

Trimming Ceiling and Floor Effects for Upper/Lower Groups: A Simulation Study

Koichi YAMAOKA

姫 路 獨 協 大 学

国際言語文化論集 第5号 抜刷

2024年2月 発行

Trimming Ceiling and Floor Effects for Upper/Lower Groups: A Simulation Study

Koichi YAMAOKA

Ceiling and floor effects have long been studied in the literature. For example, Zhu and Gonzalez (2017) investigated the impact on post-intervention outcomes, applying three different models to real data sets: (a) the traditional multilevel model, (b) the robust standard error approach, and (c) the multilevel Tobit model. By comparing the results of these three models, they illustrated that ignoring floor effects could result in misleading findings regarding the effectiveness of interventions. They further conducted a Monte Carlo simulation to show that ignoring floor effects can lead to biased results, and showed that the multilevel Tobit model worked very well and provided the least biased results in their study. Tobit regression deals with censored data, including ceiling and floor effects (McBee, 2010). Wilcox (2001) and Wilcox and Keselman (2002) presented another solution for dealing with ceiling and floor effects, especially, to trim both ends of the distribution. Liu and Wang (2021) investigated this issue of trimming and proposed formulae to calculate the mean value and variance of a normal distribution trimmed (or *truncated*) at given thresholds. They stated that, instead of the true values, trimming results in smaller variances, and smaller or larger means for ceiling and floor effects, respectively, compared to the true values. Liu and Wang have further illustrated the influence of ceiling and floor effects on the results of *t*-test and ANOVA.

This study investigates the effects of such trimming, focusing on a specific case in which the sample data set is divided into two groups, namely, the upper and lower groups. The idea is simple in that cutting off unmeasurable samples may mitigate the influence of ceiling and floor effects if such samples are distorting the distribution. The present study deals with a specific case in which the scores were broadly distributed from the minimum to the maximum edges of the measurement range at the pre-treatment time. If the scores increase after treatment, the right side of the distribution

hits the maximum measurable threshold, resulting in ceiling effects. If the scores decrease, the left side of the distribution hits the minimum threshold, resulting in floor effects.

This study is a preliminary simulation case study to establish a bench mark for the influence of ceiling and floor effects on a broad distribution, with no intent of immediate generalization. First, some insights into the ceiling and floor effects on the outcomes are introduced, followed by the simulation results of such effects. The simulation focused on ceiling effects, because the findings for the floor effects were expected to be similar in the opposite direction based on the symmetric nature of the simulated dataset, whereby ceiling effects distort the right side of the distribution, while the floor effects distort the left side.

Ceiling and Floor Effects on Outcomes

Suppose that the participants engage in pre-treatment, treatment, and post-treatment tests, consecutively. The pre- and post-treatment tests consist of 10 question items each, such that the measurement range of the tests is from 0 to 10. Figure 1 provides a conceptual image of the distributions of the scores. The pre-treatment distribution, with a mean value of 5 and σ of 1.5, is spread widely between the minimum and maximum edges of the measurement range (i.e., 0 to 10). In this example, the treatment is supposed to cause a 2.0-point worth improvement in scores. Thus, the post-treatment scores are shifted to the right by 2.0 points from the pre-treatment distribution, resulting in a mean value of 7 and σ of 1.5.

With the measurement range set as 0 to 10, values increased above 10 will be measured as 10 owing to ceiling effects, and values decreased under 0 will be measured as 0 owing to floor effects. Figure 2 provides a conceptual representation of the ceiling and floor effects from the perspective of outcomes. The outcome here is the post-treatment score minus the pre-treatment score. The horizontal axis shows the pre-treatment score, and the vertical axis shows the outcome. The measurement range of the outcome varies depending on the pre-treatment score, and only the shadowed area is measurable. For example, participants with a pre-treatment score of 10 cannot obtain higher points in their post-treatment scores. Thus, the post-treatment scores fail to measure the improvement in the outcomes of these participants; the measured outcome is zero, regardless of how much improvement is actually made. On the other hand, in the cases where participants' scores decline for some reason, those with a pre-treatment

score of zero cannot obtain lower points in the post-treatment scores. Thus, the post-treatment scores fail to measure the decline in scores of such participants; the measured outcome is zero, regardless of the extent of decline.

While this already seems problematic, the situation is worsened when the participants are divided into the upper and lower groups. Figure 3 shows an extreme case in which participants achieve a 5-point improvement after the treatment across the board. Assuming that the participants are divided into the upper group with pre-treatment scores of five points or above, and the lower group with pre-treatment scores below five, the measured mean outcome for the lower group ($= M_{outL}$) is five, because a 5-point improvement is measured for every participant. However, the measured mean outcome for the upper group ($= M_{outU}$) is less than 5, because the higher scores are cut off due to ceiling effects. This may lead to the incorrect conclusion that the treatment is more effective for the lower group than for the upper group.

