Fiscal Monitoring, School District Outcomes, and Residential Choice

Travis St.Clair*

July 21, 2022

Abstract

Under the canonical model of local political economy, taxpayers vote with their feet to find the community that provides the optimal bundle of taxes and public goods (Tiebout, 1956). One common objection to this model is that taxpayers lack perfect information about their choice sets, and consequently rely on heuristics or third-party rating agencies. This paper exploits New York state's fiscal monitoring program to examine the effect that quality labels have on school district enrollment and outcomes. Taxpayers may be especially sensitive to labels providing information about the management of school districts, as studies have shown that parents of school-age children pay close attention to information on school quality. Using a regression discontinuity design, I examine the effect of fiscal stress and "environmental" stress labels on property values, school enrollment, demographics, and test scores. While fiscal stress labels have no effect, environmental stress labels – indicating social, economic, and demographic pressures, such as a high percentage of disadvantaged students – cause enrollment and property values to decline, especially in wealthier districts. Taxpayers are more sensitive to demographic and economic indicators of school quality than to fiscal indicators.

JEL Codes: I21, D83

Keywords: fiscal monitoring, school finance, school quality, residential sorting

*Robert F. Wagner School of Public Service, New York University, 295 Lafayette Street, 2nd Floor, New York, NY 10012 (email: travis.stclair@nyu.edu).

1 Introduction

Under the classic Tiebout model of local political economy, taxpayers vote with their feet to find the community that provides the optimal bundle of taxes and public goods (Tiebout, 1956). One common objection to this model is that taxpayers lack perfect information about their choice sets. However, as the internet has given rise to an extraordinary democratization of information, the more relevant constraint in today's modern world may not be a lack of information but rather limits in taxpayers' ability to process all of the information available to them (Simon, 1990). Research in psychology and behavioral economics suggest that, in the face of such complex information, individuals frequently rely on heuristics or third-party rating agencies (Gigerenzer, 2004).

Ratings may be particularly salient for taxpayers with school age children. Local governments are the primary providers of primary and secondary education, and education is the largest spending item in local government budgets. Thus, of all the metrics that parents use to optimize their location choices, they will be especially sensitive to information about school quality. Indeed, studies have consistently shown that local school quality is an important determinant of house prices (Black, 1999; Kane, Riegg and Staiger, 2006; Fack and Grenet, 2010; Gibbons, Machin and Silva, 2013). In response to the high demand for information on school performance, and in particular for heuristics that can distill complex information, numerous public and private entities now provide ratings that better enable parents to rank and compare school districts (Hasan and Kumar, 2019).

This paper examines the sensitivity of taxpayers to heuristics about school quality by exploiting New York state's Fiscal Stress Monitoring System (FSMS). Created in 2013 by the Office of the State Comptroller (OSC), the monitoring system assesses fiscal stress in local governments and school districts and is part of a larger wave of increased oversight of local finances on the part of state governments (Nakhmurina, 2020). The FSMS calculates and publishes fiscal stress scores for all general purpose governments and school districts in the state, with the aim of measuring each government's ability to "maintain budget solvency" (Office of the New York State Comptroller, 2017*a*). In addition to financial scores, the system also publishes "environmental" scores, which measure other factors that may pose challenges to the fiscal health of a municipality but which, unlike the financial scores, are largely outside of a district's control. Among the environmental indicators that the state tracks are demographic measures of poverty and changes in the property tax base.

Because the monitoring system assigns labels of "fiscal stress" and "environmental stress" according to strict cut-offs in their scoring system, I examine the effect of these labels using a regression discontinuity (RD) design. The RD design compares school districts that received a score placing them just above the cut-off for receiving a fiscal stress (or environmental stress) label with districts that just narrowly avoided the label. Because the metrics used by the state to assign fiscal and environmental scores are based on a number of complex factors, such as property values and demographics, that are impossible for local district to precisely control, there is no way for districts to manipulate the score and influence the assignment of labels, a prerequisite for the RD design.

Although FSMS assesses stress in both general purpose governments and school districts, I focus on schools for two reasons. First, as noted above, parents of school-age children are especially sensitive to information about school quality and consequently may be more responsive to school district stress labels than municipal stress labels in their location choices. Second, school districts report on a variety of performance measures such as test scores that general purpose governments do not, allowing for a wide set of potential outcome variables.

