Promoting the performance of community childcare centers using the insight from behavioral economics: application of the last-place aversion and first-place seeking theory

Young-Gon Kim & Hosung Sohn

To cite this article: Young-Gon Kim & Hosung Sohn (2021) Promoting the performance of community childcare centers using the insight from behavioral economics: application of the last-place aversion and first-place seeking theory, International Review of Public Administration, 26:2, 192-208, DOI: 10.1080/12294659.2021.1892929

To link to this article: https://doi.org/10.1080/12294659.2021.1892929

Published online: 08 Mar 2021.
Promoting the performance of community childcare centers using the insight from behavioral economics: application of the last-place aversion and first-place seeking theory

Young-Gon Kim and Hosung Sohn

Office of Policy Development, Korea Institute for Defense Analysis, Seoul, South Korea; School of Public Service, Chung-Ang University, Seoul, South Korea

ABSTRACT
This study examines whether the performance level of community childcare centers that receive financial subsidy from governments can be improved by incorporating the insight from behavioral economics. Specifically, we test empirically whether incorporating the ‘last-place aversion’ and ‘first-place seeking’ theory into the performance evaluation system is beneficial for promoting the performance level of these governmentally funded facilities. Using a regression discontinuity design, we find that the facilities are sensitive to being placed in the lowest and first rank. The effect of being placed in the former is, however, stronger than the latter. We argue that incorporating the two theories is beneficial for making the facilities accountable for their performance.

I. Introduction
Evaluating the performance of government entities is important for promoting the effectiveness of such entities (Broom & McGuire, 1995). This is especially important for institutions that receive a large amount of government subsidies for operational purposes (Gilmour & Lewis, 2006; Grizzle & Pettijohn, 2002). As such, many efforts are put forward to improve the efficiency and effectiveness of performance evaluation system so that it can be helpful for the sustainability of these institutions (Joyce, 1993). The two most widely used systems are based on positive incentive and negative incentive (Destler, 2014). Implementation of these two systems requires additional budgetary costs (e.g. opportunity costs). If positive incentive systems are used, then additional budgets are required to reward those who outperformed. If negative incentive systems are used, then it may generate unintended consequences such as loss in self-confidence and morale.

While other types of efforts have been adopted to enhance performance evaluation system other than the two representative system mentioned above, recent research on behavioral economics provides an interesting insight on improving the existing performance evaluation system. Behavioral economics is a subfield of economics that studies
behavior of economic agents using insights from psychology such as loss aversion, present bias, inattention, and last-place aversion (Kahneman & Tversky, 1979). The debate on behavioral economics is usually framed as a question on the foundational assumptions of traditional economics. In this study, we do not engage in answering the normative aspects of such debates. Rather, the purpose of our study is positive. We test whether a behaviorally motivated theory can help develop performance evaluation system that is more effective and efficient. Specifically, this study examines whether the performance level of community childcare centers that receive financial subsidy from government can be improved by incorporating the insight from behavioral economics. We empirically test whether incorporating the 'last-place aversion' theory (Kuziemko et al., 2014) and 'first-place seeking' theory (Dutcher et al., 2015; Hardisty & Shechter, 2018) into the performance evaluation system is beneficial for promoting the performance level of these government-funded facilities.

To test whether the two theories are pervasive in performance-budgeting setting, we analyze the performance evaluation system adopted in evaluating the Community Children Center (CCC) in South Korea from year 2012 to 2017. The evaluation system suits our research purpose because the system identifies both the last-placed and first-placed centers. Most importantly, the evaluation system does not create any consequences other than ranks (i.e. grade). Therefore, the system is beneficial for testing whether the centers are sensitive to being placed in the last- or first-place. We use administrative data on all the CCCs from 2012 to 2017, and exploit the discontinuous nature of the grading system used to rank the CCCs. Exploiting such system, we employ regression discontinuity designs to isolate the causal impact of being placed in the lowest grade (last-place) and highest (first-place) grades. The results show that CCCs which are placed in the lowest grade outperformed those who are placed the next to lowest grade in the subsequent evaluation period, which coincides with the last-place aversion theory. For the tests of the first-place seeking behavior, we find that CCCs that are placed in the highest grade are more likely to secure their initial highest grade in the subsequent evaluation.