In the case of floor effects, the opposite is true. For example, assuming that the participants' scores decline by 5.0, the upper group exhibits an average outcome of -5.0, but the lower group's average outcome suffers less because the low scores are cut off. This may lead to the incorrect conclusion that the lower group is more robust against the treatment than the upper group is.

One solution for ceiling and floor effects is to trim both ends of the distribution. Studies on trimmed distributions have reported that trimming both ends by 20% is the optimal solution (Wilcox, 2001; Wilcox & Keselman, 2002). However, the purpose of the current study is to investigate a bench mark case of ceiling and floor effects and the mitigating effects of trimming, rather than to generalize a mathematical theory. The idea behind this study is simple: The influence of ceiling and floor effects may be mitigated to a certain extent, by cutting off the samples that cannot be correctly measured because of such effects. Based on this idea, this study examined the effects of trimming both ends by the same extent as the mean outcome.

It is also a common practice to divide the obtained samples into the upper and lower groups. Thus, this study investigated cases in which the samples were divided into the upper and lower groups. The research questions are as follows:

- (1) To what extent do ceiling and floor effects affect the descriptive statistics of the measured distribution?
- (2) Does trimming mitigate the influence of the ceiling and floor effects in (1)?
- (3) Does the ceiling/floor effects affect t -test results?

This study mainly focuses on ceiling effects. The influence of floor effects is expected to be similar but in the opposite direction.

Method

Data Generation

Three distributions were generated for the analysis. Table 1 provides an overview of the three distributions. “Latent distribution” here refers to the distribution to be rendered to measurement. Next, a “measured distribution” is the resultant distribution obtained by measuring the latent distribution within a measurement range of 0 to 10. Therefore, this distribution is subject to ceiling and floor effects. A factor of measurement error ($= e_i$) is added to this distribution. Lastly, a “trimmed distribution” is obtained by trimming both ends of the measured distribution.

(1) Latent Distribution

A normal distribution with a sample size of 2,000 and mean and σ values of 5.0 and 1.5, respectively, was generated using R Statistical Software (Version 3.6.3). The resultant distribution exhibited $M = 5.07$ and $SD = 1.51$. This distribution was regarded as the latent distribution at the pre-treatment test. The i th sample of the latent distribution in the pre-treatment test is expressed as follows:

$$y_i^* = x_i, \quad (1)$$

where x_i denotes i th sample of the normal distribution generated by R.

The i th sample for the latent distribution in the post-treatment test is expressed as follows:

$$y_i^* = x_i + a + e_i, e_i = N(0, 0.5^2), \quad (2)$$

where x_i is the i th sample of the normal distribution generated by R, a is a constant representing the effects of the treatment, and e_i follows a normal distribution and represents measurement errors ($M = 0$ and $\sigma = 0.5$). The distribution of e_i was generated using R.

(2) Measured Distribution

The measured distribution was obtained by setting the scores of the latent distribution above 10 to 10 (ceiling effect), and those below 0 to 0 (floor effect). The i th sample of the measured distribution (y_i) is expressed as follows:

$$y_i = y_i^* \quad \text{for } 0 \leq y_i^* \leq 10, \quad (3)$$

$$y_i = 10 \quad \text{for } y_i^* > 10, \quad (4)$$

$$y_i = 0 \quad \text{for } y_i^* < 0, \quad (5)$$

where y_i^* is the i th sample of the latent distribution in the pre- and post-treatment tests.

(3) Trimmed Distribution

Both ends of the measured distribution were cut by α to obtain the trimmed distribution. The i th sample of the trimmed distribution (y_i^t) is expressed as follows:

$$y_i^t = y_i \text{ for } \alpha \leq y_i \leq 10 - \alpha, \quad (8)$$

where y_i is the i th sample of the measured distribution of the pre- and post-treatment scores.

The sample was further divided into the upper and lower groups. The upper group had pre-treatment scores of five points or above, and the lower group had pre-treatment scores below five.

Results

Scatter Plots

Figure 4 provides scatter plots of pre-treatment scores and outcomes for the measured distribution at $\alpha = 1, 1.5, 2$, and 3 . These values correspond to $0.67, 1.0, 1.3$, and 2.0 times of the standard deviation, respectively. The right side of each plot was cut down to the maximum measurement range, clearly illustrating ceiling effects. It can be seen from Figure 4 that the larger the α value is, the larger is the affected area of the plots.

Effects of α on the Outcomes of the Upper Group

First, the effects of α on the outcomes of the upper group are illustrated, because they exhibited the most noticeable mitigating effects of trimming. Figures 5 and 6 show the mean values and standard deviations of the outcomes of the upper group, respectively, in terms of *difference in ratio*. *Difference in ratio* is the deviation from the latent distribution divided by the latent distribution value. A value of zero indicates that the measured value was the same as that of the latent distribution, thereby confirming that the latent distribution was measured correctly. If the difference in ratio is -0.1 , for example, the value is 10% less than that of the latent distribution.

