INVESTIGATING THE SUCCESS OF SCORING METHODS WHEN FACED WITH MANIPULATIVE JUDGING

By

MALIN ELISABETH RAPP-OLSSON

A Thesis Submitted to The Honors College

In Partial Fulfillment of the Bachelors Degree
With Honors in

Mathematics

THE UNIVERSITY OF ARIZONA

AUGUST 2013

Approved by:

Dr. Ning Hao
Department of Mathematics

The University of Arizona Electronic Theses and Dissertations Reproduction and Distribution Rights Form

The UA Campus Repository supports the dissemination and preservation of scholarship produced by University of Arizona faculty, researchers, and students. The University Library, in collaboration with the Honors College, has established a collection in the UA Campus Repository to share, archive, and preserve undergraduate Honors theses.

Theses that are submitted to the UA Campus Repository are available for public view. Submission of your thesis to the Repository provides an opportunity for you to showcase your work to graduate schools and future employers. It also allows for your work to be accessed by others in your discipline, enabling you to contribute to the knowledge base in your field. Your signature on this consent form will determine whether your thesis is included in the repository.

Name (Last, First, Middle)
Rapp-Olsson, Malin, Elisabeth

Degree title (eg BA, BS, BSE, BSB, BFA): B.S.

Honors area (eg Molecular and Cellular Biology, English, Studio Art): Mathematics

Date thesis submitted to Honors College:
08/07/2013

Title of Honors thesis: Investigating the success of Scoring Methods when Faced with Manipulative Judging

The University of Arizona Library Release Agreement

I hereby grant to the University of Arizona Library the nonexclusive worldwide right to reproduce and distribute my dissertation or thesis and abstract (herein, the "licensed materials"), in whole or in part, in any and all media of distribution and in any format in existence now or developed in the future. I represent and warrant to the University of Arizona that the licensed materials are my original work, that I am the sole owner of all rights in and to the licensed materials, and that none of the licensed materials infringe or violate the rights of others. I further represent that I have obtained all necessary rights to permit the University of Arizona Library to reproduce and distribute any nonpublic third party software necessary to access, display, run or print my dissertation or thesis. I acknowledge that University of Arizona Library may elect not to distribute my dissertation or thesis in digital format if, in its reasonable judgment, it believes all such rights have not been secured.

☒ Yes, make my thesis available in the UA Campus Repository!

Student signature: Malin Rapp Olsson Date: 08/07/2013

Thesis advisor signature: [signature] Date: 08/07/2013

☐ No, do not release my thesis to the UA Campus Repository.

Student signature: ______________ Date: ______________

Abstract

In judged competitions, there is always a chance that a judge is intentionally or unintentionally manipulating the scores, giving higher or lower scores to certain contestants based on their desired outcome. Is there a way to form composite scores for candidates which removes the influence of these manipulators? We will test three scoring rules associated with Olympic figure skating- The OBO Rule, the Median Rank Rule, and a New Rule proposed by Samuel Wu and Mark Yang (2004). By using Monte Carlo methods in R, we test these rules for sensitivity, accuracy, and the number of ties they result in. We test one case with no manipulative judges, and two cases with one manipulative judge each. The two manipulative judges employ different strategies to try to get the second best conestant to place first. We conclude that Wu and Yang's New Rule is most effective, based on its accuracy and rate of ties.

1 Introduction

A frequently addressed problem in statistics is that of subjective judged competitions. In particular, judged events can cause controversy when percieved unfair voting takes place, and there can be uncertainties about whether or not the best candidate won. One potential source of unfairness in subjective competitions is that there may be one or more judges who is giving scores not based on the performance of the contestants, but based on their own desire to see a particular contestant win or lose. In some cases, this biased behavior is blatant and purposeful, but in other cases we might see a judge acting slightly favorably towards certain contestants because of some underlying subtle preference. It is widely believed that in Olympic events, judges from neighboring countries tend to give each other's contestants higher scores than they give to other countries. In fact, Zitzewitz (2006) argues that figure skaters do benefit from having a compatriot on the judging panel. Zitzewitz claims that this is probably because of patriotic bias, as well as vote trading[4]. In some cases, these biases are subconscious. Is there a way to mitigate this type of biased judging, even when it may not be readily apparent? This paper will discuss three different methods of combining scores and how well each method minimizes the impact of a biased judge.

A basic judged competition, which scenarios in this paper will be dealing with, involves a panel of judges giving a score within some range to each contestant. As each contestant gets evaluated, their composite score is calculated, and this composite score is used to determine the final ranking of the contestants. The most simple and straightforward way to calculate a contestant's final score is by taking the mean of the scores. However, this has obvious flaws because each judge has the ability to drastically alter the mean. To rectify this, some competitions use a truncated mean. That is, they throw out the highest and lowest scores, and take the mean of the remaining values. This method is typically associated with Olympic events such as figure skating, though the rules for figure skating scoring have changed recently.

There are many other methods for combining scores. This paper will focus on three: The OBO Method, the Median Rank Method, and the New Method, each of which will be explained in a following section. The simulation includes different degrees of variation between the contestants, so the methods are compared when contestants are very similar and when they are more diverse in their ability. We will then see how the scoring methods fare when there is a manipulative judge on the panel. The number of accurate rankings will tell us which methods are most successful at mimimizing the impact of unfair judges.

2 Literary Review

The problem of fairness in judging is a fairly common topic in statistics, and has been approached from many different angles. This paper builds on the ideas presented in the following papers.

2.1 Previous Work

A 2010 paper, “Fairest of Them All: Using Variations of Beta-Binomial Distributions to Investigate Robust Scoring Methods”, was written by myself, Mary Good, Christopher Kinson, Mame Fatou Thiam, and Karen Nielsen related to this topic [2]. In this paper, the goal was to compare different scoring methods in order to recommend a scoring method best suited for judged competitions in which not all contestants are scored by the same judges. When contestants are judged by a different set of judges, it is difficult to compare their scores directly. By simulating different types of judging behavior, we attempted to find out if there is a way to calculate the final scores for contestants which makes the final rankings as consistent as possible regardless of the personality of the judges. In the simulation, the scores given by a standard judge are drawn from a binomial distribution, whose parameter is the contestant value drawn from a left-skewed Beta distribution. The parameters of the binomial distribution are altered to represent different judging behavior.