The results of this study indicate that incorporating the behavioral insights can help promote the performance level of social welfare facilities such as CCCs without inducing unintended consequences of performance evaluation system such as cheating and unfairness typically observed in the public sector (e.g. Georgellis et al., 2011; van Thiel & Leeuw, 2010). Moreover, our results imply that policymakers can make the facilities accountable for their performance without sacrificing additional budgets.

II. Literature review

As noted by Chetty (2015), behavioral economics offers innovative policy tools that can be used to influence economic agents’ behavior. Insights from behavioral economics offer new tools – such as changing the way performance is evaluated – that expand the set of treatment and outcome variables that can be achieved through these insights.

The existing performance evaluation system usually operates under two mechanisms: positive and negative incentives. The evaluation system that utilizes positive incentives usually provides rewards (or remuneration) for those who exceed a certain threshold that determines excellency. The evaluation system that operates under negative incentives, on
the other hand, usually impose penalties to those who have not achieved a minimum threshold that determines poor performance. These two mechanisms are adopted and operated under a widely established neoclassical economic theory called mechanism design theories.

Note that the two mechanisms can be strengthened by incorporating the insights from behavioral economics. To be more specific, because people are last-place averse and first-place seeking, people are more likely to respond to the system if the system incorporates such schemes. The fact that agents are rational is however, weakly supported empirically. In particular, recent experiments conducted by many of the behavioral economists show that economic agents are not always rational, though the findings are not always consistent with the view of behavioral economists (DellaVigna, 2009; List, 2004).

Nevertheless, if economic agents are not always rational, then it calls into question whether the existing performance evaluation system that operates under the assumption of rationality can effectively lead to an increase in the level of performance. Rather than using the traditional system, it may be more desirable to use the new system that incorporates the behavioral aspects of economic agents. In a conventional performance evaluation system, institutions are ranked based on their annual performance. Institutions are penalized or rewarded depending on their rank. This evaluation system is effective if institutions are rational and care about their relative ranking. For example, if institutions are rational, then the institution whose rank is lower than other institutions are more likely to put efforts in promoting their performance in subsequent years. To put it differently, the effort level is a linear function of ranks. As a matter of course, such linearity is conditional on the fact that some forms of positive or negative incentives are pervasive according to the relative rank of institutions.

Two recent insights from behavioral economics demonstrate that economic agents are more likely to be last-place averse and first-place seeking. Agents do not care about being lower rank, but they do care about being the lowest rank. Similarly, agents care more about being placed in the highest rank than being placed in higher ranks. Testing the two theories in performance-budgeting setting is beneficial because if the two theories is pervasive in promoting the performance level of the facilities, it implies that policymakers can make these facilities accountable for their performance without spending additional budgets.

Last-place aversion theory is pioneered by Kuziemko et al. (2014). They provide a theoretical foundation of the last-place aversion theory and experimental tests of the theory based on the gambling setting. They also validate their theory using survey data and show that people who are earning just above the minimum wage are much more likely to oppose the change in the minimum wage level. They also show that people who are located just above poverty but below median income are less likely to support redistribution than their background characteristics would predict. The seminal work of Kuziemko et al. (2014) indicates that people dislike being placed in the last-place and potentially provides useful insight with respect to improving the effectiveness and efficiency of performance evaluation system.

On the other hand, Dutcher et al. (2015) and Hardisty and Shechter (2018) propose a so-called the ‘victory effect’ which predicts that agents are first-place seeking. In their study, they test their theory based on the standard gamble method. They find that agents are risk seeking for first place, and risk-averse to avoid last place. They further find that
first-place seeking is stronger than last-place aversion. While the theoretical background of these two theories is driven by a psychological theories, their empirical basis is quite nascent. Accordingly, it is desirable to test whether these two theories are pervasive to other setting.