As is shown in Figures 5, the mean value of the measured distribution deviated more from the true value (i.e., latent distribution) as α increased, while such effects were mitigated in the trimmed distribution. Figure 6 shows the mitigating effects of trimming on the standard deviations. Considering that such an effect manifests itself at larger values of α , the next section focuses on the effects on each variable at $\alpha = 3$ to illustrate this tendency clearly.

Effects of Trimming ($\alpha = 3$)

Figure 7 illustrates the difference in ratio of variables in the measured and trimmed distributions at $\alpha = 3$. In Figure 7 (a), the mean values were relatively correct for the measured distribution, although the mean outcome for the upper group suffered the most (-0.05). The standard deviation for the upper group suffered in the post-treatment scores (-0.26) and outcomes (0.16). The ceiling effects did not affect the lower group. In summary, the ceiling effects affected the upper group most severely, followed by the overall group. The lower group was not affected by ceiling effects.

Figure 7 (b) shows that the mean values of the trimmed distributions were correct for the overall group. However, this was not the case for the upper and lower groups. The mean values suffered in the pre-treatment scores for the upper (-0.06) and lower (0.09) groups, and in the post-treatment scores of the upper (-0.04) and lower (0.05) groups. However, one noticeable feature was that the mean outcome was correct for all the groups.

The standard deviations in the trimmed distribution suffered significantly in the pre- and post-treatment scores for all the groups: -0.32, -0.41, and -0.39 in the pre-treatment scores, and -0.29, -0.32, and -0.28 in the post-treatment scores, for the overall, upper, and lower groups, respectively. However, the outcomes were not significantly affected; -0.02, -0.06, and 0.02 in the overall, upper, and lower groups, respectively.

In short, trimming was effective against the influence of ceiling effects on the mean values in the overall group, but it disturbed the means of the pre- and post-treatment scores in the upper and lower groups. Trimming affected the standard deviations of the pre- and post-treatment scores in all the groups; however, it was effective for the outcome values, including both the means and standard deviations, for the overall and upper groups.

Details of Descriptive Statistics

Table 2 provides the details of the descriptive statistics of the measured and trimmed distributions in terms of the ratio to the latent distribution. As Table 2 (a) illustrates, the influence of ceiling effects was relatively small at $\alpha = 1$: 2% or less difference for the measured distribution and 5% or less difference for the trimmed distribution, including both the means and standard deviations. On the other hand, Table 2 (d) shows that trimming improved the mean and standard deviation of the outcome at $\alpha = 3$, but worsened the situation for the other variables. For example, the

mean value of the post-treatment scores (M_{post}) was -2% different for the upper group in the measured distribution, but it was -4% different in the trimmed distribution. The standard deviation of the pre-treatment scores was 0% different for the upper group in the measured distribution, but -41% different in the trimmed distribution.

Results of *t*-tests

Table 3 shows the results of independent *t*-tests conducted to determine differences between the upper and lower group outcomes. The results were significantly different for the measured distribution at $\alpha = 3$, which in fact, is not true as the upper and lower groups had the same α value. This suggests a possibility of making a wrong decision that “the treatment was more efficient for the lower group with statistical significance when $\alpha = 3$.” For the other α values, the results were not statistically significant.

Discussion

When the α was as small as 1.0, the values of the measured distribution did not differ much from those of the latent distribution: the differences were not more than 1 % for the mean values and not more than 2% for the standard deviations. At $\alpha = 3.0$, however, the differences reached 5% for the mean values and 26% for the standard deviations. Thus, the answer to RQ 1 is that the means and standard deviations did not suffer so much (not more than 2%) at $\alpha = 1$, but the standard deviations suffered by as much as 26% at $\alpha = 3$.

Trimming affected the mean pre- and post-treatment scores in the upper and lower groups, resulting in a lower value for the upper group and a higher value for the lower group. This is a natural consequence, considering that the upper group loses high scores and the lower group loses low scores. However, trimming significantly mitigated the influence of ceiling effects on the mean outcome. For example, the measured distribution exhibited a lower mean outcome than that of the latent distribution at $\alpha = 3$, but the trimmed distribution exhibited exactly the same mean outcome as that of the latent distribution. This suggests that the difference between the measured and trimmed distributions served as a rough indicator of the influence of ceiling effects. In other words, a large difference in the mean outcome between the two distributions indicates a large influence of ceiling effects. However, similar mean outcomes in the two distributions indicate a small influence of ceiling effects. Thus, the answer to RQ 2 is that trimming mitigates the influence of ceiling and floor effects on outcome values but

at the cost of measuring many other values.

The t -test results indicated the possibility of an incorrect conclusion when the outcome was very large. Thus the answer to RQ 3 is that ceiling/floor effects can affect the t -test results when the mean outcome is large.