One particularly attractive feature of studying New York's monitoring system is that, because the system assigns two different kinds of stress labels, it is possible to compare the effects of heuristics involving two different types of information about school quality: fiscal information and social/economic information. Much of the literature on school quality has focused on the capitalization of school quality into house values. But this literature fails to identify how parents identify and measure school quality, not to mention which dimensions they consider to be of greatest importance. Indeed, a summary of the school capitalization literature concludes by acknowledging that the existing literature has little to say about what features "make a 'good' school as perceived by parents" (Black and Machin, 2011). This paper fills in some of the gaps in this literature by identifying the sorts of labels that most impact decisions about school choice, and in doing so, sheds light on whether parents are more concerned with fiscal management or economic and social trends.

I find that fiscal stress labels have no effect on a variety of demographic and economic outcomes. On the other hand, environmental stress labels cause enrollment and property values to decline. Specifically, school districts labeled as environmentally stressed experience a decline in property values of between 1.5-3 percentage points in the three years following the assignment of the label as well as a 0.9 percent drop of enrollment in the third year (t+3). These effects are concentrated in wealthier districts, as measured by property values, and to a more limited extent, in districts with higher graduation rates. Property values in districts of below average wealth experience no effects from an environmental stress label in the first two years, while districts with above average wealth experience larger effects immediately. Similarly, districts of above average wealth that are labeled as environmentally stressed experience enrollment declines in the three years following a stress label, while districts of below average wealth experience no such effects. Districts with higher graduation rates that receive an environmental stress label also experience greater declines in their enrollment relative to worse performing districts, though not in property values. These findings are robust to a variety of empirical specifications. Taxpayers are more sensitive to demographic and economic indicators of school quality than they are to fiscal indicators.

This paper contributes to a broad literature on the use of heuristics in decision-making, and also, more specifically, to the literature on school ratings. Whereas parents have historically relied on informal channels or social networks to obtain information about school quality, there is now an enormous wealth of information about school performance that is widely available online (Mikulecky and Christie, 2014). In theory, greater information should increase efficiency by improving the match between parents and school districts. However, previous research has found that school rankings and report cards tend to harm poorly rated schools (Nunes, Reis and Seabra, 2015) and increase segregation, with more advantaged populations responding to rating information to a greater extent than disadvantaged populations (Hasan and Kumar, 2019; Figlio and Lucas, 2004).[1]

This paper also contributes to the literature on fiscal federalism, and specifically to the literature on fiscal monitoring and intervention. 23 states currently have policies designed to monitor the fiscal conditions of local governments in their state (Pew Charitable Trusts, 2016; Nakhmurina, 2020). These policies exist for various reasons, including the need to forestall bankruptcy filings and other fiscal emergencies that may require intervention on the part of the state. But many municipalities that may never declare bankruptcy nevertheless suffer from chronic fiscal stress that may impair their ability to provide services to their residents. Fiscal monitoring enables state governments to identify which districts are most vulnerable and work proactively with those districts in order to improve management practices and fiscal discipline. Previous research shows that fiscal monitoring reduces public corruption (Nakhmurina, 2020) but has little effect on fiscal health (Spreen and Cheek, 2016). In Ohio, fiscal stress labels have led to recovery plans that are beneficial in the long-run but have had negative short-run impacts on home sale prices, enrollment, and math proficiency rates (Thompson, 2016). A related literature on government audits in developing countries finds that audits and enhanced monitoring reduce public corruption (Avis, Ferraz and Finan, 2018; Olken, 2007).

This paper proceeds as follows. The following section provides an overview of New York state's fiscal monitoring system. Section III describes the data sources and provides descriptive statistics. Section IV outlines the methods. Section V presents the results, including heterogeneity analyses and robustness checks. Section VI concludes.

[1]Pope (2009) finds similar results for hospital rankings.