Testing the validity of these two theories is easier said than done for many reasons. First, overcoming the endogeneity issue is difficult. Suppose governments adopted a performance evaluation system that utilizes these two theories and suppose those who are placed at the lowest rank performed poorly than those who are placed just above the lowest rank. Given the difference, it is tempting to conclude that the CCCs are not last-place averse. Note, however, that there might be other possible confounding factors that attributed to the difference.

Another possible factor that prevents researchers from establishing causality is related to a measurement error in the treatment variable. In general, many of the performance evaluation systems consist of many aspects. Suppose that the CCCs are ranked based on their baseline performance. Governments may have revealed the ranking to the public in an effort to incentivize the CCCs. If this is the case, the fact that those who are placed in the last place did not perform well in a subsequent evaluation period than those who are placed just above the last place does not necessarily imply that the agents are not last-place averse. The relatively low performance observed for those placed in the last place may have been stigmatized for being the lowest rank and that may have adversely affected their subsequent performance.

Thus, the variable of interests that researchers are trying to test should not have other aspects other than being the last place or first place. The contributions of this study are three-fold. First, this study is the first research to test the two theories in the performance evaluation system. At this point, empirical evidence regarding the two theories is mostly based on laboratory setting, and there are few studies that engage in testing the validity of the theories in policy setting. Second, by making use of the unique feature of the performance evaluation system adopted in Korea, we argue that the estimated impact we isolate is internally valid. We explain this contribution more in detail in Section III. Third, we test the impact of both the last-place aversion theory and first-place seeking theory, and determine which is more effective in promoting performance.

III. Institutional background

The legal basis for the evaluation of the performance of CCCs in Korea is Article 54 of Child Welfare Act and Article 18 of Subsidy Management Act. Based on the Acts, each CCC is evaluated on 15 indices such as quality of employee education and programs. Total number of operating CCCs as of 2017 is about 4,000. Not all the 4,000 CCCs are evaluated each year. Rather, about 1/3 of the CCCs are evaluated each year. Based on information provided by the CCCs and field evaluation results conducted by evaluation committees, each CCC is notified of their rank. The rank is consisted of five grades. The range of evaluation scores by grade is presented in Table 1.

Using the performance evaluation system of CCCs is beneficial for testing the validity of the last-place aversion and first-place seeking theory for the following reason. From the CCC perspective, there are almost no consequences of being placed in the lowest or highest grade. If there are some forms of negative or positive incentives of being placed in
IV. Research design

This study uses a regression discontinuity design (RDD) to estimate the effect of being placed on the highest or lowest grade on subsequent performance. To illustrate the treatment assigning mechanism, let $T_i$ be the indicator for whether a CCC is placed in the lowest grade or highest grade. Then for the last-place aversion analysis sample, the treatment is determined as follows:

$$T_i = 1\{X_{ic} < 60\}.$$  

where the running variable $X_{ic}$ is the performance evaluation score of CCC $i$ who are evaluated in cohort $c$. For the first-place seeking analysis sample, the treatment is determined as follows:

$$T_i = 1\{X_{ic} \geq 90\}.$$  

Note that depending on the probability of receiving the treatment, the RDD is classified into the sharp RDD or the fuzzy RDD. Because the probability of receiving the lowest or highest grade, then one cannot argue whether the difference in the subsequent performance observed during the second evaluation period is driven solely by the last-place aversion or first-place seeking behavior.

One can argue that even if there are no explicit forms of negative or positive incentives of being placed in the lowest or highest grade, there might be implicit forms of incentives attached with the grade. For example, parents who are interested in sending their children to the CCC are likely to care about the quality of the CCC, and it is also likely that parents would not send their children to the CCC whose grade is very low. If this is the case, the fact that some CCCs who are placed in the lowest grade did better in the subsequent evaluation doesn’t necessarily imply that such CCCs are last-place averse. The improvement in the performance may have been driven by the information effect (or revealing effect). This issue does not pose any problem in our setting because even though governments do provide information on the performance of CCCs on government-operated website, they do not disclose grades of each CCC on this website. To put it differently, neither do parents know the grade of each CCC, nor do each CCC knows the grade of other CCCs.