Conclusion

When the lower group exhibits more improvement than the upper group after a treatment, it is necessary to exercise caution when interpreting the results as they may have been distorted by the ceiling effect. Conversely, when the upper group shows more decrease in performance compared to the lower group after a treatment, the floor effect may be distorting the results. One way to mitigate this influence is to trim both ends of the distribution. Therefore, this study investigated the effects of trimming in a simulation case. This study has some limitations. It is a case study with a specific distribution and measurement range. The parameters of the error term in the post-treatment latent distribution was arbitrary without any strong rationale for setting the σ to 0.5. Despite the limitations, the findings provide some insights into the influence of distribution-trimming on ceiling/floor effects. The results revealed that the influence of ceiling effects was insignificant when the mean outcome was relatively small. However, when the mean outcome was relatively large, the influence was larger, especially for the standard deviations. Trimming preserved the mean values and standard deviations of the outcomes relatively well. However, this was at the cost of the mean values of the pre- and post-treatment scores for the upper and lower groups, and incorrect standard deviations of the pre- and post-treatment scores for all groups. Another useful finding is that trimming preserved the correct mean outcome value. This suggests that the difference in the mean outcomes between the measured and trimmed distributions can be a rough indicator of the influence of ceiling effects. The t -test results indicated the possibility of making an incorrect conclusion when the mean outcome was very large, although the significant/nonsignificant decision was not problematic with small mean outcomes, at least in this specific case.

References

- Liu, Q., & Wang, L. (2021). *t*-Test and ANOVA for data with ceiling and/or floor effects. *Behavior Research Methods*, 53 (1), 264–277. <https://doi.org/10.3758/s13428-020-01407-2>
- McBee, M. (2010). Modeling outcomes with floor or ceiling effects: An introduction to the Tobit Model. *Gifted Child Quarterly*, 54(4), 314–320. <https://doi.org/10.1177/0016986210379095>
- Wilcox, R. R. (2001). Pairwise comparisons of trimmed means for two or more groups. *Psychometrika*, 66 (3), 343–356. <https://doi.org/10.1007/BF02294438>
- Wilcox, R. R., & Keselman, H. J., (2002). Power analyses when comparing trimmed means. *Journal of Modern Applied Statistical Methods*, 1 (1), 24–31. <https://doi.org/10.22237/jmasm/1020254820>
- Zhu, L., & Gonzalez, J. (2017). Modeling floor effects in standardized vocabulary test scores in a sample of low SES Hispanic preschool children under the multilevel structural equation modeling framework. *Frontiers in Psychology*, 8, 2146. <https://doi.org/10.3389/fpsyg.2017.02146>

Table 1*Distributions Used in the Analyses*

Test	Distribution		
	Latent	Measured	Trimmed
Pre-treatment	Normally distributed random numbers were generated by R.	Scores of the latent distribution above 10 were set to 10 (ceiling effect), and those under 0 were set to 0 (floor effect).	Both ends of the measured distribution were trimmed by α .
	Sample Size = 2,000 $M = 5.07$ $SD = 1.51$ $y_i^* = x_i$	$y_i = y_i^*$ for $0 \leq y_i^* \leq 10$, $y_i = 10$ for $y_i^* > 10$, $y_i = 0$ for $y_i^* < 0$.	$y_i^t = y_i$ for $\alpha \leq y_i \leq 10 - \alpha$
Post-treatment	A constant value (α) and an error (e_i) were added to each score of the pre-treatment distribution.	Scores of the latent distribution above 10 were set to 10 (ceiling effect), and those under 0 were set to 0 (floor effect).	Both ends of the measured distribution were trimmed by α .
	$y_i^* = x_i + \alpha + e_i$, $e_i = N(0, 0.5^2)$	$y_i = y_i^*$ for $0 \leq y_i^* \leq 10$, $y_i = 10$ for $y_i^* > 10$, $y_i = 0$ for $y_i^* < 0$.	$y_i^t = y_i$ for $\alpha \leq y_i \leq 10 - \alpha$

Table 2*Difference in Ratio from the Latent Distribution*

Distribution	M_{pre}	M_{post}	M_{out}	SD_{pre}	SD_{post}	SD_{out}
(a) $\alpha = 1$						
Measured						
Overall	0.00	0.00	0.00	0.00	-0.01	0.00
Upper	0.00	0.00	-0.01	0.00	-0.02	0.01
Lower	0.00	0.00	0.00	0.00	0.00	0.00
Trimmed						
Overall	0.00	0.00	0.00	-0.03	-0.03	0.00
Upper	0.00	0.00	0.00	-0.04	-0.04	0.00
Lower	0.01	0.01	0.00	-0.05	-0.03	0.00
(b) $\alpha = 1.5$						
Measured						
Overall	0.00	0.00	0.00	0.00	-0.01	0.01
Upper	0.00	0.00	-0.01	0.00	-0.04	0.01
Lower	0.00	0.00	0.00	0.00	0.00	0.00
Trimmed						
Overall	0.00	0.00	0.00	-0.07	-0.06	-0.01
Upper	-0.01	-0.01	0.00	-0.09	-0.07	-0.01
Lower	0.02	0.01	0.00	-0.12	-0.09	0.00
(c) $\alpha = 2$						
Measured						
Overall	0.00	0.00	-0.01	0.00	-0.02	0.01
Upper	0.00	0.00	-0.02	0.00	-0.08	0.02
Lower	0.00	0.00	0.00	0.00	0.00	0.00
Trimmed						
Overall	0.00	0.00	0.00	-0.13	-0.12	-0.01
Upper	-0.02	-0.02	0.00	-0.18	-0.14	-0.02
Lower	0.03	0.02	0.00	-0.19	-0.14	0.00
(d) $\alpha = 3$						
Measured						
Overall	0.00	-0.01	-0.03	0.00	-0.09	0.10
Upper	0.00	-0.02	-0.05	0.00	-0.26	0.16
Lower	0.00	0.00	0.00	0.00	0.00	0.00
Trimmed						
Overall	-0.01	0.00	0.00	-0.32	-0.29	-0.02
Upper	-0.06	-0.04	0.00	-0.41	-0.32	-0.06
Lower	0.09	0.05	0.00	-0.39	-0.28	0.02