We tested the scoring methods for different judging panels, and found that the most accurate scoring method depends on the composition of the judging panel. If the judges are relatively similar to each other in their scoring behavior, then the most accurate scoring method was the mean. However, if the judges behave relatively different from each other, it is better to use the “z-rank” method of scoring, which is a method of calculating the z-scores for the scores given by each judge, and then ranking those z-scores and averaging the ranks for each contestant [2].

This paper will expand on the idea of having different panels of judges, but instead of simulating judges who always give higher or lower scores than the standard judge, we will focus on single manipulative judges, whose motivation is simply to promote one player over all the others.

2.2 Other papers

In the 2006 paper ”Nationalism in Winter Sports Judging and Its Lessons for Organizational Decision Making” Zitzewitz explores the voting patterns of judges in Olympic figure skating and Olympic ski jumping [4]. He argues that figure skating judges appear to engage in vote trading and block voting. Unlike ski jumping, the figure skating scoring system does nothing to compensate for national bias. Furthermore, ski jumping judges have an interest in remaining fair, because they are chosen by a central body, whereas figure skating judges benefit from judging along national lines because they are chosen by the national federations. He finds that the national bias results in 0.7 placements higher, on average. Zitzewitz claims that the scoring system in figure skating actually encourages manipulative voting because the truncation of extreme scores make vote trading easier to implement [4].

In his 2012 paper "Does Transparency Really Increase Corruption? Evidence from the 'Reform' of Figure Skating Judging", Zitzewitz argues that after the scoring system in Olympic figure skating became anonymous, the rate of biased scoring actually increased [5]. Though the International Skating Union believed that a new voting system in which the identity of the judges was concealed would decrease vote trading and other unethical behavior, the anonimity actually increased the amount of favor given to contestants from neighboring countries. He argues that the ISU's proposed method, where only some of the scores are randomly selected to count, does not decrease the overall amount of vote manipulation in figure skating. This would suggest that the scoring method proposed by the ISU can be improved upon.

Wu and Yang compare this ISU rule, known as the "One By One" rule, to the Median Rank rule and their own proposed New Rule [3]. These rules are tested in cases where there are one and two manipulators. They find that for three contestants and nine judges, their New Rule fares slightly better than both the OBO rule and the Median Rank. Wu and Yang make the case that the accuracy of a method is not the only factor to consider when comparing scoring methods, but that avoindance of ties is also to be desired. This paper will expand on Wu and Yangs analysis of manipulative judging, using more contestants and two different voting patterns by manipulators.

3 Judging behaviors in real competitions

The behavior of the judges is a key component to simulating a competition. In real competitions, is it common to see judges who score very differently from each other? Though there are countless examples of sports and other situations where this type of judging is applicable, this discussion will mosly focus on Olympic figure skating, as figure skating is somewhat notorious for its judging, with the scoring providing several controversies over the years. Zitzewitz claims that "while biases are often difficult to detect and predict in real-world settings, nationalistic bias in figure skating is large and predictable," [5] making figure skating a prime choice for observing bias in judging.

Analyzing actual scores from past competitions provides some insight into how a typical outcome might look. A sample of data from Olympic figure skating qualifiers in 2002 is analyzed in order to observe the scoring behavior of a handful of judges[1]. See Table 6.1 in the Appendix for the mean and median score given by each judge, and Table 6.2 for the complete data sample. From the graphs of the judges' scores (Figures 6.1 - 6.7), we can see that all scores are in the top half of possible scores. We also see that though the linear fit for each set of data is similar, there is some variation between the judges. The scores are generally more similar among the top contestants.

4 Methods

4.1 Simulation

In order to compare the scoring methods, we use a simulation in the statistical programming language R. In this simulation we model a simple competition with five contestants being scored by nine judges. The contestants in this scenario recieve scores from 0.0 to 6.0, with 6.0 being the best. These scores are referred to as the **raw scores**. Each time that scores are assigned to contestants is referred to as a new **trial**. The scoring rules are applied to the raw scores to determine a final ranking of the contestants. A large number of these trials are run using Monte Carlo repetitions, and we count the number of times each scoring method succeeds.

4.2 Modeling the Contestants and Scores

It is relatively uninteresting to see what happens when contestants have a very large difference in ability. The most controversial cases of judging happen when the field is very competitive with only slight difference in ability. Furthermore, based on our observations of actual scores from Olympic qualifiers, the range of given scores is relatively small. Therefore, we wish to model scenarios where there is very little variation between the contestants.

Though judged competitions are subjective, we assign an assumed "correct" ranking, in which Contestant 1 should get first place, Contestant 2 should get second place, etc. Our assumption that there is a correct ranking is necessary in order to quantify the performance of our scoring algorithms. Therefore, Contestant i has a **contestant value** c_i, which represents his ability. Note that Contestant 1 always refers to the best contestant, Contestant 2 to the second best, and so on.

In order to see how sensitive the scoring methods are, we run the simulation multiple times, making the constestants slightly further apart in ability each time.

This allows us to see what happens when the contestants are nearly identical (and thus harder to rank accurately) and when the difference between contestants qualities is greater. We use δ as our sensitivity, where δ is in the range $(0.001, .05)$. Each time we run the simulation, we increases δ by increments of .001, so the simulation is run for 50 different sensitivities in total.

Along with varying the sensitivity, we also want to account for different variations of ability among the contestants. We model four cases as follows:

Case 1: c = ($0.9 + \delta$, $0.9 + 0.5\delta$, 0.9, $0.9 - 0.5\delta$, $0.9 - \delta$).
Case 2: c = ($0.9 + \delta$, $0.9 + 0.5\delta$, 0.9, $0.9 - 2\delta$, $0.9 - 3\delta$).
Case 3: c = ($0.9 + 2\delta$, $0.9 + 1.5\delta$, 0.9, $0.9 - 0.5\delta$, $0.9 - \delta$).
Case 4: c = ($0.9 + 0.5\delta$, $0.9 + 0.25\delta$, 0.9, $0.9 - 0.25\delta$, $0.9 - 0.5\delta$).