Note that the Ministry of Welfare, in fact, planned on providing financial subsidies for those that received the rank of A and providing consulting to those that received the rank of F. However, the Ministry did not provide financial subsidies during our analysis period; i.e. 2012 and 2017. Regarding the consulting issue, we talked with the personnel in the Ministry, and they confirmed us that the consulting was extremely perfunctory, and so such consulting was less likely to have incentivized the CCCs.

<table>
<thead>
<tr>
<th>Grade</th>
<th>Evaluation score ($x$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>$x \geq 90$</td>
</tr>
<tr>
<td>B</td>
<td>$80 \leq x &lt; 90$</td>
</tr>
<tr>
<td>C</td>
<td>$70 \leq x &lt; 80$</td>
</tr>
<tr>
<td>D</td>
<td>$60 \leq x &lt; 70$</td>
</tr>
<tr>
<td>F</td>
<td>$x &lt; 60$</td>
</tr>
</tbody>
</table>

Table 1. Evaluation score by grade.
the treatment is equal to one when the evaluation score crosses the cutoff point, this study makes use of the sharp RDD.

To estimate the causal impact of the treatment under the sharp RDD setting, the conditional expectation functions of potential outcomes should be smooth across the running variable $X_{ic}$. Provided that the assumption holds, the average causal impact of the treatment $\tau_i$ is as follows:

$$\tau_i = \lim_{x \to 60^+} E[Y_{ic}|X_{ic} = x] - \lim_{x \to 60^-} E[Y_{ic}|X_{ic} = x]$$

for the last-place aversion analysis, and

$$\tau_i = \lim_{x \to 90^+} E[Y_{ic}|X_{ic} = x] - \lim_{x \to 90^-} E[Y_{ic}|X_{ic} = x],$$

for the first-place seeking analysis. In the equation above, $Y_{ic}$ denotes the performance outcome of CCC $i$ in cohort $c$.

In order to estimate the conditional expectation functions above, a researcher needs to estimate the conditional expectation functions near the boundary. That is, we need to estimate two conditional expectation functions: left and right of the cutoff. Many researchers propose using the local polynomial regression (e.g. Imbens & Lemieux, 2008; Lee & Lemieux, 2010). Also, the regression discontinuity estimates can be obtained by using local quadratic, cubic, or higher order specification. In our analysis, we used the local linear specification because of the point raised by Gelman and Imbens (2019). They recommend using local linear or quadratic polynomials. They also emphasize that researchers use lower-order polynomials if densities of variables are highly variable. As can be seen from the panels in Figure 3, the densities of scores left of the cutoff points are quite noisy. Moreover, according to Hahn et al. (2001), the local linear regression nonparametrically provides a consistent estimator for the treatment effect. We therefore estimated the local linear regression. Note that we also estimated the effect using local quadratic polynomials, and the effects are larger than the ones observed under the local linear specification.

While the local linear regression estimator provides a consistent estimator in the regression discontinuity design when the identifying assumptions of the regression discontinuity design are met (Gelman & Imbens, 2019), the statistical inference based on the conventional standard errors might be incorrect when the asymptotic properties are not met around the cutoff point. The regression discontinuity design, in general, makes use of observations near the neighborhood of the cutoff point arguing that the characteristics of the observations are similar between the left and right of the cutoff point. Accordingly, it is likely that the analysis sample around the cutoff point is too small for the asymptotic properties to hold. This scenario is salient in our setting especially at the 60 point cutoff that determines the lowest grade. There are about 30 CCCs who received the lowest grade. To obviate the problem induced by the finite sample, we conduct a randomization inference in the context of the regression discontinuity design developed by Cattaneo et al. (2015). When estimating the discontinuities, we provide estimates based on three bandwidth choices. We believe that providing effect estimates under several bandwidth choices is beneficial for securing transparency of the results.
V. Data