Note. Overall = the overall group; Upper = the upper group; Lower = the lower group. M_{pre} represents mean of the pre-treatment scores, M_{post} represents mean of the post-treatment score, M_{out} represents mean of outcomes, SD_{pre} represents the standard deviation of the pre-treatment scores, SD_{post} represents the standard deviation of the post-treatment scores, and SD_{out} represents the standard deviation of outcomes.

Table 3

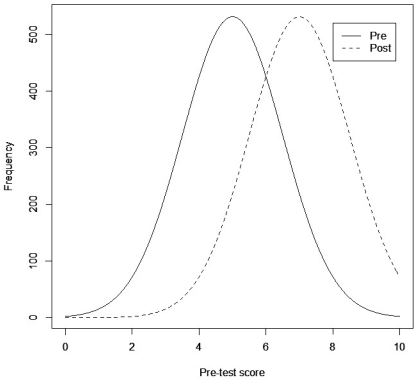
Results of t-Tests of Outcomes

Distribution	M_{out}		t	df	p
	Upper	Lower			
(a) $\alpha = 1$					
Latent	1.01	1.00	0.424	1998	0.672
Measured	1.01	1.00	0.166	1998	0.868
Trimmed	1.01	1.00	0.466	1983	0.641
(b) $\alpha = 1.5$					
Latent	1.51	1.50	0.424	1998	0.672
Measured	1.50	1.50	-0.156	1998	0.876
Trimmed	1.51	1.50	0.460	1954	0.646
(c) $\alpha = 2$					
Latent	2.01	2.00	0.424	1998	0.672
Measured	1.98	2.00	-1.014	1998	0.311
Trimmed	2.01	2.00	0.393	1897	0.694
(d) $\alpha = 3$					
Latent	3.01	3.00	0.424	1998	0.672
Measured	2.86	3.00	56.876	1998	<.001 ^a
Trimmed	3.00	3.00	0.028	1630	0.978 ^a

Note. Student's t -test. ^a Levene's test was significant ($p < .05$), suggesting a violation of the equal variance assumption.

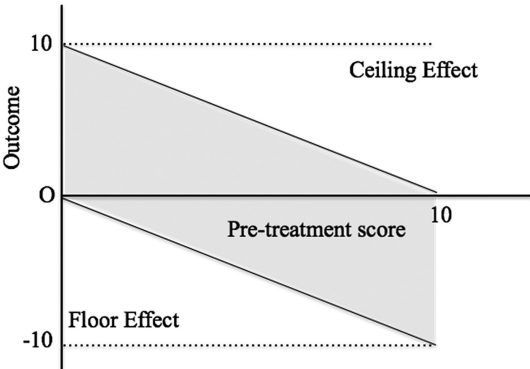
Figure 1

Conceptual Image of Distributions



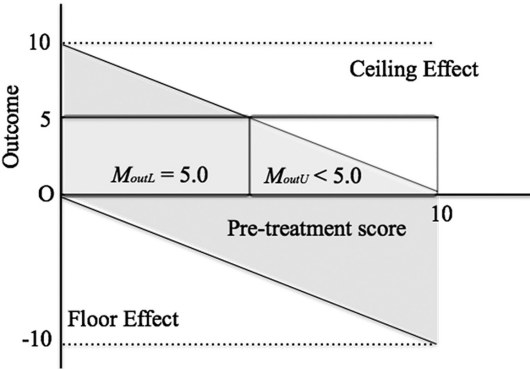
Note. Pre refers to pre-treatment distribution ($M = 5.0$, $\sigma = 1.5$), and Post refers to post-treatment distribution ($M = 7.0$, $\sigma = 1.5$)

Figure 2
Conceptual Image of Ceiling and Floor Effects on Outcomes



Note. The outcome is the post-test score minus the pre-test score. Only the shadowed area is measurable.