Case 1 is a standard case, where the contestant values are evenly spread with a difference of 0.5δ between each c_i. In Case 2, the bottom two contestant values are further from the top three contestants, with a differences of 2δ and 3δ from the median. In Case 3, the top two contestants are further from the bottom three contestants, with differences of 2δ and 1.5δ from the median. And finally, in Case 4, we have an even spread but with only 0.25δ difference between the contestant values.

For example, the following table shows the contestant values when $\delta = 0.001$.

Table 4.1: *Contestant Values for* $\delta = 0.001$

Case	c_1	c_2	c_3	c_4	c_5
Case 1	0.901	0.9005	0.9	0.8995	0.899
Case 2	0.901	0.9005	0.9	0.898	0.897
Case 3	0.902	0.9015	0.9	0.8995	0.899
Case 4	0.9005	0.90025	0.9	0.89975	0.8995

4.3 Scoring Methods

Raw scores are drawn from a binomial random variable with c_i as its parameter. We then create a rank matrix, in which we rank the scores given by each judge so that the highest raw score given by each judge is given a 1, the second highest a 2, and so on. Ties are allowed in this matrix. The scoring methods operate on this rank matrix as described by the particular rule, in order to determine a final ranking of the contestants.

4.3.1 Median Rank Method

The median rank method is relatively common in sporting events and has been used in the past in Olympic figure skating. In this method, the contestants are ordered by m, their median rank. If contestants are tied, we order them by the number of ranks that are higher than their median rank. If the contestants are still tied, they are ordered by the sum of all the ranks given to them.

4.3.2 OBO Method

The rule used by the International Skating Union since 2002 is called the "One By One" rule, or OBO for short. In the OBO scoring system, we calculate each contestant's ω, where ω is the number of pairwise wins against other contestants in the rank matrix. The OBO rule ranks the contestants by their ω, so the contestant with the largest number of pairwise wins gets first place in the competition. In the case of a tie, we order the tied contestants by the sum of their ranks.

4.3.3 New Method

The "New Method", proposed by Wu and Yang, first orders the contestants by ω, their number of pairwise wins. In the case of a tie, the lower median rank wins. If there is still a tie, we order them by the number of ranks larger than their median rank. Then, if there still is a tie, calculate p, the sum of the best n ranks, where n is half the total number of ranks, rounded up to the nearest integer. Ties are then resolved by the total sum of the ranks, and finally, if there is still a tie, order the tied contestants based on the pairwise winner.

4.4 Introducing Manipulators

By using a large number of trials, the simulation gives us the number of successes (as defined below) for each scoring method and thus lets us compare the effectiveness of each scoring method. However, we still have not addressed the case of manipulators. We will use two types of manipulators, each of which has a different strategy to get the second-best contestant to win.

The **Type A** manipulator wants Contestant 2 to win, and will place Contestant 1 in the third spot. Contestant 3 then gets bumped up to second place.

The **Type B** manipulator puts Contestant 2 in first place, and places Contestant 1 in fifth place, thus bumping up Contestants 5, 4, and 3 by one spot.

In our simulation, we check that Contestant 2 has a lower raw score than Contestant 1 before beginning to switch the scores. Note that in both cases, the manipulator attempts to not only raise Contestant 2's score, but also sabotage Contestant 1 by ranking them lower than Contestant 2's original placement. A Type A manipulation is more subtle, whereas a Type B manipulation is very noticeable and potentially controversial, as it places the best contestant in last place.

4.5 Determining a success

There can be several different ways of determining the success of an outcome. In some cases, such as awarding scholarships, it only matters that the top contestant recieves the correct ranking. In other cases, such as in qualifying rounds of an event, it only matters that the top n contestants get the top n spots, but their order within those ranks does not matter. And finally, there are cases in which the ordering of all the contestants must be correct in order for the scoring to be considered a success. In this case, in keeping with the Olympic scenario, we consider a method a success if all three top finishes are correct, and in the correct order.

5 Results

Figure 5.1 shows the sensitivity of the three scoring methods when there are no manipulators. Notice that in Case 2 and Case 3, all the scoring methods perform better than in Case 1, which can be attributed to the fact that the c_i values are further apart, so it is easier to distinguish between the contestants. The graph of Case 4 shows much lower probability of getting a successful ranking, even in the best case ($\delta = 0.05$) because the contestants are so similar. In all cases, the Median Rank is the least successful, and the success rates for the OBO Method and New Method are very similar.

Figure 5.2 shows the sensitivity of the three scoring methods when there is one Type A manipulator. Again, the OBO Method and the New Method seem more successful than the Median Rank. As we would expect, the probability of success is lower when there is a Type A manipulator than when there are no manipulators.

Figure 5.3 shows the sensitivity of the three scoring methods when there is one Type B manipulator. The Median Rank is once again the least successful, and the OBO Method and New Method very similar to each other. The probability of a success is generally slightly higher than for a Type A manipulator.

