Article

Urban Affairs Review
1–35

DOI: 10.1177/10780874251330256
journals.sagepub.com/home/uar

Does Organizational Performance Feedback Affect Public Service Outsourcing?

Jinhai Yu[1] and Xin Chen[2]

Abstract

Public service organizations are often put under performance-based accountability systems that offer organizational performance feedback. We examine whether and when public organizations use outsourcing to respond to negative performance feedback. Outsourcing can be a tool to improve future performance or escape accountability pressure. Further, public organizations with lower organizational capacity are more likely to outsource because they possess fewer resources to meet performance gaps through in-house production. Using data from the Financial Integrity Rating System of Texas, we conduct a regression discontinuity analysis of the impact of a lower performance rating on school districts' outsourcing. The results show that negative feedback increases the proportion of district expenditure on outsourcing. The effects are stronger for the districts with smaller tax bases and staff sizes. Performance-based accountability systems move public organizations toward greater use of outsourcing through negative performance feedback.

Keywords

performance feedback, outsourcing, organizational capacity, regression discontinuity design

[1]University of Connecticut, Hartford, USA
[2]Sun Yat-Sen University, Guangzhou, China

Corresponding Author:
Xin Chen, Center for Chinese Public Administration Research, School of Government, Sun Yat-Sen University, No. 135 Xingang Xi Road, Guangzhou, 510275, China.
Email: xinchen1910@gmail.com

Introduction

Governments have widely implemented performance-based accountability systems to strengthen performance management and monitoring in various policy areas, such as the test-based performance ratings under the No Child Left Behind Act (Hannaway and Hamilton 2008) and the United States managing for results reforms (Moynihan and Ingraham 2003). Public organizations not only receive regular updates on their performance but also face increasing accountability pressure when the performance information is released to the public. The so-called performance feedback has been shown to affect a wide range of organizational outcomes, including goal setting (Nielsen 2014), managerial innovations (Nicholson-Crotty, Nicholson-Crotty, and Fernandez 2017; Salge 2011), and red tape (Hong 2020). However, less is known about whether performance feedback shapes how public organizations choose the methods of public service deliveries, that is, in-house production or outsourcing. Meanwhile, the burgeoning literature on public service outsourcing has attended to the relationships between organizational performance and outsourcing (e.g., Bel, Fageda, and Warner 2010; Lee et al. 2019; Warner and Aldag 2021). Still, the empirical evidence on the effect of organizational performance on outsourcing remains inconclusive. Bridging the two strands of literature, we examine the impact of organizational performance feedback on public service outsourcing in this study.

We expect negative performance feedback to improve organizational performance and reduce accountability pressure, thus increasing outsourcing. Negative performance feedback, which shows that organizations have been performing below expectations, should prompt organizations to identify problems and search for solutions to remedy the performance gaps (Cyert and March 1963; Simon 2013). Outsourcing is an invaluable tool to achieve these ends because it has long been advocated to save costs, increase efficiency, and improve performance (e.g., Bel, Fageda, and Warner 2010; Petersen, Hjelmar, and Vrangbæk 2018). Further, negative performance feedback can increase accountability pressure and motivate blame avoidance. We expect a stronger impact of negative performance feedback on outsourcing for public organizations with lower organizational capacity. The lower the organizational capacity, the more likely public organizations will face constraints in what they can afford with existing resources and thus resort to outsourcing as a solution.

We test these propositions empirically using data from the Financial Integrity Rating System of Texas (FIRST). The FIRST system is a performance-based accountability tool for school districts, which rates their financial management performance and has made the performance ratings public since 2001. The performance ratings are based on numerical scores

covering financial conditions and financial management. This performance-based rating system targets how well a school district manages its financial resources. It differs from the test-based ratings based on student academic achievements in Texas (Craig, Imberman, and Perdue 2013). We apply the regression discontinuity (RD) design as a quasi-experimental method to estimate the impact of a lower performance rating on school districts' outsourcing. We find that a lower rating causes the proportion of district expenditure on outsourcing to increase by about 6.8 percent to 8.7 percent for an average school district in the subsequent 5 years. School districts with lower fiscal and human resource capacity, as indicated by smaller tax bases and staff sizes, see stronger effects of lower ratings.

We advance the literature by showing how organizational performance feedback under performance-based accountability systems shapes the methods of public service provision. First, we focus on public service outsourcing, while the literature on performance feedback has examined various outcomes on organizational performance or accountability. This enriches the behavioral theory of organizations (Cyert and March 1963; Simon 2013) by adding outsourcing as a distinct outcome. This study also shows that outsourcing is one critical way public organizations respond to negative performance shocks or use performance feedback information (Anderson and Nielsen 2020; Holm 2018).

Second, we add to the outsourcing literature by estimating the causal effect of organizational performance feedback using RD as a quasi-experimental method. The FIRST case offers a valuable avenue to measure financial management quality, which deviates from the outsourcing literature on fiscal performance or financial condition indicators (e.g., Bel and Fageda 2017; Fernandez, Ryu, and Brudney 2008; Zhang, Gibson, and Schafer 2018). Moreover, the literature has shown that the correlational relationship between organizational performance and outsourcing can go both ways. We use RD as a quasi-experimental design to tackle endogeneity issues in estimating the effect of organizational performance on outsourcing. This is particularly valuable, given that organizational performance has been widely recognized as a determinant of outsourcing while the empirical evidence remains inconclusive.

Third, we probe the conditions under which organizational performance feedback affects outsourcing by focusing on organizational capacity as a crucial contextual factor. Organizational capacity is particularly relevant because it indicates public organizations' resource constraints and choice sets when weighing in-house production and outsourcing as alternatives to public service delivery. We show that public organizations with lower fiscal and human resource capacities are more likely to outsource in response to negative performance feedback. Our study lays the foundation for future

research that may simultaneously use exogenous variation in performance feedback and organizational capacity, such as an experimental design.

Previous Studies

Performance Feedback and Organizational Response

Performance feedback is a critical part of performance management that concerns how an organization's past performance shapes managers' decision-making and future performance (Hong 2019). Min et al. (2021) identified three types of performance feedback: internal, external, and historical and social comparative feedback. In a typical performance-based accountability system, feedback on organizational performance is provided through an external performance rating system and made public. Examples include the student achievement ratings under the No Child Left Behind Act (Hannaway and Hamilton 2008), the performance evaluations of public agencies in Korea (Hong 2019), and the FIRST. In such contexts, performance feedback increases performance information and accountability pressure because it is not just for internal managerial use but also widely publicized.

The impact of performance feedback under performance-based accountability systems on organizational responses has received increasing scholarly attention. Organizational responses to performance feedback are often understood through the lens of the behavioral theory of the firm following Simon and March (Cyert and March 1963; Simon 2013). The theory predicts that organizations will search for solutions to address problems as indicated by negative performance feedback, which has received support from experimental evidence (Van der Voet 2023; Van der Voet and Lems 2022).

Scholars have focused on negative performance feedback and a wide range of organizational outcomes, including goal setting (Nielsen 2014), managerial innovations (Nicholson-Crotty, Nicholson-Crotty, and Fernandez 2017; Salge 2011), red tape (Hong 2020), budgetary adjustments (Flink 2019), and performance improvements (Hong 2019; Wei et al. 2023). Still, in this strand of literature, scholars have not considered how performance feedback affects organizational decisions regarding public service outsourcing. Outsourcing can be a legitimate candidate for solution search following negative performance feedback because it is widely recognized for cost savings and performance improvements (e.g., Bel, Fageda, and Warner 2010; Petersen, Hjelmar, and Vrangbæk 2018). By linking performance feedback and public service outsourcing literature, we address this research lacuna by making the case that receiving negative performance feedback shapes how public organizations deliver services.

Organizational Performance and Public Service Outsourcing

There has been extensive literature on the determinants of outsourcing in the public sector. One set of factors is related to economic efficiency. Scholars have demonstrated that outsourcing is associated with transaction costs: the lower transaction cost a service has, the more likely it is to be contracted out (Brown and Potoski 2003a; Hefetz and Warner 2012). Scholars have also shown that local governments with a non-competitive market cannot easily attract external vendors due to their small size and sparse population (Girth et al. 2012; Mohr, Deller, and Halstead 2010). Another body of research explores the influence of political factors on outsourcing decisions, finding that politics plays a significant role (Wei, Zhang, and Yang 2022). For instance, both conservative and right-leaning citizens and policymakers favor outsourcing more services (Alonso and Andrews 2020; Guo and Willner 2017; Fernandez, Ryu, and Brudney 2008).

The research highlighting fiscal constraints as a determinant is most relevant to our study. Scholars have examined the roles of the tax burden (Boyne 1998; Warner, Aldag, and Kim 2021), revenue diversification (Zhang, Gibson, and Schafer 2018), fiscal dependency (Geys and Sørenson 2016), and taxing authority restrictions in outsourcing (Bel and Fageda 2017; Fernandez, Ryu, and Brudney 2008). These studies find that outsourcing is more likely to occur when local governments face greater fiscal stress and fiscal constraints. While these studies have addressed the impact of financial condition as indicated by fiscal or financial variables, few have examined the role of financial management quality, which is more challenging to measure. We focus on the feedback on financial management quality and the resultant accountability pressure when such performance information is made publicly available, as in the case of FIRST.

Further, scholars have attended to the relationship between organizational performance and outsourcing. Much of the research focuses on the impact of outsourcing on organizational performance and finds that outsourcing enhances performance, including the efficiency and effectiveness of public organizations (e.g., Bel, Fageda, and Warner 2010; Elkomy, Cookson, and Jones 2019), governmental financial performance (Park and Noh 2021), and public employees' perception of performance (Lee et al. 2019). Others have examined the impact of organizational performance on outsourcing. For example, recent studies find that poor contracting outcomes, such as poor quality and high transaction costs, lead to contracting back in (Warner and Hefetz 2012; Warner and Aldag 2021). This suggests that the performance of past outsourcing affects future contracting decisions. Yet, it remains unclear whether external feedback on organizational performance matters.