Figure 3
An Example of Ceiling Effects on Outcome

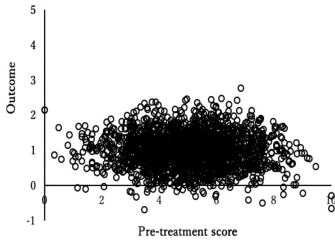


Note. The outcome is the post-test score minus the pre-test score. M_{outL} and M_{outU} are the mean values of the outcomes for the lower and upper groups, respectively.

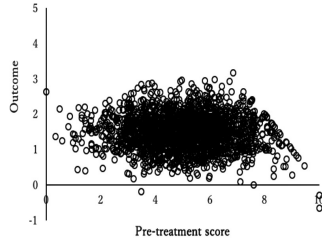
Figure 4

Scatter Plots of Outcomes for the Measured Distribution

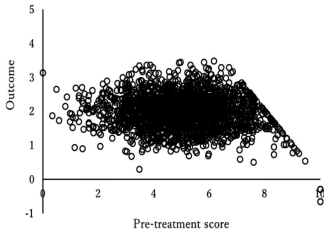
(a) $\alpha = 1$



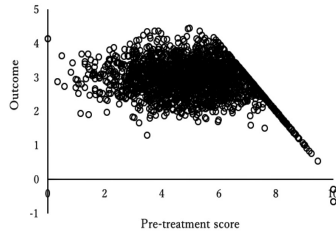
(b) $\alpha = 1.5$



(c) $\alpha = 2$



(d) $\alpha = 3$

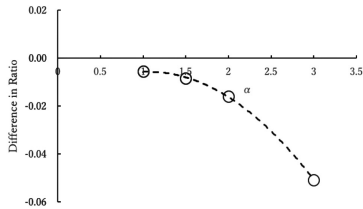


Note. $n = 2000$. Outcome = post-treatment score minus pre-treatment score. The variable α is the mean outcome of the latent distribution.

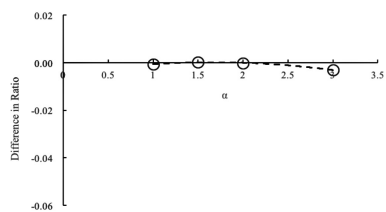
Figure 5

Differences in Mean Outcomes in Ratio (Upper Group)

(a) Measured Distribution



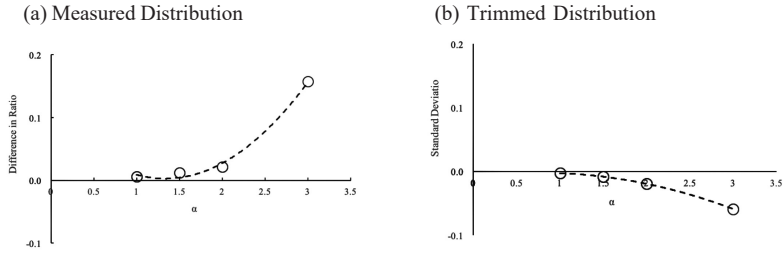
(b) Trimmed Distribution



Note. The difference in the ratio is the deviation from its latent distribution in ratio. Dotted lines show second-order approximation curves.

Figure 6

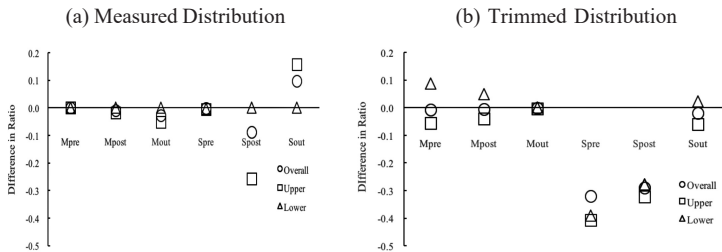
Differences in Standard Deviations of Outcomes in Ratio (Upper Group)



Note. The difference in the ratio is the deviation from its latent distribution in ratio. Dotted lines show second-order approximation curves.

Figure 7

Difference in Means and Standard Deviations ($\alpha = 3$)



Note. The difference in ratio is the deviation from its latent distribution. M_{pre} is the mean score in the pre-treatment test, M_{post} is the mean score in the post-treatment test, M_{out} is the mean outcome, S_{pre} is the standard deviation of the pre-treatment scores, S_{post} is the standard deviation of the post-treatment scores, and S_{out} is the standard deviation of the outcomes. *Overall* is the overall group, *Upper* is the upper group, and *Lower* is the lower group.