We use confidential data on all the CCCs evaluated during 2012 and 2017. The evaluation of the CCCs were conducted in three waves: 2012 vs. 2015, 2013 vs. 2016, and 2014 vs. 2017. The CCS that were evaluated in each wave were not the same across the waves. For example, in 2012, the number of CCCs that were subject to performance evaluation was about 560 and in 2015, the number of CCCs that were subject to performance evaluation was about 920. To analyze whether CCCs were sensitive to being placed in the lowest or highest rank, we restricted the analysis sample to CCCs that were observed in both years. We followed the same suit for the other two waves (i.e. 2013 vs. 2016 and 2014 vs. 2017). Furthermore, our analysis is based on the wave fixed effects so that the CCCs were compared within the wave.

The data are administered by the Evaluation Center of Community Child Center Central Support Group. In the dataset, there are rich information on the characteristics of the CCCs that can be used for testing the identifying assumptions of the regression discontinuity design. In Table 2, we provide descriptive statistics of the variables used in the analysis by the treatment status. For the last-place aversion analysis, the mean for the treated group is estimated using the CCCs whose baseline score is less than 60 (i.e. those whose grade is F). The mean for the control group is estimated using the CCCs whose baseline score is equal to 60 or above and less than 70 (i.e. those whose grade is D). For the first-place seeking analysis, the mean for the treated group is estimated using the CCCs whose baseline score is equal to or greater than 90 (i.e. those whose grade is A). The mean for the control group is estimated using the CCCs whose baseline score is less than 90 but equal to or greater than 80 (i.e. those whose grade is B).

As can be seen from Table 2, the two groups are very different from each other given that the difference in the means for most of the covariates. Table 2 implies that it is not appropriate to merely compare these two groups because their baseline characteristics are

<table>
<thead>
<tr>
<th>Table 2. Descriptive statistics.</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>% of CCCs whose rank improved</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Evaluation rank</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Evaluation score</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Age of CEOs</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>% of CCCs operating on weekend</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Share of CCCs receiving government subsidy</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Total number of enrolled children</td>
</tr>
<tr>
<td></td>
</tr>
</tbody>
</table>

Standard deviation in parentheses.
different, and it calls for a strategy to make the two groups similar. We use a regression discontinuity design so that we can compare the two groups with similar baseline characteristics but have two distinct performance grades.

VI. Results

A. Tests of identifying assumptions

Internal validity of the regression discontinuity design hinges critically on the continuity in baseline covariates at the treatment determining cutoff. We first provide graphical results as graphical analyses are focal in any regression discontinuity design application. We test the continuity in baseline characteristics based on the four variables (see Table 2). For the graphical analysis, we present the density of each variable by the running variable (i.e. baseline evaluation score). In the graph, we insert the local linear regression fit (i.e. the order of polynomial is equal to one) under the bandwidth choice of 9 for the right of the cutoff and the local linear regression fit under the bandwidth choice of 19 for the left of the cutoff.

Figure 1 presents the density of the baseline covariates by the running variable for the last-place aversion analysis. Panel A shows the density of the age of CEOs by the assignment variable. The estimated discontinuity at the 60 point cutoff is less than two years. Panel B tests for the share of CCCs that are operating on weekend. It is likely that a CCC operating their programs on weekend might be of higher quality than those not operating their programs on weekend. Hence, discontinuity in this baseline covariate is a sign that the two groups are dissimilar. As can be seen from Panel B, however, we do not see any significant discontinuity at the cutoff point. The difference in the share is about 5%. Another covariate we test for is the share of CCCs receiving subsidies from the government. Most of the share is 1, implying that most of the CCCs receive governmental subsidy. The discontinuity estimate for this variable is less than 5%. In Panel D, we examine the density of the total number of children enrolled in each CCC by the assignment variable. The discontinuity in the total number of children enrolled in each CCC observed for the cutoff point is about three children.