Finally, the empirical evidence on the impact of fiscal constraints and organizational performance on outsourcing remains inconclusive due to omitted variable biases and possible two-way causation. Unobserved or unmeasured factors, like managerial quality, may affect both organizational performance and outsourcing. As Bel and Fageda (2017, 504) note in their review of outsourcing literature: "In terms of the methodologies employed, we observed a potential problem of reverse causality affecting most existing studies … dealing with this potential problem of endogeneity should be the objective for subsequent research." Answering this call, we join a few recent studies (Alonso and Andrews 2020; Hansen, Bel, and Petersen 2022) to estimate the causal effects of financial management performance on public service outsourcing by applying an RD design as a quasi-experimental method.

Performance Feedback and Outsourcing: Why and When It Matters

Solution Search, Blame Avoidance, and Outsourcing

The logic of the behavioral theory of the firm, as adapted to public organizations, is simple (Cyert and March 1963; Simon 2013). Public organizations are boundedly rational and follow a "satisficing" principle. What is "satisfactory" depends on some reference points, either historical or social comparative. Negative performance feedback indicates that organizations have performed below their aspirational levels. This will trigger a problem definition and solution search process (Van der Voet and Lems 2022; Van der Voet 2023). Negative performance feedback can make organizations more risk-seeking and innovative in seeking solutions to improve performance (Baum and Dahlin 2007). Scholars have applied the behavioral theory of the firm to various organizational responses to performance feedback, such as organizational innovation (Nicholson-Crotty, Nicholson-Crotty, and Fernandez 2017) or performance improvements (Hong 2019).

Public service outsourcing is an organizational response that can be understood through the lens of the behavioral theory of the firm. When public organizations find their performance is less than satisfactory, as negative performance feedback indicates, they may search for options to improve future performance. Among others, one viable solution is to (re)consider the choices of public service delivery between in-house production and outsourcing. Outsourcing has long been advocated to save costs, increase efficiency, and enhance performance (Bel, Fageda, and Warner 2010; Petersen, Hjelmar, and Vrangbæk 2018). The public choice theory views traditional bureaucratic means of providing services as inherently inefficient because government monopoly inhibits the incentives to improve service quality

(Boyne 1998; Brown and Potoski 2003a). Outsourcing creates a competitive service delivery environment by introducing vendors bidding for government contracts.

Service outsourcing may become a particularly relevant organizational response when the performance feedback is focused on financial management because of the common concern about costs, as in the FIRST system. Public organizations with negative feedback on financial management are more likely to seek improvements in the financial condition, of which cost containment is a critical strategy. Outsourcing may become a salient means because it can save costs. As Fernandez, Ryu, and Brudney (2008, 447) point out, "potential cost savings and gains in efficiency should be particularly appealing to public agencies experiencing fiscal stress." More broadly, the literature has shown that public organizations are more likely to outsource under fiscal strains (Geys and Sørenson 2016; Warner, Aldag, and Kim 2021; Zhang, Gibson, and Schafer 2018).

In addition, the impact of negative performance feedback on outsourcing can be understood through the lens of accountability. The principal-agent perspective on accountability suggests negative performance feedback reduces information asymmetry and increases accountability pressure (Coats 2002; Heinrich, Lynn, and Brinton Milward 2010). For instance, James (2011) shows that reporting information about low performance decreases perceived performance and citizen satisfaction. James and Moseley (2014) find that negative performance information increases citizens' likelihood to attribute blame to local governments. Alternatively, the reputation-based perspective on accountability argues that reputation investment drives accountability holding and giving (Busuioc and Lodge 2016, 2017). Outsourcing may be used as a proactive strategy to enhance reputation because it has been associated with such positive images of cost-saving and efficiency gains. Therefore, outsourcing can offer not only a practical way to reduce accountability pressure but also a symbolic response for reputational benefits.

Specifically, outsourcing can be perceived as an effective strategy to shift the blame for low performance to external vendors. Theories of blame attribution suggest that citizens consider the level of control when attributing blame for low service performance (James et al. 2016; Leland, Mohr, and Piatak 2021). They are more likely to blame those who directly control the service delivery (Piatak et al. 2017). By outsourcing to external vendors, public organizations can shift the blame to external vendors because the latter has direct control over service provision. While some studies find no effects (James et al. 2016; Marvel and Girth 2016), a wide range of recent experimental evidence has shown the effectiveness of outsourcing in reducing blame on the government for service failures (Hung and Porumbescu 2024; Leland, Mohr, and Piatak 2021; Mohr et al. 2023; Walker et al. 2024).

Thus, public organizations may consider outsourcing a valuable strategy to reduce blame for low performance (Bach and Wegrich 2019).

Beyond blame-shifting, outsourcing can reduce the accountability pressure for public organizations by hindering responsibility attribution. Outsourcing shifts the responsibility of service provision to vendors and "blurs lines of responsibility attribution" (Hood 2011; Hung and Porumbescu 2024). Outsourcing can diffuse accountability to multiple actors, making attributing responsibility for poor performance more difficult (Bach and Wegrich 2019). For instance, Dawkins (2021) finds that outsourcing makes it "less likely to attribute responsibility for public services to government." Outsourced services may face less demanding accountability requirements than those produced in-house (Marvel and Marvel 2007; Rich 2022). Consequently, outsourcing can alleviate the accountability pressure for public organizations even when blame-shifting is not a concern.

Nevertheless, some may be skeptical of outsourcing as a means to avoid accountability. For instance, Milward and Provan (2000, 364) argue that "One does not diminish one's responsibility by paying someone else to do the work … Privatization only changes the venue, not the public responsibility." This is more likely to occur if citizens have complete information about the service performance and whether the government or service provider is in control. Thus, they can efficiently attribute blame and accountability for outsourced services (Piatak et al. 2017). With incomplete information and bounded rationality, while elected officials may eventually be blamed (James et al. 2016; Marvel and Girth 2016), blame can be directed at service providers first. If so, outsourcing can still be a viable strategy to reduce blame and accountability pressure.

To summarize, negative performance feedback can motivate public organizations to use service outsourcing to close performance gaps. Outsourcing can also allow public organizations to avoid being blamed for low organizational performance and reduce accountability pressure. Therefore, we propose the first hypothesis.

H1: Negative performance feedback increases public service outsourcing.

Organizational Capacity, Performance Feedback, and Outsourcing

While the literature shows that public organizations respond to performance feedback, the responses are not created equal. Research suggests that organizational capacity is critical in how public organizations respond to negative performance feedback (Drolc and Keiser 2021; Joaquin and Greitens 2011). Organizational capacity is a multidimensional concept that generally refers to the ability of an organization to fulfill its mission (Eisinger 2002).

It covers multiple aspects, including fiscal resources, human resources, leadership quality, and community capacity (Andrews et al. 2013; Gargan 1981). To further develop testable hypotheses, we consider two dimensions of organizational capacity—the availability of fiscal and human resources (Drolc and Keiser 2021; Hall 2008; Igalla, Edelenbos, and van Meerkerk 2020).

Organizational capacity is a crucial conditioning variable for public organizations to outsource to respond to negative performance feedback. In one sense, organizational capacity represents a "budget constraint" for organizations choosing between in-house production and outsourcing. The literature suggests that outsourcing provides a low-cost alternative to in-house production (e.g., Bel, Fageda, and Warner 2010; Petersen, Hjelmar, and Vrangbæk 2018). Organizations with higher capacity are more likely to address the performance gaps and deal with the accountability pressure with in-house production because of a more lavish budget. By contrast, organizations with lower capacity face more stringent constraints in what they can afford with existing resources and thus are more likely to resort to outsourcing as a solution.

Fiscal resource availability represents a crucial facet of organizational capacity. Public organizations without adequate fiscal resources can struggle to implement programs or policies, impeding their efforts to improve performance (Boyne 2003; Hall 2008). Empirical studies have shown that public organizations outsource services to reduce financial strain (Kim and Warner 2016; Warner, Aldag, and Kim 2021). Specifically, outsourcing can alleviate fiscal stress in at least two ways. First, outsourcing has been advocated as a strategy for cost savings and efficiency improvements through market competition (Bel, Fageda, and Warner 2010; Petersen, Hjelmar, and Vrangbæk 2018). Second, outsourcing may reduce fiscal burden by shifting legacy costs such as pension obligations to vendors (Fernandez, Smith, and Wenger 2007; Rho 2013). As noted by Traynor, Chen, and Kitterlin (2019, 79) in their interview with Florida school district managers, "[outsourcing] removed some liabilities that the school district had. It removed some cost-related retirements that the school board would have to pay. It gives us great flexibility in finance."

When negative performance feedback strikes, public organizations with lower fiscal capacity are less likely to take fiscally "expensive" steps to meet performance gaps. They are more likely to shift from in-house production to outsourcing as it is perceived to be more affordable. To the extent that public organizations with lower fiscal capacity are more sensitive to cost savings and reduction in fiscal burden, they are more likely to use outsourcing to respond to negative performance feedback. Hence, we test the following hypothesis:

H2: Negative performance feedback has a larger effect on service outsourcing when public organizations have lower fiscal capacity.

Furthermore, human resource availability is a critical element of organizational capacity. Public organizations with adequate personnel are better positioned to change organizational behaviors in response to poor performance (Moynihan 2008; Drolc and Keiser 2021). They have the staff resources to implement necessary organizational reforms. Empirically, Drolc and Keiser (2021) showed that public organizations with more staff are more likely to improve performance in response to negative performance feedback. Public organizations with ample human resources can offer more in-house service production and thus are less likely to outsource to respond to negative performance feedback.

Conversely, public organizations lacking human resources are more likely to outsource to maintain service levels. Service outsourcing has the potential to help maintain service levels through cost savings and improved efficiency without investing in staff resources heavily (Bel, Fageda, and Warner 2010; Petersen, Hjelmar, and Vrangbæk 2018). Due to civil service rules, public organizations face more constraints in hiring new employees to increase human resources. Public organizations can outsource services to complement in-house personnel. The lower the human resource capacity of a public organization, the stronger its demand for outsourcing in response to negative performance feedback. Thus, we propose the following hypothesis:

H3: Negative performance feedback has a larger effect on service outsourcing when public organizations have lower human resource capacity.