Appendix A

Table A1

Effects of Right Shift of Distribution ($\alpha = 1$)

Distribution	M_{pre}	M_{post}	$M_{outcome}$	SD_{pre}	SD_{post}	$SD_{outcome}$
Overall						
Latent	5.07	6.07	1.01	1.51	1.59	0.49
Measured	5.07	6.07	1.00	1.51	1.58	0.50
Trimmed ^a	5.07	6.08	1.01	1.47	1.54	0.49
Upper						
Latent	6.20	7.21	1.01	0.91	1.02	0.50
Measured	6.20	7.20	1.01	0.91	1.00	0.50
Trimmed ^b	6.18	7.19	1.01	0.87	0.98	0.50
Lower						
Latent	3.78	4.78	1.00	0.91	1.03	0.49
Measured	3.78	4.78	1.00	0.91	1.03	0.49
Trimmed ^c	3.80	4.80	1.00	0.87	0.99	0.49

Note. $n = 2000$. ^a ceiling = 9, trimmed = 15. ^b ceiling = 9, trimmed = 7. ^c ceiling = 0, trimmed = 8.

Table A2

Effects of Right Shift of Distribution ($\alpha = 1.5$)

Distribution	M_{pre}	M_{post}	$M_{outcome}$	SD_{pre}	SD_{post}	$SD_{outcome}$
Overall						
Latent	5.07	6.57	1.51	1.51	1.59	0.49
Measured	5.07	6.57	1.50	1.51	1.57	0.50
Trimmed ^a	5.08	6.58	1.51	1.41	1.49	0.49
Upper						
Latent	6.20	7.71	1.51	0.91	1.02	0.50
Measured	6.20	7.70	1.50	0.91	0.98	0.50
Trimmed ^b	6.14	7.66	1.51	0.83	0.94	0.49
Lower						
Latent	3.78	5.28	1.50	0.91	1.03	0.49
Measured	3.78	5.28	1.50	0.91	1.03	0.49
Trimmed ^c	3.85	5.35	1.50	0.81	0.94	0.49

Note. $n = 2000$. ^a ceiling = 24, trimmed = 44. ^b ceiling = 24, trimmed = 20. ^c ceiling = 0, trimmed = 24.

Trimming Ceiling and Floor Effects for Upper/Lower Groups A Simulation Study

Table A3

Effects of Right Shift of Distribution ($\alpha = 2$)

Distribution	M_{pre}	M_{post}	$M_{outcome}$	SD_{pre}	SD_{post}	$SD_{outcome}$
Overall						
Latent	5.07	7.07	2.01	1.51	1.59	0.49
Measured	5.07	7.06	1.99	1.51	1.55	0.50
Trimmed ^a	5.06	7.07	2.01	1.32	1.40	0.49
Upper						
Latent	6.20	8.21	2.01	0.91	1.02	0.50
Measured	6.20	8.18	1.98	0.91	0.94	0.51
Trimmed ^b	6.07	8.09	2.01	0.75	0.87	0.49
Lower						
Latent	3.78	5.78	2.00	0.91	1.03	0.49
Measured	3.78	5.78	2.00	0.91	1.03	0.49
Trimmed ^c	3.90	5.91	2.00	0.74	0.88	0.49

Note. $n = 2000$. ^a ceiling = 64, trimmed = 101. ^b ceiling = 64, trimmed = 54. ^c ceiling = 0, trimmed = 47.

Table A4

Effects of Right Shift of Distribution ($\alpha = 3$)

Distribution	M_{pre}	M_{post}	$M_{outcome}$	SD_{pre}	SD_{post}	$SD_{outcome}$
Overall						
Latent	5.07	8.07	3.01	1.51	1.59	0.49
Measured	5.07	7.99	2.93	1.51	1.45	0.54
Trimmed ^a	5.03	8.03	3.00	1.03	1.13	0.48
Upper						
Latent	6.20	9.21	3.01	0.91	1.02	0.50
Measured	6.20	9.06	2.86	0.91	0.76	0.57
Trimmed ^b	5.85	8.85	3.00	0.54	0.69	0.47
Lower						
Latent	3.78	6.78	3.00	0.91	1.03	0.49
Measured	3.78	6.78	3.00	0.91	1.03	0.49
Trimmed ^c	4.11	7.11	3.00	0.56	0.74	0.50

Note. $n = 2000$. ^a ceiling = 212, trimmed = 368. ^b ceiling = 212, trimmed = 200. ^c ceiling = 0, trimmed = 168.

Appendix B

Table B1

Effects of Left Shift of Distribution ($\alpha = -1$)

Distribution	M_{pre}	M_{post}	$M_{outcome}$	SD_{pre}	SD_{post}	$SD_{outcome}$
Overall						
Latent	5.07	4.07	-0.99	1.51	1.59	0.49
Measured	5.07	4.08	-0.99	1.51	1.58	0.49
Trimmed ^a	5.07	4.08	-0.99	1.47	1.54	0.49
Upper						
Latent	6.20	5.21	-0.99	0.91	1.02	0.50
Measured	6.20	5.21	-0.99	0.91	1.02	0.50
Trimmed ^b	6.18	5.19	-0.99	0.87	0.98	0.50
Lower						
Latent	3.78	2.78	-1.00	0.91	1.03	0.49
Measured	3.78	2.78	-0.99	0.91	1.01	0.49
Trimmed ^c	3.80	2.81	-1.00	0.87	0.99	0.49

Note. $n = 2000$. ^a floor = 14, trimmed = 15. ^b floor = 0, trimmed = 7. ^c floor = 14, trimmed = 8.