2 Institutional Background

States historically have done little to monitor the fiscal health of their local governments (Pew Charitable Trusts, 2016). However, in the aftermath of several high profile municipal bankruptcies in the 2010s, including Detroit; Jefferson County, Alabama; Stockton, California; and Central Falls, Rhode Island, many states instituted programs aimed at monitoring the fiscal health of local governments so as to pick up on signs of distress and avoid more painful fiscal emergencies. These programs appear to have been motivated in part by a fear of fiscal contagion and a potential downgrade of the state's credit rating (Pew Charitable Trusts, 2016). As an example, following Central Falls' bankruptcy filing in 2010, the state of Rhode Island passed a law giving bondholders priority over pensioners in the event of a default, which many interpreted as an effort to preserve the confidence of the bond market. The programs may also have been motivated by a desire to preserve the stability of service provision. Even if local governments are not at risk of default in the near-term, the long-term health of their economies is dependent on the ability to maintain essential services, and the state has an interest in preserving the economic vitality of its constituent governments. In response to these incentives, as of 2020, 23 states had fiscal monitoring systems in place (Nakhmurina, 2020).

The Office of the State Comptroller (OSC) in New York established the state's Fiscal Stress Monitoring System (FSMS) in 2013. The aim of the program is to assess the finances of each of the state's counties, towns, villages, and school districts on an annual basis in order to identify entities that are experiencing fiscal stress (Office of the New York State Comptroller, 2017*b*). The program provides information not only to the public, but also to local officials, who in principle can use the information to avoid disruption to vital services. The system does not include New York City in its analyses due to the city's unique fiscal structure.

Unlike programs in some other states, New York's system does not spell out a role for state intervention. California, Michigan, New Jersey, Nevada, Pennsylvania, and Rhode Island are among the states with more active assistance programs (Bowman and Zuschlag, 2022). For example, in 2013 emergency managers were running six different cities in Michigan. New York, in contrast, along with Connecticut and Massachusetts, decides on its level of involvement on a case-by-case basis depending on the severity of a local's government's distress (Pew Charitable Trusts, 2013). In each case, the state passes new legislation that establishes an emergency financial control board and outlines its powers vis-a-vis the distressed entity.

Under the FSMS, the Office of the State Comptroller calculates, separately, a fiscal stress score and an environmental stress score for each locality. The fiscal stress score measures a locality's ability to maintain budget solvency and is based on annual financial data alone. The environmental stress score measures other factors that may pose a threat to a district's fiscal health but which are not directly under the district's control, such as demographic changes to its electorate. The scoring system for general purpose governments differs slightly from the system for school districts. Because this paper focuses on school districts, I confine my description below to the system for school districts; Office of the New York State Comptroller (2017*b*) provides additional information on the system for governments.

Table 1 outlines the system for both fiscal and environmental scoring. As of 2021, the financial scoring system is based on six financial indicators in four categories. The four categories are fund balance, operating deficit, cash position, and short-term debt. The six indicators are 1) the general fund's unassigned fund balance as a percentage of gross expenditures, 2) the general fund's total fund balance as a percentage of gross expenditures, 3) the general fund's operating surplus (or deficit) as a percent of gross expenditures, 4) the general fund's total cash and short term investments as a percent of current liabilities, 5) the general fund's cash as a percent of monthly expenditures, and 6) the trend in short-term debt issuance. In 2017, the comptroller's office revised the metrics slightly, adjusting how they account for transfers in gross expenditures, and collapsing two short-term debt measures into one.

The environmental scoring system, which also underwent an overhaul in 2017, is based on six measures in six categories. The six categories are poverty, class size, teacher turnover, tax base, budget support, and English Language Learners. Prior to 2017, the indicators were 1) the percent change in property values, 2) the percent change in enrollment, 3) the number of budget vote defeats, 4) the graduation rate, and 5) the percent of students receiving free or reduced price lunch. In 2017 these measures shifted to include: 1) the percent of economically disadvantaged students, 2) the average class size, 3) the turnover rate of teachers, 4) the percent change in property values, 5), the percentage approval rate of budget measures, and 6) the percent of English language learners.

For both the fiscal and environmental scoring systems, the comptroller's office calculates an aggregate score between 0 and 100 that reflects a district's performance across all categories. Lower scores reflect lower levels of stress. Districts that score less than 25 points are given "no designation." Those scoring between 25 and 45 are labeled "susceptible [to] fiscal stress," between 45 and 65 as "moderate fiscal stress," and above 65 "significant fiscal stress." The environmental scores are on a similar 100 point scale, but the cut-offs are slightly different than for fiscal stress; districts that score between 30 and 45 points are given the label "susceptible [to] environmental stress," between 45 and 60 "moderate environmental stress", and above 60 "significant environmental stress."