Figure 2 shows the density for the first-place seeking analysis. Panel A displays the density of the age of CEOs by the running variable. As can be seen from the figure, the density is extremely smooth across the running variable, and there are no observable discontinuity at the cutoff point. In Panel B, we display the density of the share of CCCs operating in weekend. On average, the share of CCCs operating in weekend is about 50%. The range of the share of CCCs operating in weekend is also quite small. Panel C presents the density of the share of CCCs receiving subsidy from the government. For the left of the cutoff, the estimated average in the share is about 85%. For the right of the cutoff, the estimated average in the share is also around 85%. Finally, the density of the total number of enrolled children is extremely stable across the running variable.

While the results from the graphical analyses indicate that observable differences between the two groups are small, graphical analyses are limited in testing for the statistical significance of the estimated discontinuity formally. We provide the results of the tests of continuity in baseline covariates for the last-place aversion analysis in Panel A of Table 3. We present discontinuity estimates derived from three different bandwidth choices. \( h = 3 \) indicates that the discontinuity is estimated by comparing the CCCs whose score is between
Figure 1. Baseline covariates by the running variable (cutoff = 60). Panel A: Age of CEOs, Panel B: % operating on weekend, Panel C: % receiving government subsidy, Panel D: Total number of enrolled children.

57 and 59 with the CCCs whose baseline evaluation score is between 60 and 62. We see, from Panel A of Table 3, that none of the covariates turned out to be statistically significant. The fact that the two groups are observably similar does not automatically imply that the two groups are unobservably similar. We argue, however, that because observable characteristics are similar, and given that the treatment assignment mechanism is as good as random because CCCs are not notified of their evaluation score, there are few reasons to believe that the two groups are unobservably dissimilar.

Panel B of Table 3 presents the tests of continuity in covariates for the first-place seeking analysis. Two points are noteworthy from Panel B. First, some estimates are statistically significant at the five percent level. We argue that this does not invalidate the identifying assumptions. Note that the estimated discontinuities for the ‘Age of CEOs’ and ‘Total number of enrolled children’ variables are quite small. The discontinuity estimate for the ‘Age of CEOs’ variable is only 1.3 years. This difference is, in our opinion, very small given that the average age of CEOs in our analysis sample is around 46 years (see Panel B of Table 2). Also, the estimated discontinuity for the ‘Total number of enrolled children’ variable is 1.68 children. This difference is quite small given that the average total number of enrolled children in the analysis sample is around 31.4 children
Panel A: Age of CEOs
Panel B: % operating on weekend
Panel C: % receiving government subsidy
Panel D: Total number of enrolled children

Figure 2. Baseline covariates by the running variable (cutoff = 90). Panel A: Age of CEOs, Panel B: % operating on weekend, Panel C: % receiving government subsidy, Panel D: Total number of enrolled children

(see Panel B of Table 2). Hence, while the discontinuity estimates are statistically significant, these are practically insignificant.

Second, the discontinuity estimates are relatively sensitive to the choice of bandwidth. Accordingly, it is quite difficult, from statistical perspective, whether the estimated discontinuities reflect true deviation. Rather, the sensitivity of the effect estimates reflects the noisiness of the densities of these covariates. Because most of the estimated discontinuities are practically and statistically insignificant, we believe that Table 3 points toward the fact that observable characteristics, and possibly unobservable characteristics of the two groups are qualitatively similar, especially around the cutoffs.

B. Outcome analysis

In order to test whether the CCC is sensitive to being placed in the lowest grade or first grade, we use three outcomes: i) A dummy variable indicating whether the grade of a CCC is improved in the subsequent evaluation; ii) The evaluation rank received by the CCC in subsequent evaluation; iii) The evaluation score received by the CCC in subsequent evaluation. For CCCs that received the rank of A, we defined the share improved
Figure 3. Outcomes by the running variable. Panel A: % improved (last-place), Panel B: % improved (first-place), Panel C: Evaluation rank (last-place), Panel D: Evaluation rank (first-place), Panel E: Evaluation score (last-place), Panel F: Evaluation score (first-place)

as those that received the rank A. For CCCs that received the rank of F, we defined the share improved as those that received the rank of D or higher. Figure 3 shows the density
of the share of CCCs that received a grade that is higher than the baseline evaluation grade.