Data and Sample

Data

We use data from the FIRST in the empirical analysis. Since 2002, the FIRST system has held school districts accountable for their financial management performance. The FIRST system covers various indicators of the school district's financial condition and financial management. For financial conditions, districts are rated based on financial solvency indicators such as debt-to-asset ratio and general fund balance. On the financial management dimension, districts are evaluated on such procedural grounds as how well they follow financial disclosure requirements, whether they make timely debt payments, or whether they discuss the impacts of budget changes before budget adoption. Overall, the FIRST system features a management-based rating scale that

focuses on the internal financial management of school districts, which differs from the ratings based on student scores or academic achievements in Texas (Craig, Imberman, and Perdue 2013).

Since 2007, the FIRST system has assigned numerical scores[1] from 1 to 100 to school districts based on a predetermined grading scale. The score is then converted to one of four ratings for each district, including "A = superior achievement," "B = above standard achievement," "C = meets standard achievement," and "F = substandard achievement." F means that school districts fail to meet the minimum standard of financial management. The ratings are released to the public on the Texas Education Agency (TEA) website around August each year. FIRST assigns total performance scores to districts based on 15–20 performance indicators. The performance indicators are categorized as capping, critical, and regular, each carrying a different weight in determining the final performance score. Districts may have imperfect control over performance scores on some indicators. For example, tax revenues or fiscal surpluses are only partly within the control of districts because they also depend on the tax base and the local economy.

In the financial management within school districts, superintendents play a pivotal role as the top public managers (Davidson, Schwanenberger, and Wiggall 2019; Rho 2013). Operating under the oversight of school boards, they are responsible for daily operations, the preparation and administration of the district's budget, and the organization and hiring of central administration staff. The superintendent typically has a significant role in appointing the finance manager, although this can vary depending on the district's specific governance structure. In a typical school district's budget, instructional expenditures represent the largest budgetary allocation, accounting for ~60 percent of total expenditures. Administrative and support services, as well as plant services, each command over 10 percent of overall budget allocations. Sound financial management can lead to enhanced resource allocation and potential improvements in academic performance for school districts, and it can positively affect superintendents' reputations and career prospects (Grissom and Andersen 2012; Grissom and Mitani 2016). Conversely, poor financial management can result in budget cuts, reduced staffing, or even state intervention, damaging a superintendent's reputation and career advancement.

The FIRST system features a typical performance-based accountability design because the FIRST ratings contain both performance and accountability components. The performance dimension results from the indicators that evaluate school districts' financial health and their performance in the financial management process. The accountability dimension is manifested in publishing the performance ratings to the general public. The disclosure of the FIRST ratings means they are not only for internal managerial use but also

remain accessible to external stakeholders such as parents, voters, taxpayers, or news media. The accountability side distinguishes the FIRST system from the performance management or measurements intended for managerial use only.

The accountability pressure under the FIRST system can come from several sources. First, school districts with high ratings can be rewarded, and those with low ratings can be sanctioned. For instance, the Texas Education Code stipulates that the state will consider the FIRST ratings in the designation for outstanding performance for school districts. Meanwhile, the districts with low ratings face additional compliance burdens because they must submit corrective action plans to address the performance gaps identified in the FIRST ratings. Second, within two months of the release of the performance ratings, school districts must "hold a public meeting to distribute a financial management report that explains the rating and its performance".[2] The public meetings can further expand the reach of the information contained in the performance ratings to the local community.

Sample

We collect the FIRST and outsourcing data from TEA between 2012 and 2020.[3] The FIRST data contain school districts' performance scores, performance ratings, the thresholds for each rating, and the rating releasing dates. We use a categorical variable from 1 to 4 to indicate the four performance ratings—F, C, B, and A—respectively. We construct the dependent variable *percentage outsourced* as the percentage of the school district budgetary expenditure on outsourced services, following prior studies on school district outsourcing (e.g., O'Toole and Meier 2004; Rho 2013). This measurement covers outsourcing in various service areas or contractor types. As a continuous variable, the measurement captures the degree of outsourcing a school district uses.

Outsourcing constitutes an average of about 9.77 percent of total expenditures in the sample. This proportion of outsourcing is consistent with findings from similar studies (Rho 2013), emphasizing its significant role in school district financial management. Figure A1 in the Appendix shows the average percentage of outsourcing across years, which shows a slightly upward but stable trend from 2012 to 2020. Among the districts in the sample, Harmony district has the lowest level of outsourcing, with only 0.79 percent of its expenditures going to outsourced services. In contrast, the Port Aransas district has the highest, with 63.98 percent of its spending outsourced. Outsourced services typically include professional services such as medical, maintenance, and repair. Additionally, school districts may outsource instructional services and staff training programs to other

Table 1. Summary Statistics.

Variables	(1) *N*	(2) Mean	(3) SD	(4) Min	(5) Max
Percentage outsourced	19,096	9.77	4.33	0.79	63.98
Negative feedback	19,096	0.20	0.40	0	1
Performance rating	19,096	3.56	0.72	1	4
Rating A	19,096	0.66	0.47	0	1
Rating B	19,096	0.25	0.44	0	1
Rating C	19,096	0.05	0.22	0	1
Rating F	19,096	0.03	0.16	0	1
Number of staff per pupil	19,096	0.17	0.05	0.07	0.63
Taxable property values per pupil(log)	19,096	12.92	0.89	10.16	17.01

Note. We measure the performance ratings with a categorical variable ranging from 1 to 4, representing the four ratings of F, C, B, and A, respectively.

public schools or institutions that provide specialized educational programs, such as career and technology education or training for staff. They may contract with private firms, nonprofit organizations, or other school districts.

Further, we create two moderator variables to test the moderating effects of fiscal and human resources. First, we construct a continuous variable that equals taxable property values per pupil (log) to measure districts' fiscal resource capacity. Taxable property value represents the primary tax base for school districts and is a suitable measurement for school districts' fiscal capacity (Martinez-Vazquez and Jameson Boex 1997; Nelson and Balu 2014). For example, about 80 percent of the Texas districts' revenues came from property taxes in 2017, making it the single largest source of revenue, according to TEA.[4] Property-wealthy school districts can raise more revenues than other districts with the same tax rate. They also have greater potential to collect more financial resources to the degree that property tax increases are allowed.

Second, consistent with prior studies (AbouAssi and Jo 2017; Kim and Peng 2018; O'Toole and Meier 2010), we use school districts' number of administrative staff[5] per pupil to indicate human resource capacity. A larger number of staff per pupil reflects a higher stock of human resources, which is critical for public organizations to implement programs or policies.

Table 1 shows the summary statistics. The dependent variable, the *percentage outsourced,* has a mean of 9.77 and a standard deviation of 4.33. The key variable of interest, *negative feedback*, has a mean of 3.56 and a standard deviation of 0.72. One moderator variable, the *taxable property values per pupil* (*log*), has a mean of 12.92 and a standard deviation of 0.89. Another moderator variable, the *number of staff per pupil*, has a mean of 0.17 and a standard deviation of 0.05.

Empirical Strategy

The RD Design

In the empirical analysis, we seek to estimate the causal impact of performance feedback on school districts' outsourcing. The performance feedback can be endogenous to outsourcing due to omitted variable biases or reverse causation. For instance, school districts with superintendents caring more about efficiency may have more robust financial management performance and heavier use of outsourcing. This may create an omitted variable bias if the motivation for superintendents to improve efficiency is not observed or measured. Moreover, the direction of causality may go from outsourcing to organizational performance because districts that pursue more service outsourcing may have better financial management performance if outsourcing saves costs or improves efficiency.

To address the endogeneity concerns, we employ an RD design to estimate the causal impact of a lower performance rating on school districts' outsourcing. The RD design hinges on the assumption that subjects cannot precisely manipulate the treatment assignment in the vicinity of the cutoff on a running variable (Lee and Lemieux 2010). If so, the subjects right below the cutoff constitute valid control groups for the subjects right above the cutoff. Because subjects cannot self-select into treatment status in the narrow range around the cutoff, the treatment and the control groups should be balanced on both observable and unobservable characteristics. Consequently, the RD design contrasts with a conventional regression approach that relies on a list of control variables to reduce omitted variable bias. Adding control variables in an RD design does not help to obtain unbiased estimates, but it may increase efficiency and lead to more precise estimates (Lee and Lemieux 2010).

Our RD design deviates from a conventional RD model in two ways. First, we employ the RD design with multiple cutoffs (Cattaneo et al. 2016). In the FIRST system, the cutoff points are the thresholds of the numerical performance scores that determine the performance ratings. The FIRST system has three cutoffs that generate four performance ratings. The school districts whose performance scores fall between two rating cutoffs may be viewed as either treated by one rating or untreated by the other rating. To categorize school districts into treatment and control groups, we use the middle point of the differences between two cutoffs as the bandwidth,[6] following Cattaneo et al. (2016). For instance, suppose the cutoff for an A rating is 90, and the cutoff for a B rating is 80. A school district with a performance score of 87 that receives a B rating could be viewed as treated by the B rating or untreated by the A rating. In this case, the bandwidth is 5 because

the difference between the two cutoffs is 10. The A rating will include performance scores between 95 and 85, and the B rating will include those between 85 and 75. Thus, the school district with a performance score of 87 belongs to the control group of the A rating because its score falls within the bandwidth of the A rating.

As is common in a multi-cutoff RD design, we normalize the cutoffs by subtracting the closest cutoff score from the performance scores and pooling the results for multiple cutoffs together (Cattaneo et al. 2016). We use the pooled sample for the main analysis and then estimate the RD effects for the three rating cutoffs separately to explore the heterogeneity of effects. The running variable is the difference between the performance scores and the closest cutoffs. The treatment variable, *negative feedback*, is defined as a dummy equaling 1 when the running variable is negative and 0 otherwise.