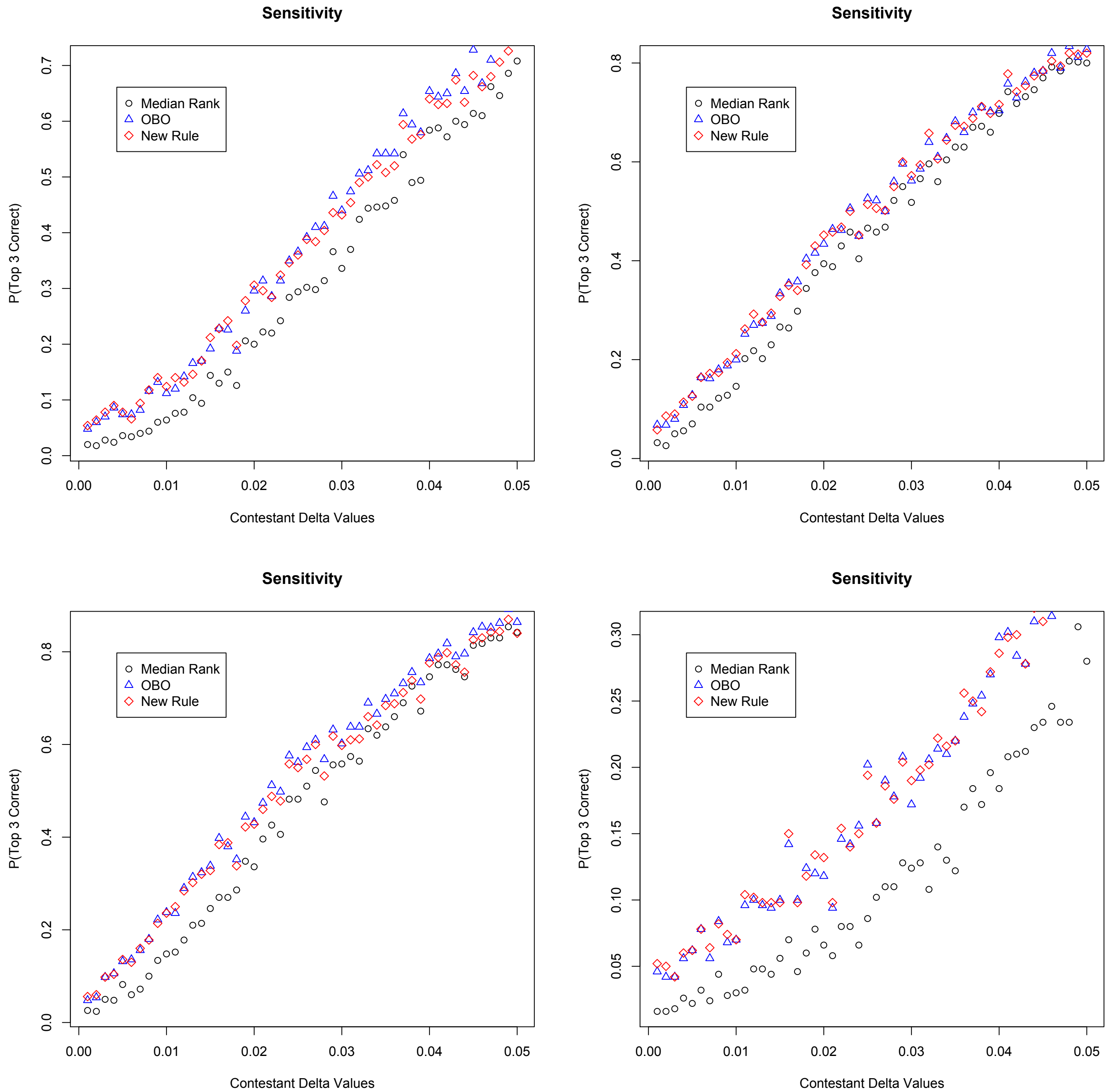

Figure 5.1: *The probability of a success using each scoring method when there are no manipulators. Top left corresponds with Case 1, top right is Case 2, bottom left is Case 3, and bottom right is Case 4.*

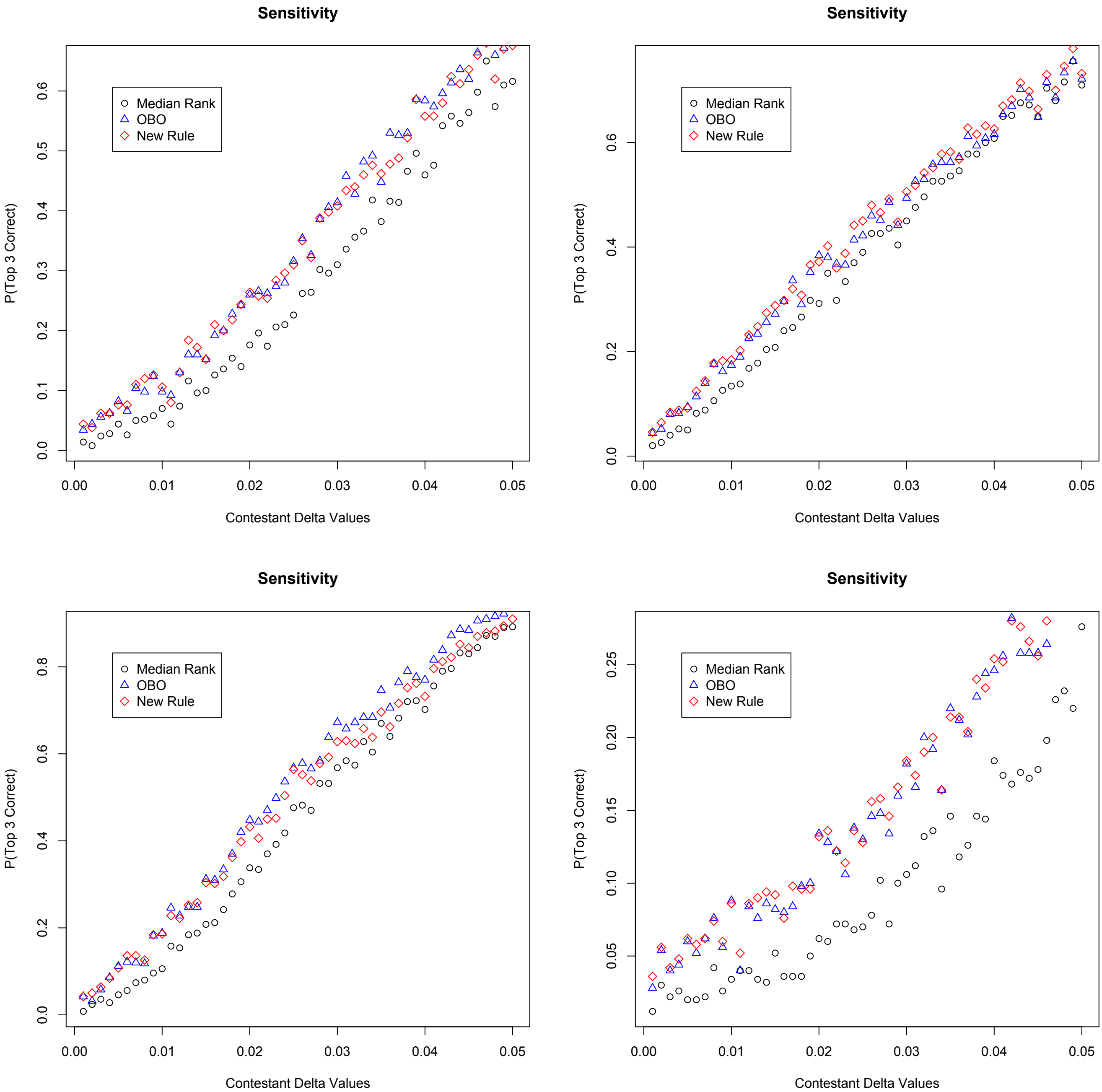

Figure 5.2: *The probability of a success using each scoring method when there is one Type A manipulator. Top left corresponds with Case 1, top right is Case 2, bottom left is Case 3, and bottom right is Case 4.*