Although districts can fall into four different categories for both fiscal and environmental reasons, in practice the number of districts that fall into the "significant stress" categories is low. Between 2013-2021, 92.4 percent of district-years (5,586 out of 6,066 observations) were given no designation; 5 percent were "susceptible to fiscal stress", 1.7 percent were "moderate fiscal stress", and 0.8% percent "significant fiscal stress." The corresponding percentages for environmental stress were 81 percent, 15 percent, 2.8 percent, and 1.3 percent. For these reasons, as I discuss further below, I focus on the difference between organizations that are given no designation and those that are labeled as susceptible to (environmental or fiscal) stress, i.e. those that receive no label and those that receive some form of label that identifies them as in distress.

2.1 Salience of Labels

It is unclear how salient these ratings are to taxpayers. Because local officials must submit annual financial data to the OSC to be in compliance with the FSMS, it is likely that local administrators are knowledgeable about the various designations, and many school districts that score well on the FSMS's metrics advertise their scores on their websites. The OSC issues a press release and announces scores three times a year: late January for school districts, April and September for governments; because local governments do not all have fiscal years that correspond with the school fiscal years (ending June 30), school districts scores are released on a different schedule. However, the New York state department of education also produces school "report cards," which provide information about enrollment, demographics, graduation rates, expenditures per pupil, and staff qualifications, and the FSMS stress ratings do not form part of the report cards. In this sense, the OCS appears to regard the FSMS as a fiscal management tool rather than as a measure of school performance that might be of interest to parents. However, the school report cards contain a vast amount of data that may be difficult for many parents to make sense of. One distinguishing factor of the FSMS labels is that they provide an easy-to-understand heuristic that may stand out among a large amount of complex information.

To further investigate the salience of the FSMS labels, I examine google search data for terms that might provide some indication of the labels' relevance. While the term "fiscal stress" does not produce enough data to analyze searches that are specific to New York, Appendix Figure A1a shows the trend over time in searches for the term "ny school rating" within New York state. The figure most's revealing feature is that the number of searches peaked immediately after the OCS announced school districts labels for the first time in January 2014. For the most part, searches increased to local maxima each year after the OCS' announcement of school district stress labels in late January, with the exception of 2016 when searches appear to have peaked just prior. A similar analysis of the term “nj school rating” (Figure A1b) reveals no such pattern; searches for the term in New Jersey peaked in February of 2017 instead of January 2014. While these results are merely suggestive, they do provide some evidence that parents are aware of the rating system and seek out information pertaining to it.

2.2 Prior Research on FSMS

Two prior pieces of research have studied New York’s system. Both focus on general purpose governments rather than schools. Chung and Williams (2021) use a regression discontinuity design to show that districts that received the mildest fiscal score label were able to improve their fund balance. Yang (2022) uses an event study approach and finds that fiscal stress labels have no effect on municipal bond prices but that a designation of significant fiscal stress lowers housing prices. Neither study examines school districts, which, for reasons discussed above, may be especially responsive to performance indicators, nor do they explore the difference between fiscal and environmental stress labels.

3 Data

This paper analyzes state monitoring data from 2013, when the Fiscal Stress Monitoring System (FSMS) was first rolled out to school districts, through 2021. Fiscal scores, environmental scores, and property values come from the office of the New York state comptroller and are all at the school district-year level. Demographic, enrollment, and assessment data come from the New York state department of education. Test scores have been standardized by grade and by year.[2]

[2]Test scores are missing for the 2019-2020 school year due to coronavirus-related school closured and the cancellation of state testing.

One of the key outcome variables is property values. The FSMS measures property values as the aggregate value of all properties in each school district based on assessment data. In theory, it would be preferable to have property data based on market transactions, and in fact, for local governments the FSMS employs data from the American Community Survey on the median value of owner occupied housing. However, this information is not available for school districts on an annual basis, and consequently this study uses the assessment data instead. The drawback of using assessment data is that it may not accurately reflect market conditions, especially since assessment practices can vary from district to district. However, these shortcomings are mitigated somewhat by the fact that New York State law requires that assessors estimate the market value of all taxable properties every year (New York State Office of Real Property Tax Services, 2021); while assessments are sometimes levied at a percentage of market value, the FSMS data captures the estimated market value ("full value"). Moreover, since the dataset contains repeated observations from districts over time, it is straightforward to adjust for time-invariant differences in assessment practices across districts (see the methods section below). Thus, despite being measured with error, the variable still reflects changes in market values and thus captures taxpayer's willingness to pay for school quality.

