

## INFORMATION ITEM

**DATE:** June 30, 2021  
**TO:** Technical Advisory Committee  
**PREPARED BY:** Joe Barbeau, Senior Planner (651-602-1705)  
Steve Peterson, Manager of Highway Planning and TAC/TAB  
Process (651-602-1819)  
Elaine Koutsoukos, TAB Coordinator (651-602-1717)  
**SUBJECT:** 2022 Regional Solicitation: Scoring Measure Outlier Adjustments

Many scoring measures are scored by awarding full points to the top-scoring application and distributing points proportionately to other applications. This sometimes results in one application scoring 100% of the points, while all others score less than 50% or even less than 10% of the points, the latter in particular providing almost no separation among any applications other than the top-scorer. Scoring committees have corrected for that by adjusting the remaining applications to the second highest-scoring application. This tends to improve the overall spread among applications, though arguably at the expense of a deserved advantage of the top-performing application.

There is no definition of what constitutes an outlier, no standard for when an adjustment is needed, and no standard for how to adjust. Up until now, scorers and committees have relied on “knowing an outlier when we see it” and have tended to adjust projects proportionately to the second-ranked application. Partners have provided feedback that this should be standardized. In 2020, there was even a scoring challenge suggesting that an outlier adjustment should have occurred on a measure.<sup>1</sup>

The purpose for adjusting low-scoring applications within a measure is not to improve scores; it helps create separation among applications. The standards that would have to be set are:

1. What is an outlier (or outliers) and when is an adjustment needed? Key discussion points:
  - a. Is a minimum number of total applications needed for an outlier adjustment to be used?
  - b. Would a minimum proportion of applications (e.g., 50%) need to be below a certain number of points (e.g., 20% of the maximum) to necessitate an adjustment?
  - c. How should the need for an adjustment be determined if there are multiple high-scoring applications (e.g., three applications score 90% and above with the remaining eight applications scoring 10% and below)?
  - d. Is it likely that adherence to a strict standard could tie the hands of a scoring committee, preventing it from adjusting when it makes sense or forcing it to adjust when it does not make sense?
2. How should an adjustment be made? Assuming this should be standardized, key discussion points include:
  - a. Is adjusting to the second-ranked application the best method? Sometimes this creates a spread, but other times it would still leave most applications with almost

---

<sup>1</sup> The score was not changed, as Funding & Programming determined that the scorer and scoring committee did not have an obligation to adjust for an outlier since there are no standards.

no spread. It could also effectively eliminate the top-ranked application's advantage.

- b. How (or should) scoring committees keep the top-performing application from losing too much of the advantage it has earned? For example, a nominal total above 100% could be granted for top-performing outliers, which would partly offset the lost advantage. Another idea could be increasing that number above 100% based on its previous margin over the top adjusted application.

An example outlier policy: For quantitative proportionate scoring measures with at least eight applications, when the second-place application scores less than 50% of the total points possible, the scores should be adjusted to the second-highest scoring application.

Another approach could be to set a minimum (e.g., transit ridership, ADT) for getting the maximum points, as was suggested internally to address high ridership numbers on systemwide transit projects. The downsides to this approach are that it would need to be customized for each measure and would need to be determined prior to receipt of applications (i.e., the total counts would be unknown).

The following pages provide case studies of outliers used or not used in recent Regional Solicitations. Scenarios 2 and 3 used the same adjustment, but from different starting points and may provide evidence that a standard is needed, as neither scenario probably should have had an adjustment. Scenario 3 solved a problem that didn't exist and while there was an outlier in Scenario 2, the adjustment did not help address it.

If members view the standards as too strict and unknown, the policy for scoring committees to use outliers when and how they see fit could be maintained. In that case, use or lack thereof of an adjustment should not be challengeable.

#### Funding & Programming Committee Comments

At the June 17, 2021 meeting, members shared several thoughts:

- While it is valuable to make sure that the measures are achieving their task of separating projects, adjustments tend to harm the highest-rated project. This is a difficult balance.
- There was split opinion on whether to trust the judgment of the committees versus trying to prescribe when and how to adjust. Unforeseen circumstances could occur in advance to the scoring.
- Generally, members favored an adjustment when no application scores 50% of the top-scoring project, by adjusting it to 50% to 80% of the top score (as a single number or a range). There was some interest in mitigating the advantage lost by the top application.
  - However, Scenario 2 shown below is not an effective adjustment, in that the spacing among most projects does not change a lot. The adjustment cost the top-scoring application much of its advantage, greatly rewarded the second-rated application, and was fairly neutral towards the rest.
- An outlier adjustment could be viewed as a last resort.
- A log scale could be used.

## 1. 2016 Un-Addressed Outliers

Regional Solicitation: 2016

The 2016 Regional Solicitation included several measures for which scoring committee members felt that one or more outlier applications impacted a scoring measure's effectiveness. Staff identified the below potential outliers.