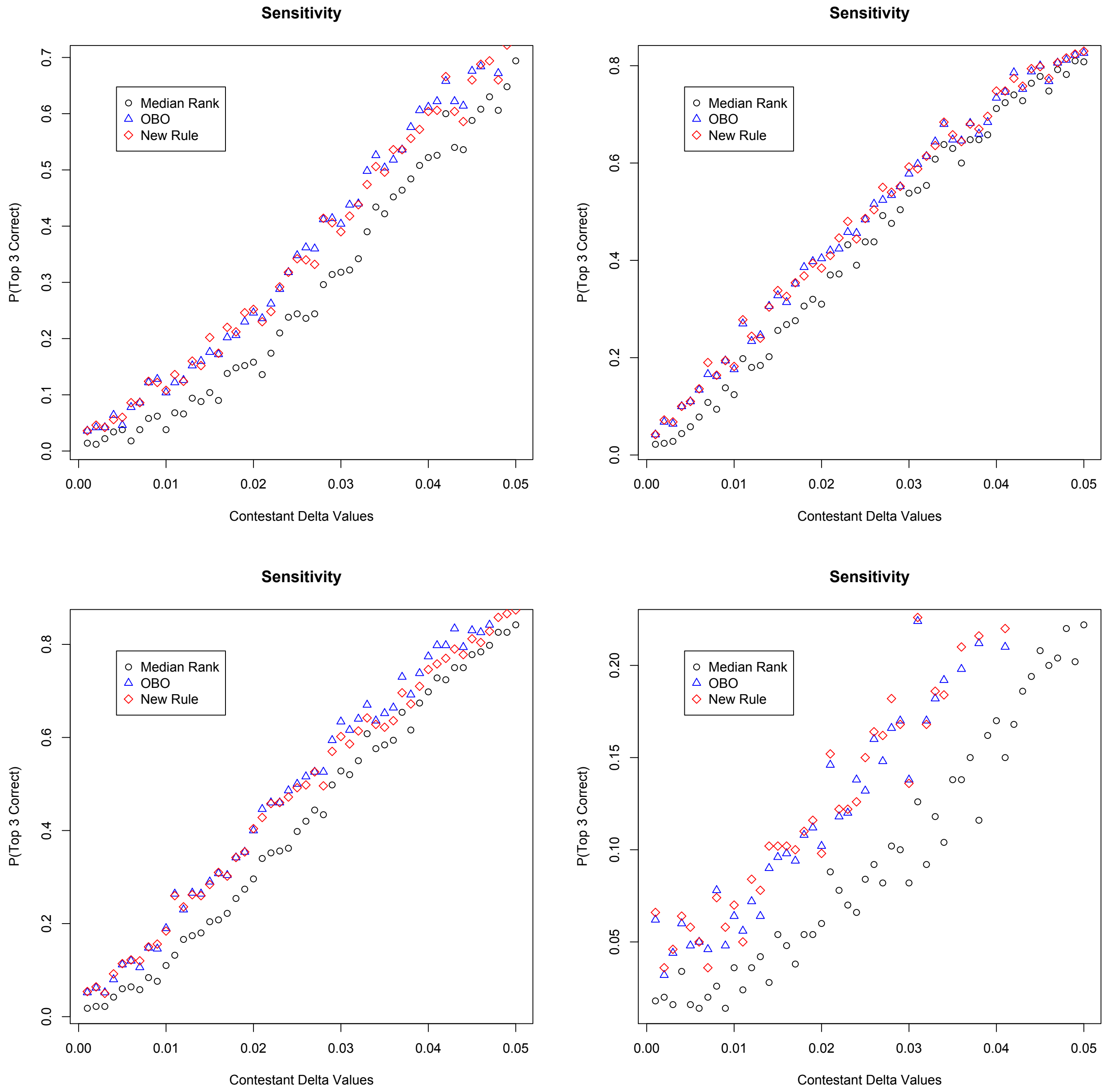

Figure 5.3: *The probability of a success using each scoring method when there is one Type-B manipulator. Top left corresponds with Case 1, top right is Case 2, bottom left is Case 3, and bottom right is Case 4.*

Tables 5.1, 5.2 and 5.3 show the total number of successes when there are no manipulators, a Type A manipulator, and a Type B manipulator. The highest number of successes for each case has been bolded. The total number of trials is $25,000$.

In Case 1 and Case 3, the OBO rule has slightly more successes than the New Rule. This is true whether there is a manipulator or not. Case 2 and Case 4 fare slightly better under the New Rule. The difference between the New Rule and the OBO Rule is very small.

For all cases, the highest number of successes occurs when there are no manipulators. The lowest number of successes occur when there is a Type A manipulator.

Table 5.1: *Total Successes With No Manipulators*

Case	Median Rank	OBO	New Rule
Case 1	7,766	**9,569**	9,379
Case 2	11,135	12,209	**12,239**
Case 3	11,327	**12,960**	12,626
Case 4	2,838	4,470	**4,532**

Table 5.2: *Total Successes With One Type-A Manipulator*

Case	Median Rank	OBO	New Rule
Case 1	6,913	**8,683**	8,581
Case 2	9,591	10,476	**10,743**
Case 3	11,147	**12,960**	12,443
Case 4	2,416	3,852	**3,952**

Table 5.3: *Total Successes With One Type-B Manipulator*

Case	Median Rank	OBO	New Rule
Case 1	7,083	**8,991**	8,862
Case 2	10,791	11,948	**12,016**
Case 3	10,339	**12,228**	11,906
Case 4	2,382	3,983	**4,031**

In addition to counting the number of successes, we also want to know how many times there is a tie in the final rankings. This is shown in Table 5.4. See Tables 6.2-6.5 in the Appendix for details on the number of ties in each case. The OBO method results in the most ties, with the Median Rank method having slightly fewer. The New Rule has a much lower rate than either of the other two methods.