Second, we use the dynamic RD design to account for the timing of outsourcing in response to negative performance feedback. School districts may take time to respond to the performance rating changes by outsourcing. Theory suggests that public organizations may not attend to performance gaps if they are within their "zone of acceptance" (Meier, Favero, and Zhu 2015). Further, the timing of outsourcing may depend on the supply and demand sides, where school districts must search for suitable external vendors for their services. It is thus unclear ex-ante how long it takes from the negative performance feedback to the outsourcing decisions. We apply a dynamic RD design to account for the potentially changing effects of negative feedback on outsourcing over time following Cellini, Ferreira, and Rothstein[7] (Cellini, Ferreira, and Rothstein 2010). Specifically, we use the following model specification to estimate the effects of negative feedback on outsourcing over five[8] years.

$$y_{i,t+\tau} = \theta_\tau d_{it} +_\tau f(v_{it}) + \alpha_\tau + \beta_t + \gamma_{it} + \varepsilon_{i,t+\tau}, \quad (1)$$

where $y_{i,t+\tau}$ is the service outsourcing of district i in τ years after negative feedback in year t. On the right-hand side, d is an indicator for whether a school district i has a lower performance rating in year t, and θ is the coefficient of interest, which captures the dynamic effects of negative feedback on district outsourcing. For example, θ_2 captures the negative feedback effects two years after the performance rating of a school district. Moreover, $f()$ is a linear polynomial function[9] of the running variable v_{it}. In addition, α, β, and γ represent the fixed effects for years relative to negative feedback, calendar years, and negative feedback, respectively. The γ_{it} effect absorbs any across-district variation. Finally, ϵ is the error term.

As in the extant studies (e.g., Cellini, Ferreira, and Rothstein 2010; Kogan, Lavertu, and Peskowitz 2017), we stack the observations around each performance rating over the period from $t-2$ to $t+5$ to implement the RD design.

For instance, a school district receiving a rating in 2014 will include all observations for 2012–2019. We combine the stacked datasets for each performance rating into a single dataset covering the sample period from 2012 to 2020. Following the literature (e.g., Cellini, Ferreira, and Rothstein 2010), we assume that there are no anticipation effects of negative feedback on district outsourcing in the vicinity of the rating cutoffs by constraining θ to 0 when $\tau < 0$ but allow it to vary otherwise.[10]

Moreover, to estimate the moderating effects of organizational capacity, we modify Equation (1) by pooling together all five years after the negative feedback and estimating the average effects of the negative feedback. We then estimate the interactive effects between organizational capacity and negative feedback as follows[11]:

$$y_{it} = d_{it} + \delta f(v_{it}) + d_{it}M_{it} + M_{it}f(v_{it}) + \varphi M_{it} + \beta_t + \gamma_{it} + \varepsilon_{it}, \quad (2)$$

where M_{it} represents the moderator variables, and the coefficient of interest σ shows the moderating effects of M_{it} on outsourcing. The estimation of Equation (2) is based on the same sample as Equation (1).

Validity of the RD Design

The validity of an RD design depends on the assumption that the treatment assignment is as good as random near the cutoff (Lee and Lemieux 2010). We take three measures to assess the validity of the RD design. First, we examine the distribution of the performance scores near the cutoff. We expect to find a smooth distribution at the cutoff if the RD design is valid. As shown in Figure A2 in the Appendix, there are no discernable jumps in the distribution of performance scores around the cutoff. Second, we conduct a formal density test proposed by Cattaneo, Jansson, and Ma (2020). The estimated discontinuity is 0.70, with a p-value of .48. This suggests that the density distribution is continuous, supporting the validity of the RD design. Finally, we test the balance of multiple covariates measuring school district characteristics by the treatment status. We expect no statistically significant differences if the RD design is valid. As shown in Table A2 in the Appendix, the covariates are balanced.

Before presenting the results of the econometric analysis, we plot the average levels of outsourcing by negative feedback in Figure 1. The dashed vertical line shows the centered cutoff. The school districts on the right of the vertical line constitute the treatment group; the districts on the left represent the control group. We fit a linear curve on each side of the cutoff. Figure 1 shows that service outsourcing is higher for the treatment group school districts than the control group. This graphic evidence is consistent with H1 that negative feedback increases service outsourcing.

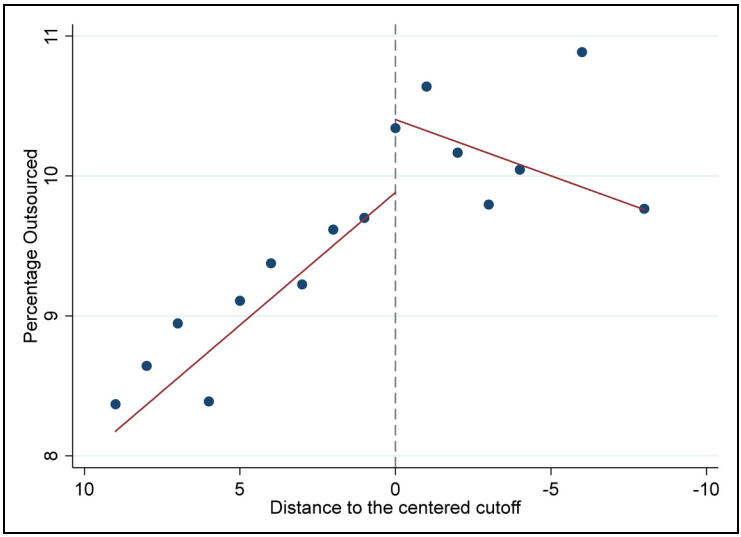

Figure 1. The distribution of service outsourcing by negative feedback.

Results

Baseline Results

Table 2 shows the effects of negative feedback on outsourcing estimated based on Equation (1). Column (1) of Table 2 shows the results for the full sample. Columns (2)–(4) show the results estimated by the performance rating. Each row of Table 2 shows the dynamic effects of negative feedback in the subsequent five years and then the average effects in that period.

As column (1) of Table 2 shows, negative feedback has positive and statistically significant effects ($p < .05$) on service outsourcing in all but the first years after the negative feedback. The sizes of the effects range from 0.66 to 0.85. The average effect of negative feedback in the five years is 0.475 and statistically significant ($p < .01$). Thus, negative feedback causes an increase between 0.66 and 0.85 percentage points in the subsequent two to five years for school districts. Since the dependent variable of service outsourcing has a sample mean of 9.77, these effects amount to an increase of 6.8 percent to 8.7 percent for an average school district. Based on the average total expenditure of 42 million dollars in the sample, negative feedback results in additional spending of over 2 million dollars on service outsourcing. The size of this effect is economically meaningful and comparable to prior studies on outsourcing expenditures (Rho 2013).

In column (2), negative feedback shows positive and statistically significant effects on service outsourcing in all five subsequent years for the subsample of

Table 2. Effects of Negative Feedback on Service Outsourcing Over 5 Years.

Relative years after negative feedback	(1) Full sample	(2) A to B	(3) B to C	(4) C to F
1	0.108	0.500*	−0.239	−0.963
	(0.159)	(0.297)	(0.787)	(1.419)
2	0.662***	1.144***	−0.473	0.408
	(0.231)	(0.341)	(1.047)	(1.710)
3	0.847***	1.089***	−0.268	6.271
	(0.268)	(0.400)	(1.250)	(5.045)
4	0.806***	1.121***	0.973	−0.222
	(0.283)	(0.382)	(1.263)	(2.198)
5	0.658**	1.024**	−0.497	−1.109
	(0.306)	(0.499)	(1.236)	(2.681)
5 years total	0.475***	0.876***	−0.139	0.658
	(0.174)	(0.291)	(0.768)	(1.205)
Observations	19,096	15,505	2,638	953
Negative feedback cases	2,838	2,086	343	409
No. of school districts	977	965	337	147

Note. Fixed effects of relative years, calendar years, performance rating, and a linear function of score differences are included but not shown. Because one school district may have experienced multiple performance ratings, the sum of the number of school districts in the latter three columns is not equal to that of the first column. Robust standard errors, clustered at the school district level, are in parentheses. ***$p < .01$, **$p < .05$, *$p < .1$.

lower ratings, particularly between A and B. The effect sizes are about two times that of the full sample, as shown in column (1). By contrast, as shown in columns (3) and (4), negative feedback shows no statistically significant effects on service outsourcing for the subsamples of lower ratings, specially from B to C and from C to F. This suggests that the negative feedback from a rating A to a rating B drives the effects in the full sample. Overall, these results support H1 that negative feedback increases public service outsourcing for school districts.

Moderating Effect of Organizational Capacity

Table 3 shows the effect of negative feedback on the organizational capacity of school districts estimated based on Equation (2). Column (1) shows the effects of negative feedback by fiscal capacity indicated by taxable property value per pupil; column (2) shows the effects of negative feedback by human resource capacity proxied by the number of staff per pupil. Column (1) shows that the interaction term between negative feedback and taxable property value per pupil has a negative sign and has a t value of about 1.4, which is close but not statistically significant at the $p < .1$ level. Column (2)

Table 3. The Negative Feedback Effects by Organizational Capacity.

Relative years after negative feedback	(1) Taxable property value	(2) Number of staff per pupil
Negative feedback	4.303 (2.683)	1.643** (0.644)
Taxable property value per pupil	−0.909* (0.549)	
Negative feedback × Taxable property value per pupil	−0.293 (0.209)	
Number of staff per pupil		1.946 (6.856)
Negative feedback × Number of staff per pupil		−7.253** (3.646)
Observations	19,034	19,092
No. of school districts	974	977

Note. Fixed effects of relative years, calendar years, performance rating, and a linear function of score differences are included but not shown. Robust standard errors, clustered at the school district level, are in parentheses. *** $p < .01$, ** $p < .05$, * $p < .1$.

shows that the interaction term between negative feedback and the number of staff per pupil has a negative and statistically significant effect at the $p < .05$ level.