For the last-place aversion analysis (the three panels on the left of Figure 3), we see that, on average, the CCCs who were placed in the lowest grade in the initial evaluation period performed better in the subsequent evaluation period than those who were placed just above the lowest grade. Also, we see a large discernible discontinuity in the share of CCCs with improvement at the 60 point cutoff. This discontinuity implies that compared with those who received grade D in the initial evaluation, more of the CCCs who received grade F in the initial evaluation improved in the subsequent evaluation.

For the first-place seeking analysis (the three panels on the right of Figure 3), the three outcomes are increasing in the baseline evaluation score. Note, however, that in the neighborhood of the 90 point cutoff, we see a little bit of dip in the densities right before the cutoff point, followed by a discernible jump in the density just right after the cutoff point, indicating that the CCCs who were placed in the highest grade are more likely to maintain their highest grade than those who were placed in the lowest grade, though the exact conclusion cannot be made using graphical analyses. Hence, we turn to estimating the exact discontinuity at the 90 point cutoff using local polynomial regression estimators.

We present the results of randomization inference for the last-place aversion analysis in Panel A of Table 4. We see that all the estimated discontinuities are statistically significant for the first outcome variable. For example, under the uniform kernel specification, the estimated discontinuities are 0.356, 0.298, and 0.318, indicating that the CCCs who are placed in the lowest grade performed significantly better than those placed just above the lowest grade. Hence, the evidence point toward the fact that the CCCs are
last-place averse. We also find statistically significant impact on the evaluation rank. The estimated discontinuity is quite large, though not all estimates turned out to be statistically significant. Nevertheless, given the large practical significance of the estimated discontinuities, we argue that being placed in the lowest grade positively affects the evaluation rank. The estimated impact on the evaluation score is also quite large, though the practical significance of effect estimates is sensitive to the choice of bandwidth.

The results of the randomization inference for the first-place seeking analysis are presented in Panel B of Table 4. The discontinuity estimates observed for the first outcome variable are around 0.10, indicating that the CCCs who were placed in the highest grade is about 10 percentage points more likely to secure their highest grade in the subsequent evaluation than those who are similar in observable characteristics but were placed just below the highest grade. Moreover, all the estimated discontinuities are statistically significant at the 1% level. We also see similar impact with respect to the subsequent evaluation rank. The estimated discontinuities observed for the second outcome variable are around 0.13. Furthermore, all the effect estimates are statistically significant. The discontinuity estimates observed for the subsequent evaluation score are also highly statistically and practically significant. On average, we find that the evaluation score is about 1 point higher for those who received the highest grade in the initial evaluation. All in all, the results presented in Table 4 are favorable for the validity of the last-place aversion and first-place seeking theory.

Note, however, that there is one factor that may induce bias in our estimated effects. That is, the last-place aversion or first-place seeking behavior might have been driven by the regression to the mean phenomenon because transitory noise may have contributed to the overstatement of the performance level of CCCs. While the mean reversion may have occurred in our setting, we provide two arguments that such reversion is less likely
in our setting. First, as Kane and Staiger (2002) note, the mean reversion is most likely when the analysis sample is consisted of a random draw from a population. Our setting does not suffer from such issue because our analysis sample is not consisted of a random draw from a population. Rather, all the CCCs that are evaluated in the first evaluation period are evaluated in the subsequent evaluation period. So we believe that our estimate is less likely to have been driven by the first issue that creates the mean reversion. Second, Chay et al. (2005) show that a regression discontinuity design is effective in controlling for the mean reversion. The logic behind their argument is that because those who are just below the cutoff point and those who are just above the cutoff point are similar in terms of unobservable characteristics including the noise that may contribute to the mean reversion, the difference between these two groups would cancel out the contribution of the mean reversion factor. Because our setting makes use of the RDD, we think that our estimate is less likely to have been driven by the second issue that induces the regression to the mean. Regardless, our setting may still be susceptible to the regression to the mean issue, so we caution on drawing a definitive conclusion about the impossibility of the mean reversion issue in our setting.