Table 5.4: *Rate of Ties*

Case	Median Rank	OBO	New Rule
No Manipulator	12.2%	15.3%	4.0%
Type-A	12.6%	16.1%	4.2%
Type-B	12.8%	15.9%	4.2%
All	**12.5%**	**15.7%**	**4.1%**

6 Discussion

We found that including one manipulator of either Type A or Type B does not affect which scoring method to recommend. Based on the number of successes of the methods, the Median Rank scoring method does not perform better than the other two rules. This is true for all sensitivities we tested and for all of our contestant cases. The number of successes for the New Method and the OBO Method are too close to be able to judge which method is preferable.

As expected, all methods produce the most successes when there are no manipulators. More interestingly, the methods produce the fewest successes when a manipulator of type A is used. A judge who wants to make Contestant 2 win might think that giving the front-runner the worst possible score would decrease their composite score, giving Contestant 2 a better chance of winning. However, all three of these scoring methods minimize the impact of outliers. The Median Rank method is first interested in the median, so one extreme value (such as a very low score for Contestant 1) has very little impact on the final ranking. In the OBO Method and the New Method, the first comparison is of pairwise wins. Again, one extreme value will not have very much impact on a contestant's composite rank. When one of these rules is in effect, it seems that making smaller changes to the rankings (like the Type A manipulator) is a more effective manipulation tactic.

Of course, the number of successes is not the only important factor in deciding which method to use. It may be more desirable in some situations to use a simpler method. In sports especially, it can be useful to have a method that can be understood by the viewers, so as to avoid confusion and controversy. Another important factor in deciding on a scoring method is the number of ties it produces. A method that produces many ties is not an effective scoring method because it is failing to properly rank the contestants. We found that the New Rule produces much fewer ties than the other scoring methods. For this reason, the New Rule seems to be the best choice out of these three scoring methods.

In this paper, we expand Wu and Yang's number of contestants used from 3 to 5. In future, a similar simulation could be run using more contestants, to see if manipulation is more impactful when there are more contestants. Additionally, we would like to test for having two or three manipulators at once. We only have two types of manipulators, and both want to see Contestant 2 win. It would be interesting to see a simulation with manipulators with different strategies, to see which manipulation technique is the most effective. We agree with Wu and Yang's conclusion that their proposed rule is better than the OBO Rule and the Median Rank Rule, and hope that organizations like the ISU continue to revisit their scoring systems in order to make them stronger against manipulators.

References

[1] IceCalc. 2002 Men's Qualifying Results. IceCalc, 2002, http://www.icecalc.com/events/owg2002/results

[2] Mary Good, Christopher Kinson, Karen Nielsen, Malin Rapp-Olsson, Mame Fatou Thiam. (2010). Fairest of Them All: Using Variations of Beta-Binomial Distributions to Investigate Robust Scoring Methods. SUMSRI Journal, 2010, http://www.units.muohio.edu/sumsri/sumj/2010/Fairestststats.pdf

[3] Samuel S Wu, Mark C. K Yang. (2004). Evaluation of the Current Decision Rule in Figure Skating and Possible Improvements. The American Statistician, 58:1, 46-54.

[4] Zitzewitz, E. (2006). Nationalism in Winter Sports Judging and Its Lessons for Organizational Decision Making. Journal of Economics & Management Strategy, 15: 6799.

[5] Zitzewitz, E. (2012). Does Transparency Reduce Favoritism and Corruption? Evidence from the Reform of Figure Skating Judging. Journal of Sports Economics, May 2012.

Appendix

Table 6.1: *Scores From Olympic Qualifiers*

Judge	Mean Score Given	Median Score Given
Judge 1	10.0	10.1
Judge 2	10.1	10.1
Judge 3	9.9	10.1
Judge 4	10.0	10.2
Judge 5	9.6	9.9
Judge 6	9.9	9.8
Judge 7	9.6	9.8