To facilitate the interpretation of the interactive effects, we plot the marginal effects of negative feedback by the two moderators in Figures 2 and 3, along with the distribution of the moderators (Brambor, Clark, and Golder 2006). In Figure 2, we observe a statistically significant effect of negative feedback when the taxable property value per pupil is below 13.30, accounting for about 73 percent of the observations of our sample. When the taxable property value per pupil is relatively low (25th percentile, or 12.39), negative feedback shows a statistically significant effect of 0.78 on outsourcing. In contrast, when the taxable property value per pupil is relatively high (75th percentile, or 13.35), negative feedback shows a statistically insignificant effect of 0.48 on outsourcing. This is consistent with H2 in that negative feedback has a greater effect on service outsourcing when the fiscal capacity is relatively low.

Similarly, in Figure 3, we observe a statistically significant effect of negative feedback when the number of staff per pupil is below 0.18, or about 76 percent of the observations of our sample. Negative feedback shows a statistically significant effect of 0.89 on outsourcing when the number of staff per pupil is relatively low (25th percentile, or 0.14). It shows a statistically insignificant effect of 0.53 when the number of staff per pupil is relatively high (75th percentile, or 0.18). This finding supports H3 that negative feedback matters more when the school district's human resource capacity is limited.

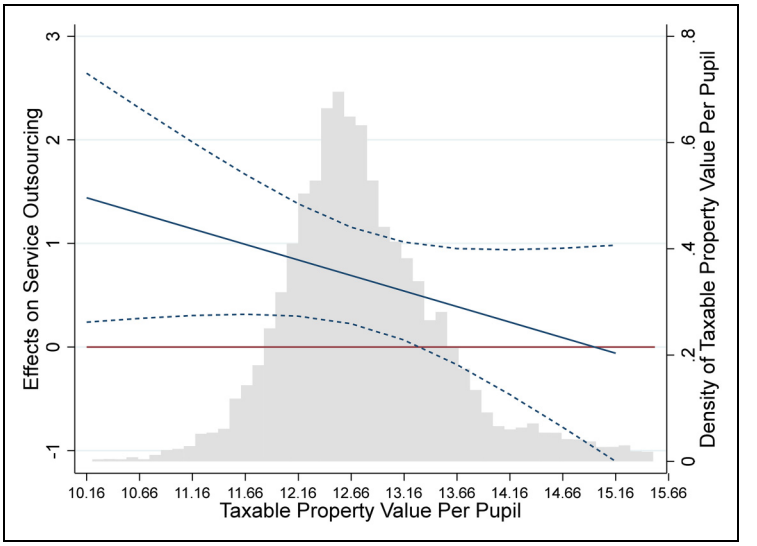

Figure 2. The marginal effects of negative feedback by taxable property value per pupil.

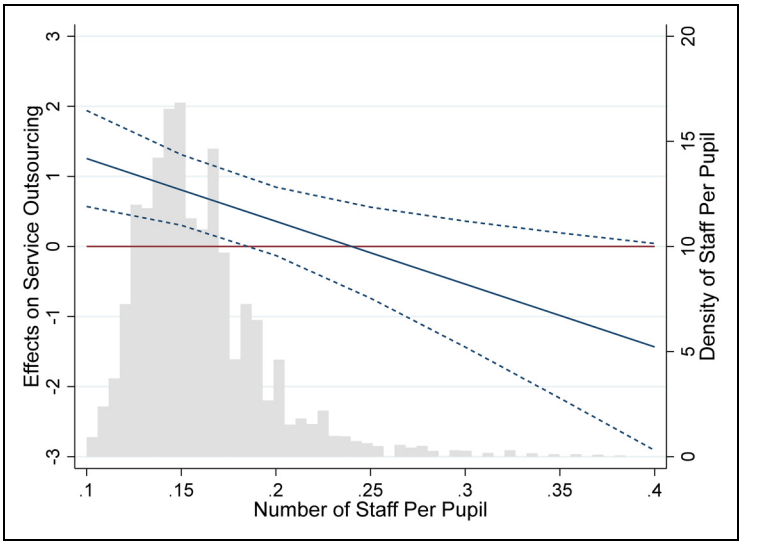

Figure 3. The marginal effects of negative feedback by the number of staff per pupil.

Robustness Checks

We conduct three robustness checks for the main results. First, as is conventional in the RD literature (Lee and Lemieux 2010), we use alternative bandwidths to rerun the main model based on Equation (1). We use a smaller bandwidth of 3 points of performance scores and a data-driven bandwidth of 10 points of performance scores, following Calonico, Cattaneo, and Farrell (2020). Table A3 in the Appendix shows that the main results remain robust. Second, we use alternative specifications of polynomial functions to re-estimate the main model. Table A4 in the Appendix shows that the main results are robust with controlling the running variable's quadratic and cubic functions.

Third, we explore the heterogeneity of the effects of the timing of school board elections. Local political institutions can be critical for the relationship between outsourcing and accountability. One such factor is when school board elections are held: on-cycle or off-cycle. Off-cycle elections have lower turnout but higher voter interest, voter knowledge, and interest group involvement, making incumbents more likely to lose than on-cycle elections (Anzia 2012; de Benedictis-Kessner 2018). As a result, off-cycle elections should be associated with higher accountability pressure for school board members, who are more likely to use outsourcing to reduce it than those under on-cycle elections. Therefore, we expect a larger impact of lower FRIST ratings on outsourcing for school districts with off-cycle elections than on-cycle elections. However, as reported in Table A5 in the Appendix, we find no statistically significant moderating effects of the timing of school board elections. This suggests the main finding remains invariant to the accountability concerns of school board members, as indicated by election timing.

Discussion and Conclusion

In this study, we have examined the impact of organizational performance feedback on public service outsourcing. We posit that negative performance feedback increases outsourcing through performance and accountability effects. Public organizations may treat outsourcing as a key means to remedy performance gaps because it may help cut costs and improve service efficiency. Additionally, public organizations may resort to outsourcing to circumvent the pressure of accountability or avoid blame. Further, public organizations with lower organizational capacity may have a stronger demand for outsourcing. In comparison, those with higher organizational capacity are more likely to address the performance gaps with in-house production of public services.

We use an RD design as a quasi-experimental method to estimate the effect of performance feedback on public service outsourcing, which allows the identification of causal effects near the rating cutoffs. The empirical tests with data from the Texas FIRST system generate evidence consistent with

these expectations. We find that negative performance feedback causes Texas school districts to increase the proportion of expenditures on public service outsourcing by about 6.8 percent to 8.7 percent, or about 2 million dollars, for an average school district in the subsequent five years.

Moreover, the impact of negative feedback on outsourcing is larger when the school districts have lower organizational capacity, as indicated by smaller tax bases and staff sizes. While we find statistically significant moderating effects of both fiscal and human resource capacity, future research can further compare their impact. One possible direction is to test if human resource capacity has a weaker moderating effect than fiscal capacity, to the degree that school districts may face more stringent constraints in personnel than fiscal decisions.

In addition, we find critical dynamic effects of performance feedback on public service outsourcing. We find that FIRST ratings have statistically significant effects on outsourcing except for the first year in the five-year window. This suggests that outsourcing takes time to materialize in response to negative performance feedback. In Texas, school board members serve on a two-year term with no term limit. In our sample period, the average tenure of school board members is 5.76 years, suggesting that an average board member serves more than two terms. For a school board member serving a single term, while the first year is too short to see the benefit of outsourcing in reducing accountability on the ballot, observing a negative effect in the second year is still assuring and promising. Moreover, the school board members serving longer than two years have an even more extended time window to benefit from lower accountability pressure should they choose to outsource.

From the political perspective, while outsourcing may facilitate blame-shifting, it may not help elected officials claim credit for outsourced services. For example, Dawkins (2021) finds that citizens are more likely to blame the government for outsourced service failures than praise it for service success. This suggests a limited potential for outsourcing as a tool for elected officials to claim credit. However, it remains unknown if the same asymmetry between blame and credit exists for in-house produced services and how it compares to outsourced services. If elected officials are subject to the same negativity bias (Weaver 1986), they may care more about blame avoidance than credit taking and thus still resort to outsourcing.

The results also show that the effect of negative feedback is driven by the performance ratings changing from A to B. There are several possible interpretations for this finding. First, the relatively small sample sizes for the subsamples of performance ratings from B to C and C to F may prevent the precise estimation of the effects. Second, suppose the statistical power is sufficient, and thus, the true effects for the rating changes from B to C and from C to F are zero. In that case, we observe larger impacts of negative feedback on outsourcing for high-performing organizations than for low-performing

organizations. One explanation is that negative feedback from an A rating may be more salient to public organizations, thus prompting outsourcing as a response, given that A ratings are common for many districts in the sample. Nonetheless, this finding contradicts the expectation that low-performing organizations are more likely to respond to negative performance signals (e.g., Hong 2019; Wei et al. 2023). Future research may test these competing explanations using a research design with higher statistical power or qualitative designs, such as case studies for the districts with C or F ratings, to uncover how lower ratings affect their outsourcing decisions. Researchers can then extend the theory to compare the effect of negative feedback on outsourcing in high-performing versus low-performing organizations.

Further, our data and research design allow us to compare lower and higher ratings within and between school districts. These cases include rating downgrades or upgrades by the same districts across years. While they constitute a relatively low proportion in our sample, future research may compare the outsourcing effects of rating downgrades and upgrades, exploring whether rating downgrades from A to B has similar effects as rating upgrades from B to A.

These findings have critical policy implications. We offer evidence that public organizations use outsourcing to address performance gaps, as negative performance feedback indicates. This suggests that performance feedback triggers outsourcing as a viable organizational strategy that may save costs and improve efficiency, which points to future performance improvements. Moreover, public organizations with lower fiscal and human capacities are more likely to use outsourcing as a response. Thus, outsourcing can be a particularly attractive option to public organizations with limited resources to meet performance gaps through in-house production.

To the extent that public organizations may use outsourcing to reduce the accountability pressure resulting from negative performance feedback, the increase in outsourcing may raise concerns about accountability (Dawkins 2021). Recent studies have shown that outsourcing may decrease service quality and lower accountability without sufficient monitoring and contract management (Brown and Potoski 2003b). Previous studies suggest that the monitoring cost could reach as high as 20 percent of the total outsourcing cost (Prager 1994; Marvel and Marvel 2007). As public organizations with lower capacity are prone to outsource services under negative performance feedback, they may need to invest more resources to monitor and hold external vendors accountable. Therefore, future research may examine the service performance contracted out due to negative performance feedback.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: Xin Chen acknowledges receiving funding from Guangdong Basic and Applied Basic Research Foundation under grant number 2025A1515010467.