Table 1 presents summary statistics. The sample includes 678 school districts over the period 2013-2021. The average fiscal score is 8.5 out of 100, well-below the 25 point cut-off for "susceptible fiscal stress." The average environmental score is 15.8, also well below the 30 point cut-off for "susceptible environmental stress." As noted above, 92 percent of the (district-year) observations in the sample received no fiscal stress label, and 81 percent of observations received no environmental stress label. However, despite the fact that the number of districts receiving a stress label is small in any given year, over the sample period a much higher percentage of districts (30 percent) have received a fiscal stress label at least once, while 50 percent have received an environmental stress label at least once.

4 Methods

To evaluate the effect of stress labels on school district outcomes, I use a regression discontinuity (RD) design. The RD design exploits the fact that the FSMS uses explicit thresholds in its scoring system to assign fiscal and environmental stress labels. A key requirement of the RD design is that individuals are unable to precisely manipulate the assignment variable, a requirement that is satisfied in this case due to the nature of the indicators used and the multivariate nature of the scoring. To manipulate their fiscal scores, districts would need to exert precise control over their fund balance, surplus, and cash on hand. In the case of environmental scores, districts would need to control the value of properties in the district and demographic characteristics of the population. Even if districts were able to exert control over one of the variables constituting either their fiscal or environmental score, such as teacher turnover in the case of the environmental scores, it is not plausible that they would be able to control all of the measures simultaneously. Appendix Figure 2 plots the density distribution of districts around the thresholds for "susceptible fiscal stress" and "susceptible environmental stress," confirming that there is no bunching just below the cut-offs. Standard density manipulation tests (Cattaneo, Jansson and Ma, 2018) also reveal no systematic evidence of manipulation.[3]

This paper focuses on the comparison between districts that received a stress label and those that did not, i.e. it focuses on the impact of a change in the treatment along the extensive margin. Although it is also possible to estimate the impact of changing from one label to another, eg. the effect of a district receiving a "significant" stress label rather than a "moderate" stress, this movement along the intensive margin is likely to be less salient to taxpayers. Moreover, because of the small number of observations receiving the "moderate" or "significant" label, RD analyses at these thresholds are severly underpowered. As a result,

[3]To conduct the manipulation test, I use the default settings of the rddensity command in Stata (Cattaneo, Jansson and Ma, 2018) along with the epanechnikov kernel. This yields p-values of 0.12 for the fiscal scores and 0.86 for the environmental scores, rejecting the null in both cases that there is manipulation at the cut-off.

I confine the results of RD analyses at these higher thresholds to an appendix.

Thus, this paper examines the local average treatment effect (LATE) at the lowest stress label cut-off. Because of the relatively small size of the sample, the baseline specifications use a parametric approach that includes all observations, including those further from the cut-off. In addition, due to the panel nature of the data, the model includes district and year fixed effects. Specifically, I estimate a regression model of the following form:

$$y_{i,t+k} = \alpha + \beta_1 1[r \geq c]_{it} + f(r-c)_{it} + 1[r \geq c]_{it} \times f(r-c)_{it} + \gamma_i + \delta_t + \epsilon_{it} \tag{1}$$

where $y_{i,t+k}$ represents school district outcomes measured at the district-year level k years in the future, r is the running variable (either the fiscal or environmental score), and c is the cut-off between a stress label and no stress label. f is a polynomial function of the running variable, which I estimate separately on both sides of the cut-off. γ_i and δ_t represent district and year fixed effects respectively. Standard errors are clustered by school district. The baseline specification uses a quadratic polynomial, however in robustness checks I show results for a parametric model with linear polynomials as well as non-parametric estimation.

