Direction" were dropped since the first appeared only sporadically until 1984, and the second was awarded only twice.

\textsuperscript{17} Only one category "Special Effects" until 1962.
\textsuperscript{18} Only one category "Special Effects" until 1962.
Appendix 2. Regression Analysis

Regression analysis is concerned with the relation between one variable (say overall values, V) and a set of other variables (the x_k, here the scores on property k, k = 1, 2, 3), leading to a relation in which V is obtained as a weighted sum (the \( \beta \) are weights) of the scores of the properties:

\[
V = \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3.
\]

The V and the x_k are given, while the weights \( \beta_k \) will be determined through a calculation (or, as is often said, will be "estimated"). This is easy to explain if there is a single property, since the procedure can be represented graphically.

In Figures 1a-1c, a work is represented by a point (an observation). On the vertical axis one can read the overall value, on the horizontal one, the rate obtained for the unique property. In Figure 1a, the scatter of points leads us to conclude that there exists a positive, upward sloping relation between the two variables (the larger the score on the property, the larger the overall value of the work), that can approximately be represented by the line which goes through the scatter of points.\(^{20}\) If all the points were exactly on the line (as in Figure 1b), one could obtain the value of a work by simply reading its score on the property. This is the ideal case, since, more generally, the relation between the two variables will be more fuzzy, but if the scatter of points is reasonably flat in one direction and elongated in the other direction (as it is in Figure 1a), one can conclude that there exists a relation, and compute, using some criterion,\(^{21}\) the slope of the line. In Figure 1c, we also represent a scatter that does not have the same southwest-northeast pattern, and there is no obvious choice for the slope: it can take any direction. This underlines two extreme cases, the first in which all points lie on a line (a perfect adjustment) and the second, where any choice of slope is as good, or as bad as any other (almost every line is possible). The case illustrated in Figure 1a is intermediate, and is due to the fact that V is measured with some error or that the unique property does not explain value in a perfect way so that, instead of \( V = \beta_1 x_1 \), the relation should be written \( V = \beta_1 x_1 + u \), where \( u \) is a random disturbance, which originates from elements that we ignore.

These considerations lead us to define a coefficient which will measure the quality of the adjustment of the line to the scatter of points. This coefficient (called R-square) is defined in such a way that it will vary between one (perfect adjustment,

\(^{19}\)Note that this set may consist of a unique variable.

\(^{20}\)This implies that the equation is written \( V = \beta_1 x_1 + u \), where \( u \) is the distance between a point and the line.

\(^{21}\)Such as minimizing the sum of distances (or of squared distances) between the line and the points.
Figure 1b) and zero (any adjustment is possible, Figure 1c), while intermediate values will hold for cases such as the one in Figure 1a.

The slope of the line, $\beta_1$, is a number that can be estimated. It will come with two more numbers which describe an "interval" in which the slope can vary, a measure of the relative fuzziness with which the slope is estimated. A narrow interval will correspond to a good adjustment (the slope is equal to 0.20, but it can vary between 0.19 and 0.21, which is pretty accurate). A wide interval is the sign of a poor adjustment. For instance, if the calculated slope is equal to 0.20, but the interval goes from -0.30 to 0.70, then the direction of the line is not determined with much accuracy. It could be downward (negative slope) instead of upward (positive slope) sloping so that even if the estimated coefficient is equal to 0.20, there is some likelihood that it could also be equal to zero, since zero belongs to the interval [-0.30, 0.70]. If so, then the variable has little or no influence (a non significant influence) on $V$, and can thus be ignored.

This reasoning can easily (at least in mathematical terms, not in graphical ones) be extended to the case of a relation between $V$ and a set of variables, say $x_1$, $x_2$, and $x_3$. Then $V = \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + u$, where again, $u$ is a disturbance term due to non measured variables (for example $x_4$) or to some randomness, or measurement errors in $V$. For each variable of this set, one can determine a coefficient (a slope) together with the interval in which it can vary. If this interval is narrow, the variable has an influence on $V$. If it is large and contains zero it does not contribute to explaining $V$.

Dickie\(^{22}\) and Vermazen\(^{23}\) also evoke the situation in which some properties can interact within a work, making its overall value different from the "simple sum of the work's independently valuable properties." This sum will be larger if they interact positively and smaller if they interact negatively. The regression method that is suggested here makes it very easy to deal with such interactions also. For instance, if properties 1 and 2 are believed to interact, one can introduce a multiplicative term $\gamma x_1 x_2$ in the equation, where $\gamma$ is an additional weight to be estimated. Note that the interaction weight will affect in the same way all the works, and not, as assumed by Dickie and Vermazen, some works only, and maybe some in a positive way, others negatively.

Figure 1. Illustrating Regression Analysis