ORCID iD

Xin Chen https://orcid.org/0000-0002-1499-2958

Notes

1. The FIRST system assigned the performance ratings based on the number of questions with positive answers from 2002 to 2006.
2. For details, see https://tea.texas.gov.
3. We exclude the performance rating data in 2015 because they do not provide information on each of the four performance ratings; they only report if a school district passed or failed the performance evaluation.
4. For details, please see https://comptroller.texas.gov/economy/fiscal-notes/2019/jan/funding.php.
5. The data on administrative staff includes the superintendent, assistant superintendents, and administrative support staff in the school district administration.
6. We later show that our results are robust with alternative bandwidths.
7. Equation (1) below is analogous to Equation (7) for the intent-to-treatment specification by Cellini, Ferreira, and Rothstein (2010, 228).
8. The results are robust with alternative specifications of the window of time, such as three or seven years. The results of this robustness check are available upon request.
9. We report the results of robustness checks with alternative forms of polynomial functions in the Appendix.
10. This assumption is consistent with the empirical test showing no statistically significant effects of negative feedback on outsourcing in years $t-1$, $t-2$, and $t-3$, as reported in the Appendix.
11. The RD literature is inconclusive on whether to allow the polynomial function to vary by the moderator variable. Our results remain robust with respect to either specification.

References

AbouAssi, Khaldoun, and Suyeon Jo. 2017. "Partnerships among Lebanese Nonprofit Organizations: Assessing the Impact on Human Resource Capacity." *American Review of Public Administration* 47 (6): 687–98.

Alonso, José M., and Rhys Andrews. 2020. "Political Ideology and Social Services Contracting: Evidence from a Regression Discontinuity Design." *Public Administration Review* 80 (5): 743–54.

Andersen, Simon Calmar, and Helena Skyt Nielsen. 2020. "Learning from Performance Information." *Journal of Public Administration Research and Theory* 30 (3): 415–31.

Andrews, Rhys, George Boyne, Larry O'Toole, Kenneth Meier, and Richard Walker. 2013. "Managing Migration? EU Enlargement, Local Government Capacity and Performance in England." *Public Administration* 91 (1): 174–94.

Anzia, Sarah F. 2012. "The Election Timing Effect: Evidence from a Policy Intervention in Texas." *Quarterly Journal of Political Science* 7 (3): 209–48.

Bach, Tobias, and Kai Wegrich. 2019. "The Politics of Blame Avoidance in complex Delegation Structures: The Public Transport Crisis in Berlin." *European Political Science Review* 11 (4): 415–31.

Baum, Joel AC, and Kristina B. Dahlin. 2007. "Aspiration Performance and Railroads' Patterns of Learning from Train Wrecks and Crashes." *Organization Science* 18 (3): 368–85.

Bel, Germà, and Xavier Fageda. 2017. "What Have We Learned from the Last Three Decades of Empirical Studies on Factors Driving Local Privatisation?" *Local Government Studies* 43 (4): 503–11.

Bel, Germà, Xavier Fageda, and Mildred E. Warner. 2010. "Is Private Production of Public Services Cheaper Than Public Production? A Meta-Regression Analysis of Solid Waste and Water Services." *Journal of Policy Analysis and Management* 29 (3): 553–77.

Boyne, George A. 1998. "The Determinants of Variations in Local Service Contracting: Garbage In, Garbage Out?" *Urban Affairs Review* 34 (1): 150–63.

Boyne, George A. 2003. "Sources of Public Service Improvement: A Critical Review and Research Agenda." *Journal of Public Administration Research and Theory* 13 (3): 367–94.

Brambor, Thomas, William Roberts Clark, and Matt Golder. 2006. "Understanding Interaction Models: Improving Empirical Analyses." *Political Analysis* 14 (1): 63–82.

Brown, Trevor L., and Matthew Potoski. 2003a. "Transaction Costs and Institutional Explanations for Government Service Production Decisions." *Journal of Public Administration Research and Theory* 13 (4): 441–68.

Brown, Trevor L., and Matthew Potoski. 2003b. "Managing Contract Performance: A Transaction Costs Approach." *Journal of Policy Analysis and Management* 22 (2): 275–97.

Busuioc, E. M., and Martin Lodge. 2016. "The Reputational Basis of Public Accountability." *Governance* 29 (2): 247–63.

Busuioc, Madalina, and Martin Lodge. 2017. "Reputation and Accountability Relationships: Managing Accountability Expectations Through Reputation." *Public Administration Review* 77 (1): 91–100.

Calonico, Sebastian, Matias D. Cattaneo, and Max H. Farrell. 2020. "Optimal Bandwidth Choice for Robust Bias-corrected Inference in Regression Discontinuity Designs." *The Econometrics Journal* 23 (2): 192–210.

Cattaneo, Matias D., Michael Jansson, and Xinwei Ma. 2020. "Simple Local Polynomial Density Estimators." *Journal of the American Statistical Association* 115 (531): 1449–55.

Cattaneo, Matias D., Rocío Titiunik, Gonzalo Vazquez-Bare, and Luke Keele. 2016. "Interpreting Regression Discontinuity Designs with Multiple Cutoffs." *Journal of Politics* 78 (4): 1229–48.

Cellini, Stephanie Riegg, Fernando Ferreira, and Jesse Rothstein. 2010. "The Value of School Facility Investments: Evidence from a Dynamic Regression Discontinuity Design." *Quarterly Journal of Economics* 125 (1): 215–61.

Coats, Jennifer C. 2002. "Applications of Principal-Agent Models to Government Contracting and Accountability Decision Making." *International Journal of Public Administration* 25 (4): 441–61.

Craig, Steven G., Scott A. Imberman, and Adam Perdue. 2013. "Does It Pay to Get an A? School Resource Allocations in Response to Accountability Ratings." *Journal of Urban Economics* 73 (1): 30–42.

Cyert, Richard M., and James G. March. 1963. "A Behavioral Theory of the Firm." *Englewood Cliffs, NJ* 2 (4): 169–87.

Davidson, Frank D., Michael Schwanenberger, and Richard Wiggall. 2019. "What Matters Most in Superintendent Evaluation." *Education Leadership Review* 20 (1): 217–33.

Dawkins, Ryan. 2021. "Private Contracting and Citizen Attitudes Toward Local Government." *Urban Affairs Review* 57 (5): 1286–311.

de Benedictis-Kessner, Justin. 2018. "Off-Cycle and Out of Office: Election Timing and the Incumbency Advantage." *The Journal of Politics* 80 (1): 119–32.

Drolc, Cody A., and Lael R. Keiser. 2021. "The Importance of Oversight and Agency Capacity in Enhancing Performance in Public Service Delivery." *Journal of Public Administration Research and Theory* 31 (4): 773–89.

Eisinger, Peter. 2002. "Organizational Capacity and Organizational Effectiveness among Street-Level Food Assistance Programs." *Nonprofit and Voluntary Sector Quarterly* 31 (1): 115–30.

Elkomy, Shimaa, Graham Cookson, and Simon Jones. 2019. "Cheap and Dirty: The Effect of Contracting Out Cleaning on Efficiency and Effectiveness." *Public Administration Review* 79 (2): 193–202.

Fernandez, Sergio, Jay Eungha Ryu, and Jeffrey L. Brudney. 2008. "Exploring Variations in Contracting for Services among American Local Governments: Do Politics Still Matter?" *American Review of Public Administration* 38 (4): 439–62.

Fernandez, Sergio, Craig R. Smith, and Jeffrey B. Wenger. 2007. "Employment, Privatization, and Managerial Choice: Does Contracting out Reduce Public Sector Employment?" *Journal of Policy Analysis and Management* 26 (1): 57–77.

Flink, Carla M. 2019. "Predicting Budgetary Change: The Effect of Performance Gaps." *Journal of Public Administration Research and Theory* 29 (2): 227–37.

Gargan, John J. 1981. "Consideration of Local Government Capacity." *Public Administration Review* 41 (6): 649–58.

Geys, Benny, and Rune J. Sørensen. 2016. "Revenue Scarcity and Government Outsourcing: Empirical Evidence from Norwegian Local Governments." *Public Administration* 94 (3): 769–88.

Girth, Amanda M., Amir Hefetz, Jocelyn M. Johnston, and Mildred E. Warner. 2012. "Outsourcing Public Service Delivery: Management Responses in Noncompetitive Markets." *Public Administration Review* 72 (6): 887–900.

Grissom, Jason A., and Stephanie Andersen. 2012. "Why Superintendents Turn Over." *American Educational Research Journal* 49 (6): 1146–80.

Grissom, Jason A., and Hajime Mitani. 2016. "Salary, Performance, and Superintendent Turnover." *Educational Administration Quarterly* 52 (3): 351–91.

Guo, Ming, and Sam Willner. 2017. "Swedish Politicians' Preferences Regarding the Privatization of Elderly Care." *Local Government Studies* 43 (1): 1–21.

Hall, Jeremy L. 2008. "Assessing Local Capacity for Federal Grant-Getting." *American Review of Public Administration* 38 (4): 463–79.

Hannaway, Jane, and Laura Hamilton. 2008. *Performance-Based Accountability Policies: Implications for School and Classroom Practices*. Washington: Urban Institute and RAND Corporation.

Hansen, Gustav Egede, Germá Bel, and Ole Helby Petersen. 2022. "The Unequal Distribution of Consequences of Contracting Out: Female, Low-Skilled, and Young Workers Pay the Highest Price." *Journal of Public Administration Research and Theory* 33 (3): 434–52.

Hefetz, Amir, and Mildred E. Warner. 2012. "Contracting or Public Delivery? The Importance of Service, Market, and Management Characteristics." *Journal of Public Administration Research and Theory* 22 (2): 289–317.