Measures with potential outlier concerns included the following characteristics:

- Roadway Expansion Measure Performance (21 applications submitted)
  - 1C. Top application scored 50. Others scored from 0 to 23.
  - 5B. Top application scored 50. Second-ranked application scored 23. Others from 0 to 8.
  - 6. Top application scored 150. Other 20 applications scored from 0 to 55.
  - 9. Top application scored 100. Other 20 applications scored from 10-48.
- Roadway Reconstruction / Modernization (34 applications)
  - Measure 2B. Top application scored 65. Others scored from 4 to 31.
  - Measure 5B. Top application scored 30. Second-ranked application scored 18. Others scored from 0 to 2.
- Roadway System Management (4 applications)
  - Measure 1C. Top application scored 30. Others scored 10, 5, and 6.
  - Measure 5A. Top application scored 150. Others scored 28, 15, and 0.
  - Measure 5B. Top application scored 50. Others scored 17, 9, and 0
  - Measure 6 Top application scored 200. Others scored 88, 0 and 0.
- Bridges Measure Performance (8 applications submitted)
  - Measure 1B. Top two applications scored 30. Others scored from 0 to 8.
- Safe Routes to School (3 applications)
  - Measure 2A. Top application scored 170. Others scored from 31 to 46.
  - Measure 6. Top application scored 100. Others scored 32 to 47.
- Transit Expansion Measure (10 applications)
  - Measure 2. Top application scored 350. Second application scored 247. Others scored from 10 to 76.
  - Measure 7. Top application scored 100. Others scored from 4 to 16.
- Transit System Modernization (13 applications)
  - Measure 2. Top application scored 300. Others scored from 1 to 96.
  - Measure 7. Top application scored 100. Others scored from 0 to 16.
- Travel Demand Management (6 applications)
  - Measure 2. Top application scored 100. Others scored from 6 to 23.

These 18 potential outliers led to scoring committees being able to adjust for outliers, starting with the 2018 Regional Solicitation. Staff only identified three outliers in 2018, following the adjustments.

## 2. Generally Successful Adjustment? Probably Not.

Regional Solicitation: 2018

Funding Category: Roadway Strategic Capacity

Scoring Measure: 1B Connection to Total Jobs, Manufacturing/Distribution Jobs, and Students (Connection to Total Jobs Component)

Employment w/i 1 mile	Score (Max 50) - Per Scoring Guidance	Final Score - Per Removal of High Scoring Outlier	Change in Gap Over Below Score	Change in Gap vs. Top
72,624	50	50	-40	N/A
13,974	10	50	+10	-40
10,291	7	37	+2	-30
9,813	7	35	0	-28
9,373	6	34	+6	-28
7,705	5	28	+1	-23
7,546	5	27	+3	-22
6,585	5	24	+1	-19
6,172	4	22	+2	-18
5,460	4	20	+2	-16
5,044	3	18	0	-15
5,001	3	18	+8	-15
2,609	2	9	+4	-7
1,064	1	4	+1	-3
787	1	3	+1	-2
440	0	2	+1	-2
276	0	1	N/A	-1

The original scoring spread resulted in one application scoring 50 points while 16 applications scored 0 to 10 points, providing almost no differentiation among the applications not ranked first. The adjustment in the third column was meant to address this, though the change in scoring gap was marginal, primarily impacting the advantage of the top-performing application. This is shown in the far-right column, which shows the loss of margin between each project and the top-rated project.

Overall, this adjustment was most beneficial to the second-ranked project and most damaging to the top-ranked project.

### 3. Potentially Unnecessary Adjustment

Regional Solicitation: 2020

Funding Category: Roadway Strategic Capacity

Scoring Measure: 1B Connection to Total Jobs, Manufacturing/Distribution Jobs, and Students (Connection to Total Jobs Component)

Employment w/i 1 mile	Score (Max 50) - Per Scoring Guidance	Final Score - Per Removal of High Scoring Outlier	Change in Gap Over Below Score	Change in Gap vs. Top
10,285	50	50	-4	N/A
9,363	46	50	0	-4
6,461	31	35	+1	-4
6,435	31	34	+1	-3
4,709	23	25	0	-2
4,495	22	24	0	-2
4,131	20	22	+1	-2
3,427	17	18	0	-1
2,094	10	11	0	-1
1,864	9	10	0	-1
1,734	8	9	0	-1
1,678	8	9	0	-1
1,635	8	9	-1	-1
1,064	5	6	-1	-1
695	3	4	+1	-1
579	3	3	0	0
555	3	3	0	0
276	1	1	N/A	0

The original scoring spread included some bunching toward the bottom. However, the overall spread does not appear to lack separation. Further, the adjustment has a negligible effect on the overall spread and is most impactful on the reduced advantage for the top-performing application.

Overall, the adjustment was minimally impactful.