For outcomes, I focus on variables that reflect taxpayers' location choices, such as property values and district enrollment, as well as standard school performance metrics, such as test scores and graduation rates. For the main specifications, I look at property values, district enrollment, the percent of economically disadvantaged students, and third grade math scores. In appendices, I also report results for the high school graduation rate, the percent of students with limited English proficiency (LEP), and additional grade-specific test scores. While some of the outcome variables that I focus on (property values, percent of economically disadvantaged students) also appear in the calculation of the running variable, the running variables uses the realization at time t while the outcome is a realization at time $t+k$. This is standard in the RD literature; see, for example, the literature on the incumbency advantage, in which the running variable is the vote share in an election while the outcome variable is the vote share in a future period (Lee, 2008). However, to confirm that there is no mechanical relationship between the running variable and any of the outcomes of interest, I also conduct falsification tests.

5 Results

5.1 Falsification Tests

Before discussing the main findings, I first present the results of falsification tests in order to validate the regression discontinuity design. I use the baseline RD model to examine the effect of stress labels in years t-1 and year t. Since the OCS cannot tabulate scores for a school district until after the conclusion of the year, it is not possible for a stress label based on results in year t to affect outcomes in either year t or year t-1. Consequently, any significant finding would suggest that the regression model is misspecified. Tables 3 and 4 present the results for the effect of fiscal and environmental labels respectively on the four main outcomes that I concentrate on below: property values, district enrollment, the percent of economically disadvantaged students, and grade 3 math scores. In both tables, none of the coefficients are statistically significant and all are close to zero, thereby affirming the validity of the research design.

5.2 Main Results

Figures 1-2 show graphical RD plots measuring the effect of a fiscal stress label in year t+1 (Figure 1) and the effect of an environmental stress label in year t+1 (Figure 2). The y-axis shows residuals from a regression of the outcome variable on district and year fixed effects; because covariates can improve the precision of RD designs, the residuals provide a clearer picture of the discontinuity in the outcome than a naive plot. Tables 5 and 6 show the corresponding regression results for year t+1, t+2, and t+3. The plots in Figure 1 indicate that there is no discernible effect of a fiscal stress label on any of the outcomes; the confidence intervals, represented by the gray lines, overlap in each case. The results in Table 4 are consistent with this finding. None of the coefficients are statistically significant.

On the other hand, the plots in Figure 2 indicate that the environmental stress label does affect school district outcomes. Figures 2a and 2b show a discernible drop in property values and log enrollment in the year after a district is labeled as environmentally stressed. In table 6, the RD results show that property values decline by approximately 1.5 percent, 1.4 percent, and 2.7 percent in the three years following the label relative to districts that just narrowly avoid being categorized as stressed. All three coefficients are statistically significant. Table 6 also indicates a statistically significant drop in enrollment; three years after a label, environmental stressed districts have experienced a 0.9 percent decline in enrollment relative to districts that narrowly avoided the label. According to both the RD plots and the regression results in Tables 5-6, there are no discernible effects on either the percent of economically disadvantaged students or on math test scores in grade 3.

Tables A1 and A2 present results for additional outcomes, including the high school graduation rate, the percentage of students with limited English proficiency (LEP), and grade 8 test scores. Out of 30 coefficients, only two are significant, and these two show opposing findings. Thus, for the duration of the paper, I focus on property values and enrollment as the primary outcomes of interest. Tables A3 and A4 show results for the thresholds between 1) susceptible and moderate stress and 2) moderate and significant stress. Only one of the 24 coefficients across the two tables is significant, indicating that there are no statistically significant effects of moving from one type of stress label to another, validating the choice to focus on changes at the extensive margin of the treatment rather than the intensive margin.

To summarize, the RD results indicate that a fiscal stress label has no effect on school district outcomes (relative to no label), but that an environmental stress label causes property values to decline on the order of 1-2 percent and enrollment to decline by approximately 0.9 percent.

5.3 Alternative Specifications

To demonstrate the robustness of the finding that an environmental stress label lowers property values and enrollment, I conduct a variety of robustness checks and alternative specifications. First, I employ a parametric specification similar to the baseline approach with a linear rather than a quadratic functional form. Next, I use a nonparametric approach that limits the analytic sample to a small range around the cut-off and uses a triangular kernel. Although it is standard in the literature to employ an "optimal" bandwidth, given the smaller number of observations above the cut-off, it is not possible to compute the mean squared error (MSE) optimal bandwidth above the threshold for every specification, and consequently I use a uniform bandwidth of 10 points.[4] Table 7 presents the results. The results for the linear parametric specifications are extremely similar to the results in Table 6. The coefficients for local linear regression are also similar in magnitude but with much larger standard errors.