Heinrich, Carolyn J., Laurence E. Lynn Jr, and H. Brinton Milward. 2010. "A State of Agents? Sharpening the Debate and Evidence Over the Extent and Impact of the Transformation of Governance." *Journal of Public Administration Research and Theory* 20 (suppl_1): i3–i19.

Holm, Jakob Majlund. 2018. "Successful Problem Solvers? Managerial Performance Information use to Improve low Organizational Performance." *Journal of Public Administration Research and Theory* 28 (3): 303–20.

Hong, Sounman. 2019. "A Behavioral Model of Public Organizations: Bounded Rationality, Performance Feedback, and Negativity Bias." *Journal of Public Administration Research and Theory* 29 (1): 1–17.

Hong, Sounman. 2020. "Performance Management Meets red Tape: Bounded Rationality, Negativity Bias, and Resource Dependence." *Public Administration Review* 80 (6): 932–45.

Hood, Christopher. 2011. *The Blame Game: Spin, Bureaucracy, and Self-Preservation in Government by Christopher Hood*. Princeton, NJ: Princeton University Press.

Hung, Wan-Ju, and Gregory A. Porumbescu. 2024. "Who Does the Public Blame for Inequities in Public Service Provision? Examining the Effect of Contracting and Cost Information in a Survey Experiment." *Public Performance & Management Review* 1–25.

Igalla, Malika, Jurian Edelenbos, and Ingmar van Meerkerk. 2020. "What Explains the Performance of Community-Based Initiatives? Testing the Impact of Leadership, Social Capital, Organizational Capacity, and Government Support." *Public Management Review* 22 (4): 602–32.

James, Oliver. 2011. "Performance Measures and Democracy: Information Effects on Citizens in Field and Laboratory Experiments." *Journal of Public Administration Research and Theory* 21 (3): 399–418.

James, Oliver, Sebastian Jilke, Carolyn Petersen, and Steven Van de Walle. 2016. "Citizens' Blame of Politicians for Public Service Failure: Experimental Evidence about Blame Reduction Through Delegation and Contracting." *Public Administration Review* 76 (1): 83–93.

James, Oliver, and Alice Moseley. 2014. "Does Performance Information About Public Services Affect Citizens' Perceptions, Satisfaction, and Voice Behavior?

Field Experiments with Absolute and Relative Performance Information." *Public Administration* 92 (2): 493–511.

Joaquin, M. E., and Thomas J. Greitens. 2011. "The Accountability-Performance Link: An Attempt at Distilling Some Mechanisms in a Management Reform Initiative." *Public Performance & Management Review* 34 (3): 323–49.

Kim, Mirae, and Shuyang Peng. 2018. "The Dilemma for Small Human Service Nonprofits: Engaging in Collaborations with Limited Human Resource Capacity." *Nonprofit Management and Leadership* 29 (1): 83–103.

Kim, Yunji, and Mildred E. Warner. 2016. "Pragmatic Municipalism: Local Government Service Delivery after the Great Recession." *Public Administration* 94 (3): 789–805.

Kogan, Vladimir, Stéphane Lavertu, and Zachary Peskowitz. 2017. "Direct Democracy and Administrative Disruption." *Journal of Public Administration Research and Theory* 27 (3): 381–99.

Lee, Gyeo Reh, Shinwoo Lee, Deanna Malatesta, and Sergio Fernandez. 2019. "Outsourcing and Organizational Performance: The Employee Perspective." *American Review of Public Administration* 49 (8): 973–86.

Lee, David S., and Thomas Lemieux. 2010. "Regression Discontinuity Designs in Economics." *Journal of Economic Literature* 48 (2): 281–355.

Leland, Suzanne, Zachary Mohr, and Jaclyn Piatak. 2021. "Accountability in Government Contracting Arrangements: Experimental Analysis of Blame Attribution Across Levels of Government." *American Review of Public Administration* 51 (4): 251–62.

Martinez-Vazquez, Jorge, and L. F. Jameson Boex. 1997. "Fiscal Capacity: An Overview of Concepts and Measurement Issues and their Applicability in the Russian Federation." GSU Andrew Young School of Policy Studies Working Paper 97-3.

Marvel, John D., and Amanda M. Girth. 2016. "Citizen Attributions of Blame in Third-Party Governance." *Public Administration Review* 76 (1): 96–108.

Marvel, Mary K., and Howard P. Marvel. 2007. "Outsourcing Oversight: A Comparison of Monitoring for In-House and Contracted Services." *Public Administration Review* 67 (3): 521–30.

Meier, Kenneth J., Nathan Favero, and Ling Zhu. 2015. "Performance Gaps and Managerial Decisions: A Bayesian Decision Theory of Managerial Action." *Journal of Public Administration Research and Theory* 25 (4): 1221–46.

Milward, H. B., and Keith G. Provan. 2000. "Governing the Hollow State." *Journal of Public Administration Research and Theory* 10 (2): 359–80.

Min, Byung Hee, Youngmin Oh, and Ralph S. Brower. 2021. "The Effects of Diverse Feedback Dynamics on Performance Improvement: A Typology of Performance Feedback Signals." *Administration & Society* 53 (1): 123–57.

Mohr, Robert, Steven C. Deller, and John M. Halstead. 2010. "Alternative Methods of Service Delivery in Small and Rural Municipalities." *Public Administration Review* 70 (6): 894–905.

Mohr, Zachary, Jared McDonald, Jaclyn Piatak, and Suzanne Leland. 2023. "Is Government Escaping Blame? The Effect of Contracting and Victim Attributes in a Prisoner Transportation Experiment." *Public Management Review* 1–22.

Moynihan, Donald P. 2008. *The Dynamics of Performance Management: Constructing Information and Reform*. Washington, DC: Georgetown University Press.

Moynihan, Donald P., and Patricia W. Ingraham. 2003. "Look for the Silver Lining: When Performance-Based Accountability Systems Work." *Journal of Public Administration Research and Theory* 13 (4): 469–90.

Nelson, Ashlyn Aiko, and Rekha Balu. 2014. "Local Government Responses to Fiscal Stress: Evidence from the Public Education Sector." *Public Administration Review* 74 (5): 601–14.

Nicholson-Crotty, Sean, Jill Nicholson-Crotty, and Sergio Fernandez. 2017. "Performance and Management in the Public Sector: Testing a Model of Relative Risk Aversion." *Public Administration Review* 77 (4): 603–14.

Nielsen, Poul A. 2014. "Performance Management, Managerial Authority, and Public Service Performance." *Journal of Public Administration Research and Theory* 24 (2): 431–58.

O'Toole, L. J., and Kenneth J. Meier Jr. 2004. "Parkinson's law and the new Public Management? Contracting Determinants and Service-Quality Consequences in Public Education." *Public Administration Review* 64 (3): 342–52.

O'Toole, L. J., and Kenneth J. Meier Jr. 2010. "In Defense of Bureaucracy: Public Managerial Capacity, Slack and the Dampening of Environmental Shocks." *Public Management Review* 12 (3): 341–61.

Park, Ji Hyung, and Shihyun Noh. 2021. "Does Contracting out Matter to Financial Condition? The Case of New York Local Governments." *Public Performance & Management Review* 44 (6): 1191–214.

Petersen, Ole Helby, Ulf Hjelmar, and Karsten Vrangbæk. 2018. "Is Contracting out of Public Services Still the Great panacea? A Systematic Review of Studies on Economic and Quality Effects from 2000 to 2014." *Social Policy & Administration* 52 (1): 130–57.

Piatak, Jaclyn, Zachary Mohr, and Suzanne Leland. 2017. "Bureaucratic Accountability in Third-party Governance: Experimental Evidence of Blame Attribution During Times of Budgetary Crisis." *Public Administration* 95 (4): 976–89.

Prager, Jonas. 1994. "Contracting Out Government Services: Lessons from the Private Sector." *Public Administration Review* 54: 176–84.

Rho, Eunju. 2013. "Contracting Revisited: Determinants and Consequences of Contracting Out for Public Education Services." *Public Administration Review* 73 (2): 327–37.

Rich, Jessica AJ. 2022. "Outsourcing Bureaucracy to Evade Accountability: How Public Servants Build Shadow State Capacity." *American Political Science Review* 1–16.

Salge, Torsten O. 2011. "A Behavioral Model of Innovative Search: Evidence from Public Hospital Services." *Journal of Public Administration Research and Theory* 21 (1): 181–210.

Simon, Herbert A. 2013. *Administrative Behavior*. New York: Simon and Schuster.

Traynor, Mark, Can Chen, and Miranda Kitterlin. 2019. "Contracting School Meals in a Rural Florida District." *International Hospitality Review* 33 (2): 73–90.

Van der Voet, Joris. 2023. "Search in Response to Negative Performance Feedback: Problem-Definition and Solution-Generation." *Journal of Public Administration Research and Theory* 33 (1): 43–62.

Van der Voet, Joris, and Elvira Lems. 2022. "Decision-Makers' Generation of Policy Solutions Amidst Negative Performance: Invention or Rigidity?" *Public Administration Review* 82 (5): 931–45.

Walker, Richard M., Oliver James, M. Jae Moon, and Wen Wen. 2024. "A Cross-Cultural Replication of 'Citizens' Blame of Politicians for Public Service Failure: Experimental Evidence About Blame Reduction Through Delegation and Contracting." *Public Administration Review*.

Warner, Mildred E., and Austin M. Aldag. 2021. "Re-municipalization in the US: A Pragmatic Response to Contracting." *Journal of Economic Policy Reform* 24 (3): 319–32.

Warner, Mildred E., Austin M. Aldag, and Yunji Kim. 2021. "Privatization and Intermunicipal Cooperation in US Local Government Services: Balancing Fiscal Stress, Need and Political Interests." *Public Management Review* 23 (9): 1359–76.

Warner, Mildred E., and Amir Hefetz. 2012. "Insourcing and Outsourcing: The Dynamics of Privatization among US Municipalities 2002–2007." *Journal of the American Planning Association* 78 (3): 313–27.