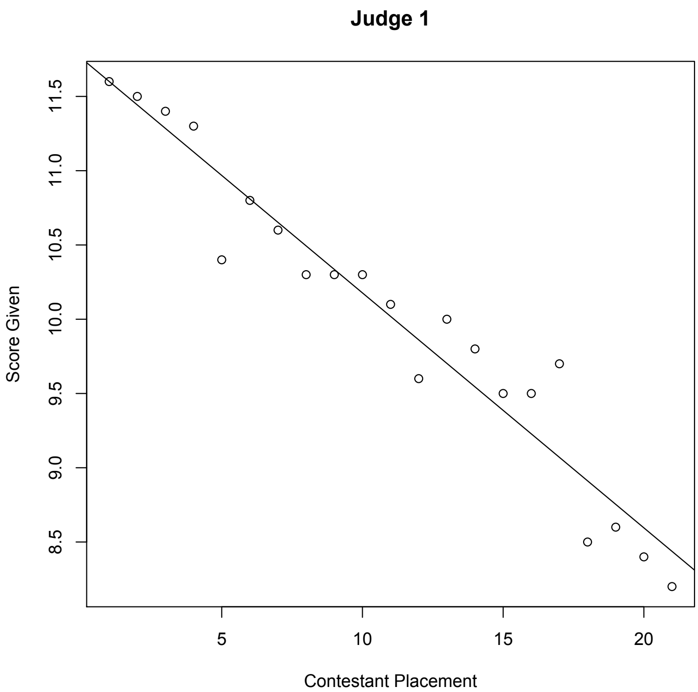

Figure 6.1: *Actual scores given by a judge in a 2002 Olympic qualifying event*

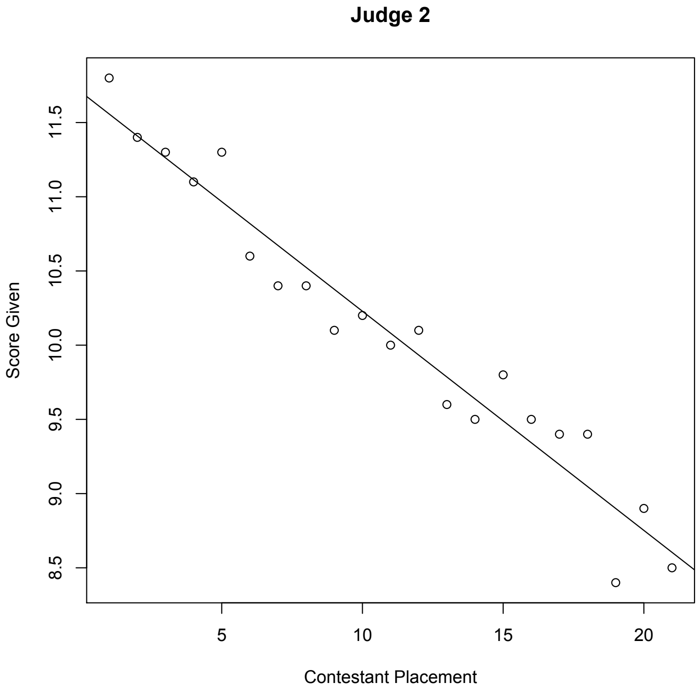

Figure 6.2: *Actual scores given by a judge in a 2002 Olympic qualifying event*

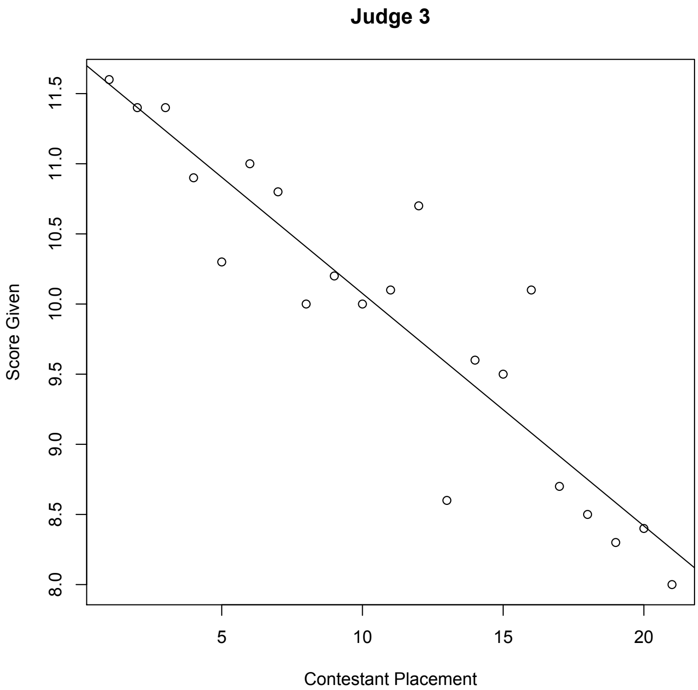

Figure 6.3: *Actual scores given by a judge in a 2002 Olympic qualifying event*

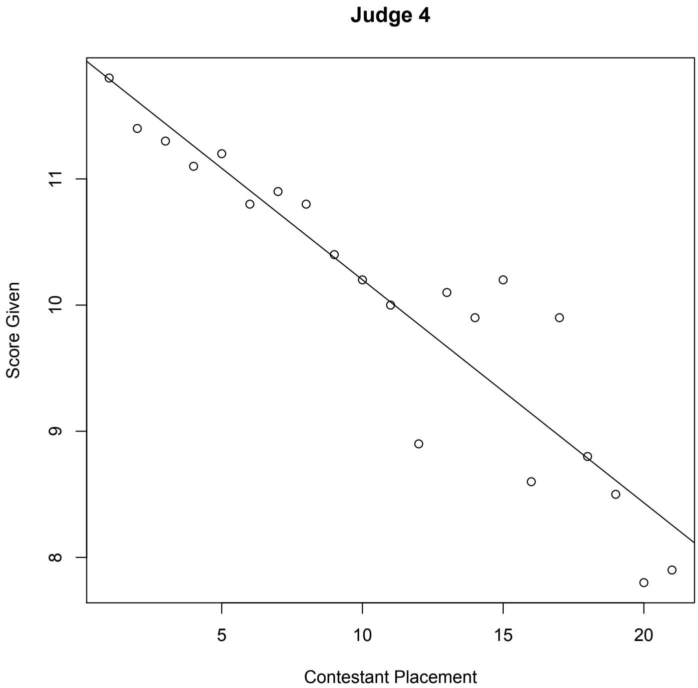

Figure 6.4: *Actual scores given by a judge in a 2002 Olympic qualifying event*

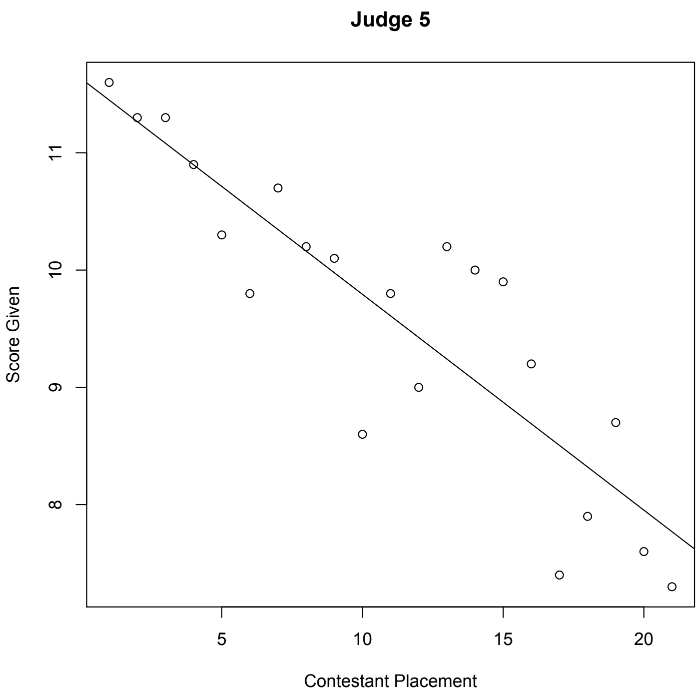

Figure 6.5: *Actual scores given by a judge in a 2002 Olympic qualifying event*

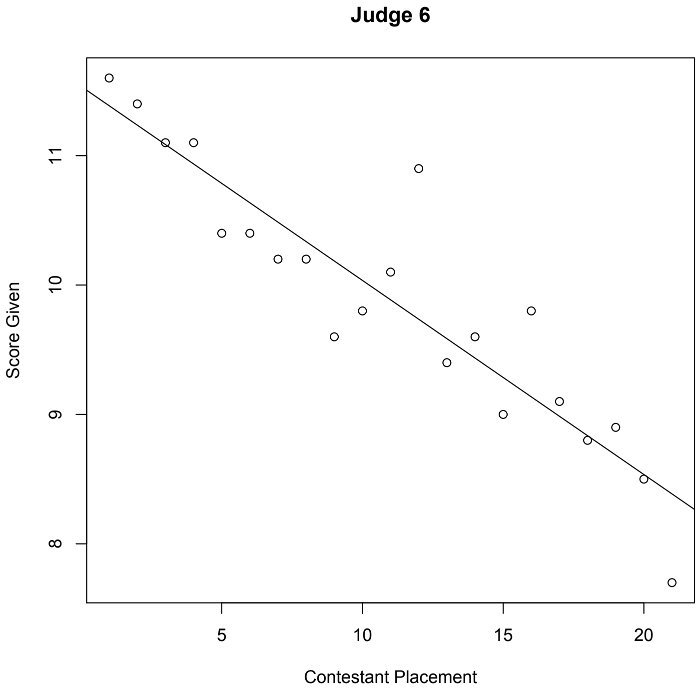

Figure 6.6: *Actual scores given by a judge in a 2002 Olympic qualifying event*

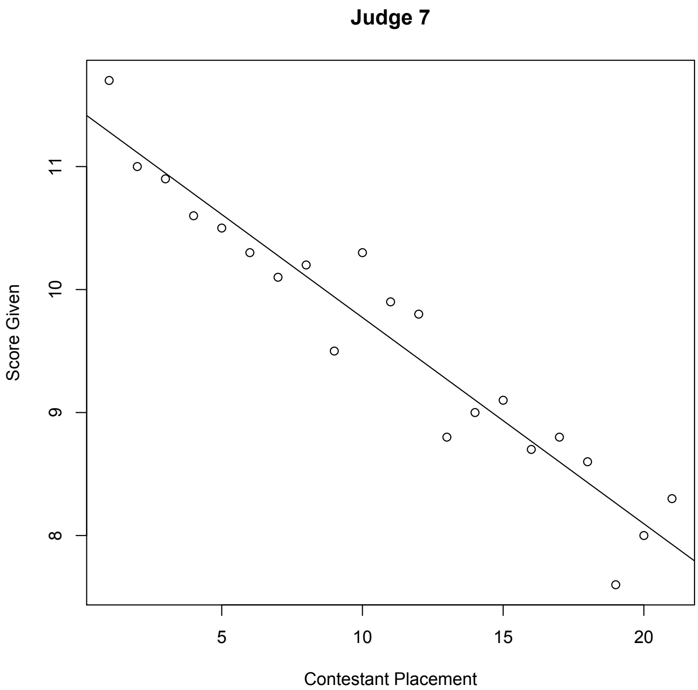

Figure 6.7: *Actual scores given by a judge in a 2002 Olympic qualifying event*

Table 6.2: *Raw Scores From Olympic Qualifiers*

Judge 1	11.6	11.5	11.4	11.3	10.4	10.8	10.6	10.3	10.3	10.3	10.1	9.6	10.0	9.8	9.5	9.5	9.7	8.5	8.6	8.4	8.2
Judge 2	11.8	11.4	11.3	11.1	11.3	10.6	10.4	10.4	10.1	10.2	10.0	10.1	9.6	9.5	9.8	9.5	9.4	9.4	8.4	8.9	8.5