Up to this point, I have estimated the effects of a fiscal stress label and an environmental stress label separately. However, as the labels are assigned simultaneously, it may be appropriate to estimate the effects jointly in order to improve precision. The main reason for including covariates in RD analysis is to increase the precision of the estimator rather than to reduce bias, although in some cases covariates may also affect identification if there is concern that the potential outcomes are discontinuous at the cut-off (Calonico et al., 2019). Here, the baseline specifications already include some covariates in the form of the district and year fixed effects. Table 8 shows the results from joint estimation. Once again, the

[4] To compute the optimal bandwidth, I use the rdrobust package in STATA.

pattern of results is unchanged.

Finally, I explore the effects of including treatment effect lags. Because the FSMS recalculates scores and assigns labels on an annual basis, a school district that is labeled as stressed in year t may also have received a similar label the year before. In some cases, fiscal and environmental stress designations remain in place for several years. The RD design estimates an unbiased local average treatment effect, meaning that it estimates the *average* difference in the outcome variable between stressed and non-stressed districts at the cut-off, regardless of their history of previous designations. Districts that receive persistent stress designations mostly appear further from the cut-off, and thus the effect of a stress label on these districts may differ from the LATE that this paper estimates. Nevertheless, to more fully account for treatment effect dynamics, I also estimate a regression that includes indicators for stress labels in previous years. The regression takes the following form:

$$
\begin{aligned}
&y_{i,t+k} = \alpha + \beta_1 1[r \geq c]_{it} + f(r-c)_{it} + 1[r \geq c]_{it} \times f(r-c)_{it} + \\
&\beta_2 1[r \geq c]_{it-1} + f(r-c)_{it-1} + 1[r \geq c]_{it-1} \times f(r-c)_{it-1} + \\
&\beta_3 1[r \geq c]_{it-2} + f(r-c)_{it-2} + 1[r \geq c]_{it-2} \times f(r-c)_{it-2} + \gamma_i + \delta_t + \epsilon_{it}
\end{aligned}
\tag{2}
$$

where β_2 and β_3 represent treatment effects in years t-1 and t-2 respectively. Table 9 presents the results. Once again, the pattern of results remains similar, although the effect on log property values in year t+1 is no longer statistically significant. Thus, the results in this section indicate that the main finding – that an environmental stress label lowers property values and enrollment – is robust to a variety of alternative specifications.

5.4 Heterogeneity

In this section, I explore heterogeneity in the effects. Prior research suggests that affluent families are better positioned to incorporate new information in their assessement of school districts and consequently their location choices (Hasan and Kumar, 2019). Thus, I first explore whether or not the effect of stress labels is more pronounced in wealthier districts. I measure wealth based on property values in the first year of the sample (2013), and examine effects on districts above and below the median. Although I find no effects from fiscal labels in the results discussed above, I explore heterogeneity in the effects of both fiscal and environmental labels, as the effects of fiscal labels may be concentrated in specific subpopulations. The results are in Tables 10 and 11. Table 10 shows that there are no effects of fiscal labels in districts of either above or below average wealth. However, Table 11 shows that, consistent with prior work, the effects of environmental labels are indeed driven by the effects on wealthier districts. For both property values and enrollment, the effects are larger in every year for districts of above average wealth. Only in one case is the coefficient significant for below-average districts.

Next, I explore heterogeneity by school district performance, which I measure using high school graduation rates. Although student performance is correlated with wealth, recent work has shown that student achievement suffers in localities where the gap between teacher wages and private sector wages is high (Marchand and Weber, 2020) and consequently that the two outcomes do indeed represent independent measures of different constructs. Once again, I use the earliest available measure of the variable that captures heterogeneity (in this case, 2014) so that districts remain fixed in a category (above or below the median). Tables 12 and 13 present the results. Once again the fiscal stress label does not appear to have any effect; coefficients are statistically indistinguishable from zero for districts with both above and below average graduation rates. On the other hand, there is evidence of heterogeneity in the effect of an environmental stress label on enrollment; total enrollment declines by 2.5

percent in year t+1 for districts with above average high school graduation rates, while there is no statistically significant effect for below average districts. The effect on property values is less clear, with no systematic pattern in the coefficients; in fact, in year t+1 districts with below average graduation rates appear to suffer greater declines.