Weaver, R. K. 1986. "The Politics of Blame Avoidance." *Journal of Public Policy* 6 (4): 371–98.

Wei, Wenchi, Chengwei Wang, Wenkang Zhai, and Wenzhao Li. 2023. ""Honor List" and "Shame Roll": Quasi-Experimental Evidence of the Effect of Performance Feedback under Political Control." *Journal of Public Administration Research and Theory* 33 (1): 122–38.

Wei, Wenchi, Zhiwei Zhang, and Kaifeng Yang. 2022. "Local Institutions and Public Service Outsourcing: Managerial Professionalism, Mayoral Strength, and Electoral Rules." *Public Administration Review* 82 (6): 1095–123.

Zhang, Zhiwei, Bryan Gibson, and Josephine Gatti Schafer. 2018. "Reexamining the Relationship Between Fiscal Stress and Outsourcing." *Public Performance & Management Review* 41 (1): 22–46.

Author Biographies

Jinhai Yu is assistant professor in Public Administration at University of Connecticut, USA. His research interests include public organizational performance, budgetary politics and policy, and fiscal transparency and accountability. His research has been published in Journal of Public Administration Research and Theory, Public Administration Review, and Public Management Review, among others.

Xin Chen is assistant professor in Public Administration at Sun Yat-sen University, China His research interests include public budgeting and finance, local governance, and organizational performance. His research has been published in Journal of Public Administration Research and Theory, Public Management Review, and American Review of Public Administration, among others.

Appendix

Heterogeneity of Effects by Election Timing

We collected data on the election timing of Texas school districts by following the methodology of Anzia (2012). We first used Anzia's (2012) data, which covers election timing up to 2009, as the baseline. We then manually updated the election timing for each district using data from Ballotpedia and individual district websites. For a few smaller districts where election information was unavailable, we relied on Anzia's data, assuming that their election timing had remained unchanged from 2009 to 2020. This assumption is reasonable, as most districts did not change their election timing during this period. We construct the variable of election timing as a dummy, with 1 indicating on-cycle elections and 0 indicating off-cycle elections.

Table A1. Testing the Pre-Rating Trends.

Dependent variables	(1)	(2)
Percentage outsourced		
$t-1$	0.309 (0.241)	−0.067 (0.155)
$t-2$	0.499** (0.234)	0.264 (0.176)
$t-3$	0.357 (0.234)	−0.264 (0.176)
Fixed effects	Y	Y
Linear function of running variable	Y	Y
Pooled sample over years	N	Y

Note. Each entry shows the negative feedback effects on service outsourcing from a separate specification. The three rows show the negative feedback effects in years $t-1$, $t-2$, and $t-3$, respectively. The specifications in column (1) include fixed effects and a linear function of the running variable. Column (2) further pools the stacked samples. Robust standard errors, clustered at the school district level, are in parentheses. *** $p < .01$, ** $p < .05$, * $p < .1$.

Table A2. Balance Tests.

Covariates	(1) Average effects 5 years total	(2) Dynamic effects 1	(3) 2	(4) 3	(5) 4	(6) 5
Number of students	18.940 (34.508)	9.466 (21.600)	16.920 (33.072)	28.385 (49.571)	24.771 (55.032)	45.195 (53.549)
Number of schools	0.028 (0.050)	0.045 (0.041)	−0.021 (0.059)	0.029 (0.064)	0.048 (0.071)	0.060 (0.079)
Total revenues (log)	0.005 (0.011)	−0.004 (0.009)	0.006 (0.012)	0.004 (0.014)	0.020 (0.017)	0.021 (0.018)
Total expenditures (log)	−0.014 (0.019)	−0.011 (0.017)	−0.005 (0.020)	−0.037 (0.024)	−0.006 (0.027)	−0.004 (0.030)
Dropout rate	−0.016 (0.075)	0.100 (0.081)	−0.037 (0.097)	−0.097 (0.095)	−0.181* (0.105)	−0.011 (0.108)
SAT scores	−32.209 (28.426)	−13.595 (32.022)	−67.053* (34.201)	−53.753 (36.749)	−8.658 (35.981)	−3.749 (40.943)
ACT scores	0.097 (0.320)	0.442 (0.360)	0.047 (0.371)	−0.025 (0.448)	−0.710 (0.461)	0.086 (0.519)
Teacher salaries (log)	−0.001 (0.004)	−0.002 (0.003)	−0.001 (0.004)	−0.002 (0.005)	0.003 (0.005)	0.002 (0.006)

Note. Each row reports the regression discontinuity estimates of the effects of negative feedback on an outcome based on Equation (1) in the main text. Column (1) reports the average effects for the five years, while other columns show the dynamic effects by year. Fixed effects of relative years, calendar years, performance rating, and a linear function of the running variable are included but not shown. Robust standard errors, clustered at the school district level, are in parentheses. SAT = Scholastic Aptitude Test; ACT = American College Testing.***$p < .01$, **$p < .05$, *$p < .1$.

Table A3. Negative Feedback Effects on Outsourcing With Alternative Bandwidths.

Relative years after negative feedback	(1) Bandwidth of 3 points	(2) Data-driven bandwidth of 10 points
1	0.424 (0.303)	0.233 (0.236)
2	0.995** (0.412)	0.810*** (0.294)
3	0.851 (0.527)	0.825** (0.371)
4	1.090** (0.491)	0.987** (0.405)
5	0.953* (0.544)	0.703* (0.417)
5 years total	0.784** (0.364)	0.605** (0.251)
Observations	15,461	18,620
No. of school districts	946	974

Note. The bandwidth in column (1) is the middle point of the smallest range between any two ratings in the sample. The bandwidth in column (2) is a data-driven bandwidth calculated following Calonico, Cattaneo, and Farrell (2020). Fixed effects of relative years, calendar years, performance rating, and a linear function of the running variable are included but not shown. Robust standard errors, clustered at the school district level, are in parentheses. ***$p < .01$, **$p < .05$, *$p < .1$.

Table A4. Negative Feedback Effects on Outsourcing With Alternative Polynomials.

Relative years after negative feedback	(1) Quadratic	(2) Cubic
1	0.221 (0.219)	0.231 (0.216)
2	0.746*** (0.272)	0.755*** (0.271)
3	0.834** (0.344)	0.850** (0.343)
4	0.985** (0.389)	0.994** (0.391)
5	0.729* (0.395)	0.769* (0.397)
5 years total	0.575** (0.231)	0.589** (0.229)
Observations	19,096	19,096
No. of school districts	997	977

Note. Fixed effects of relative years, calendar years, performance rating, and a polynomial function of the running variable are included but not shown. Robust standard errors, clustered at the school district level, are in parentheses. ***$p < .01$, **$p < .05$, *$p < .1$.

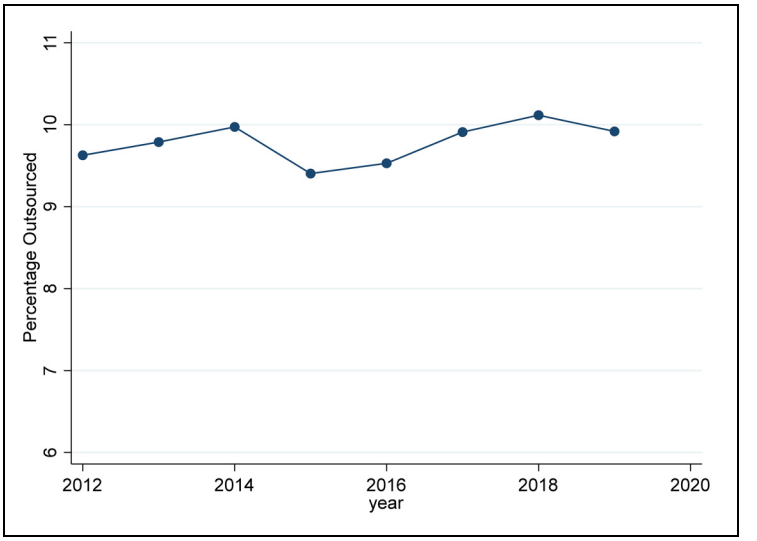

Figure A1. The average percentage of outsourcing across years.

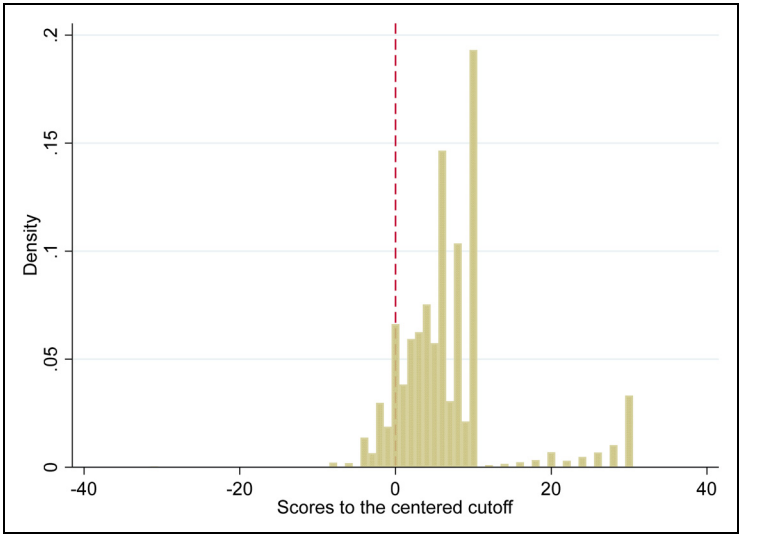

Figure A2. Distribution of performance scores relative to the centered cutoff.

Table A5. The Negative Feedback Effects by Election Timing.

Relative years after negative feedback	(1) Election timing
Negative feedback	0.463** (0.234)
On cycle election	−0.250 (0.159)
Negative feedback × On cycle election	0.026 (0.283)
Observations	18,853
No. of school districts	968

Note. Fixed effects of relative years, calendar years, performance rating, and a linear function of score differences are included but not shown. Robust standard errors, clustered at the school district level, are in parentheses. ***$p < .01$, **$p < .05$, *$p < .1$.