Judge 3	11.6	11.4	11.4	10.9	10.3	11.0	10.8	10.0	10.2	10.0	10.1	10.7	8.6	9.6	9.5	10.1	8.7	8.5	8.3	8.4	8.0
Judge 4	11.8	11.4	11.3	11.1	11.2	10.8	10.9	10.8	10.4	10.2	10.0	8.9	10.1	9.9	10.2	8.6	9.9	8.8	8.5	7.8	7.9
Judge 5	11.6	11.3	11.3	10.9	10.3	9.8	10.7	10.2	10.1	8.6	9.8	9.0	10.2	10.0	9.9	9.2	7.4	7.9	8.7	7.6	7.3
Judge 6	11.6	11.4	11.1	11.1	10.4	10.4	10.2	10.2	9.6	9.8	10.1	10.9	9.4	9.6	9.0	9.8	9.1	8.8	8.9	8.5	7.7
Judge 7	11.7	11.0	10.9	10.6	10.5	10.3	10.1	10.2	9.5	10.3	9.9	9.8	8.8	9.0	9.1	8.7	8.8	8.6	7.6	8.0	8.3

Table 6.3: *Number of Ties With No Manipulators*

Case	Median Rank	OBO	New Rule
Case 1	3109	3926	909
Case 2	2188	3154	814
Case 3	2425	3395	1002
Case 4	4471	4778	1250

Table 6.4: *Number of Ties With One Type-A Manipulator*

Case	Median Rank	OBO	New Rule
Case 1	3188	4029	937
Case 2	2561	3453	989
Case 3	2384	3477	983
Case 4	4473	5080	1340

Table 6.5: *Number of Ties With One Type-B Manipulator*

Case	Median Rank	OBO	New Rule
Case 1	3210	4069	955
Case 2	2322	3165	867
Case 3	2581	3531	1027
Case 4	4703	5099	1374