VII. Discussion and conclusions

We provide two implications from the empirical results. First, given the estimated results in the last-place aversion and first-place seeking analyses, we argue that the degree of the last-place aversion behavior is stronger than that of the first-place seeking behavior, at least in the performance evaluation system. The estimated impact regarding the last-place aversion is, on average, 35 percentage points, while the estimated effect regarding the first-place seeking is, on average, 10 percentage points. The difficulty of getting out of the lowest grade is not exactly the same as retaining the highest grade, so the comparison of the effect estimates between the two scenarios is not easy. But given that we are estimating the impact at the cutoff by comparing the two groups that are observably and unobservably similar, we believe that our setting is favorable for making such comparison.

Second, if policymakers were to adopt the behavioral insights such as the last-place aversion and first-place seeking theory, one must think carefully about the opportunity costs of being placed in the last place or first place. It is quite likely that if the opportunity costs of being placed either in the last place or first place are significant from the economic agents, then the effectiveness of such policy would be promoted to a great extent. We believe that examining the relative impact of the last-place aversion and first-place seeking depending on the opportunity costs faced by economic agents should be the focus for future research.

Regarding the policy implication, we think that the two policy implications can be drawn from our study. First, the government should disclose the list of CCCs that are in the lowest grade without attaching any consequences for being placed in the lowest grade. Many evaluation systems in the public sector attach negative incentives to those that are underperforming. Penalizing underperforming agents is criticized for inducing unintended negative consequences and being too stringent and unfair for those that lack resources to promote their performance (van Thiel & Leeuw, 2010). Therefore, merely disclosing the performance level of CCCs and resorting to behaviorally motivated insight such as last-


place aversion is more likely to be effective in preventing unintended consequences of penalizing CCCs.

Second, we argue that government should focus more on disclosing the list of CCCs that are in the lowest rank rather than disclosing the list of CCCs that are in the highest rank. Rewarding those who are outperforming is criticized for crowding out intrinsic motivation (e.g. Georgellis et al., 2011) and for being inefficient in terms of government budget because such reward is oftentimes not linked to improvement in subsequent evaluation (e.g. Bacolod et al., 2012).

This paper is not free of limitation. Above all, the external validity of the findings is worth mentioning. This study tests the validity of the two theories using the performance evaluation system adopted specifically for evaluating the CCCs. Because incentives faced by economic agents differ by the setting that the agents are placed, the findings may not be applicable to settings other than CCCs. Another limitation of our study is that we did not answer the question on whether incorporating behaviorally motivated insights into performance evaluation systems would promote the overall performance level of CCCs. We direct future research to engage in addressing such research question.

Notes

2. Note that there are three cohorts. The first cohort consists of CCCs who received the initial performance in 2012 and the next performance in 2015. The second cohort consists of CCCs who received the initial performance in 2013 and next performance in 2016. The third cohort consists of CCCs who received the initial performance in 2014 and the next performance in 2017.

Funding

This research was supported by the Chung-Ang University Research Grants in 2020.

Notes on contributors

Young-Gon Kim, Ph.D., is an Associate Research Fellow at the Korea Institute for Defense Analyses. He previously worked for the Korea National Open University as a researcher. His current research and research interests include organizational and manpower management, military culture and policy of human rights in the military. He is a graduate of Kyung Hee University (B.A, MPA and Ph.D. in Public Administration)

Hosung Sohn is an assistant professor in the School of Public Service at Chung-Ang University. His research area is in the field of public finance and program evaluation using experimental and quasi-experimental research designs.

ORCID

Hosung Sohn http://orcid.org/0000-0002-7031-4430
References