Table B2

Effects of Left Shift of Distribution ($\alpha = -1.5$)

Distribution	M_{pre}	M_{post}	$M_{outcome}$	SD_{pre}	SD_{post}	$SD_{outcome}$
Overall						
Latent	5.07	3.57	-1.49	1.51	1.59	0.49
Measured	5.07	3.58	-1.48	1.51	1.57	0.49
Trimmed ^a	5.08	3.59	-1.49	1.41	1.49	0.49
Upper						
Latent	6.20	4.71	-1.49	0.91	1.02	0.50
Measured	6.20	4.71	-1.49	0.91	1.02	0.50
Trimmed ^b	6.14	4.66	-1.48	0.83	0.95	0.50
Lower						
Latent	3.78	2.28	-1.50	0.91	1.03	0.49
Measured	3.78	2.29	-1.48	0.91	0.99	0.49
Trimmed ^c	3.85	2.35	-1.50	0.81	0.93	0.49

Note. $n = 2000$. ^a floor = 28, trimmed = 44. ^b floor = 0, trimmed = 20. ^c floor = 28, trimmed = 24.

Trimming Ceiling and Floor Effects for Upper/Lower Groups A Simulation Study

Table B3

Effects of Left Shift of Distribution ($\alpha = -2$)

Distribution	M_{pre}	M_{post}	$M_{outcome}$	SD_{pre}	SD_{post}	$SD_{outcome}$
Overall						
Latent	5.07	3.07	-1.99	1.51	1.59	0.49
Measured	5.07	3.09	-1.97	1.51	1.55	0.50
Trimmed ^a	5.06	3.07	-1.99	1.32	1.40	0.49
Upper						
Latent	6.20	4.21	-1.99	0.91	1.02	0.50
Measured	6.20	4.21	-1.99	0.91	1.02	0.50
Trimmed ^b	6.07	4.09	-1.98	0.75	0.89	0.49
Lower						
Latent	3.78	1.78	-2.00	0.91	1.03	0.49
Measured	3.78	1.81	-1.96	0.91	0.94	0.51
Trimmed ^c	3.90	1.91	-1.99	0.74	0.87	0.49

Note. $n = 2000$. ^a floor = 59, trimmed = 101. ^b floor = 0, trimmed = 54. ^c floor = 59, trimmed = 47.

Table B4

Effects of Left Shift of Distribution ($\alpha = -3$)

Distribution	M_{pre}	M_{post}	$M_{outcome}$	SD_{pre}	SD_{post}	$SD_{outcome}$
Overall						
Latent	5.07	2.07	-2.99	1.51	1.59	0.49
Measured	5.07	2.15	-2.92	1.51	1.45	0.54
Trimmed ^a	5.03	2.06	-2.98	1.03	1.13	0.49
Upper						
Latent	6.20	3.21	-2.99	0.91	1.02	0.50
Measured	6.20	3.21	-2.99	0.91	1.02	0.50
Trimmed ^b	5.85	2.87	-2.98	0.54	0.73	0.49
Lower						
Latent	3.78	0.78	-3.00	0.91	1.03	0.49
Measured	3.78	0.94	-2.84	0.91	0.76	0.59
Trimmed ^c	4.11	1.13	-2.98	0.56	0.70	0.48

Note. $n = 2000$. ^a floor = 202, trimmed = 368. ^b floor = 0, trimmed = 200. ^c floor = 202, trimmed = 168.

Trimming Ceiling and Floor Effects for Upper/Lower Groups

A Simulation Study

Koichi YAMAOKA

Abstract

Trimming has long been investigated as a way to deal with ceiling/floor effects. This study focuses on the influence of trimming on ceiling/floor effects when the samples are divided into the upper and lower groups. Latent, measured, and trimmed distributions are compared to investigate the effects of trimming. The pre-treatment distributions, spread across the measurement range, are shifted by a mean outcome to obtain the post-treatment scores. The thresholds for trimming are set to the same value as the mean outcome because excluding unmeasurable samples can mitigate the ceiling/floor effects. The results reveal that the influence of the ceiling effects is insignificant when the mean outcome was relatively small. However, when the mean outcome is relatively large, the influence is larger, especially for the standard deviations. Trimming preserves the mean values and standard deviations of the outcomes relatively well. However, this comes with incorrect mean values of the pre- and post-treatment scores for the upper and lower groups, and incorrect standard deviations in the pre- and post-treatment scores for all groups. The findings suggest that the difference between the mean outcomes for the measured and trimmed distributions can roughly indicate the influence of ceiling effects. The t -test results indicate the possibility of making an incorrect conclusion when the mean outcome is very large, although the significant/nonsignificant decision is not problematic with small mean outcomes, at least in this specific case.

Keywords: ceiling effects, floor effects, trimming.