Thus, the negative effect of an environmental stress label on strict district enrollment appears to be driven by the effect on wealthier and more high-performing districts. The negative effect on property values is driven by the effect on wealthier districts, with heterogeneity in performance playing less of a role.

5.5 Effects on Enrollment by Grade

The results above show that an environmental stress label – indicating negative changes to a school district's property base and demographic make-up – causes property values and enrollment to decline, especially in wealthier districts. But what is driving this effect? One possible way to explore the underlying mechanism is to examine the effects on enrollment in specific grades. If new parents are forward-looking and base their location choice on all possible information about a school district, then one might expect that the effect of an environmental stress label would be visible primarily on enrollment in lower grades. On the other hand, to the extent that parents are sensitive about school quality, they may be more sensitive about the quality of their children's high school than their elementary school. Prior work shows that the earnings of workers are dependent on the high schools that they attended (Betts, 1995; Dearden, Ferri and Meghir, 2002) and that there is a strong correlation between housing wealth and college choice (Lovenheim and Reynolds, 2013). On the other hand, there are few, if any, documented effects of the choice of elementary school on later-life outcomes.

Table 14 shows the effect of an environmental stress label on elementary school enrollment (grades 1-6) and secondary school enrollment (grades 7-12). The results indicate that the overall enrollment effect is driven by a decline in secondary school enrollment. In years t+1 and t+2, an environmental stress label causes secondary school enrollment to decline by 3.2 and 2.5 percent, while the coefficients on elementary school enrollment are positive and not statistically significant. The effect on secondary school enrollment is thus consistent with a model in which parents are particularly sensitive to the quality of their children's high school, which they regard as playing a key role in determining their later life outcomes. Parents are less sensitive about the quality of their children's elementary school.

6 Discussion and Conclusion

This paper exploits New York state's fiscal monitoring system to examine the effect of stress labels on school district outcomes. Using a regression discontinuity design, it shows that while labeling a school district as fiscally stressed has no effect, labeling a school district as "environmentally stressed" – indicating that the district is experiencing demographic and economic pressures – causes enrollment and property values to decline. Property values decline by 1.5-2.7 percent, and enrollment declines by 0.9 percent. There are no statistically significant effects on test scores or demographic variables. The effects on enrollment and property values appear to be driven by declines in wealthier – and in the case of enrollment, higher performing – districts, where parents may be more sensitive to measures of school quality. The enrollment effects are concentrated at the secondary school level, suggesting that parents are especially sensitive to the quality of the high school that their children attend.

In addition to shedding light on the use of heuristics in school choice, this paper's main contribution is to highlight the importance of economic and demographic information relative to fiscal information in residential and school choice. We know little about how taxpayers respond – either electorally or via location choice – to the fiscal condition of their local governments. In many respect, an important measure of management performance is the ability to maintain a reasonable financial cushion and low levels of debt. However, the results above suggest that taxpayers may not in fact place much weight on effective fiscal management. Instead, they place significantly more weight on "environmental" factors, such as trends in property values, that are leading indicators of a community's economic health in the future. In this sense, the results provide evidence that taxpayers are forward-looking in their assessments of school districts.

The results also dovetail with accumulating evidence on increasing income segregation within the United States (Owens, Reardon and Jencks, 2016; Reardon et al., 2018). Insofar as wealthier residents are better able to incorporate and process new information as part of their decision-making around where to live and what school to send their children to, school quality indicators may not ultimately help to alleviate inequality in the way that they are intended. This is especially true with "environmental" indicators of the sort studied in this paper; unlike fiscal stress labels, which are intended to encourage local districts to act in such a way as to improve their financial condition, there are no obvious tools at the disposal of local school districts that enable them to respond to stressors such as increasing poverty. This suggests that the negative equity consequences of such information provision may outweigh the efficiency gains.

References

Avis, Eric, Claudio Ferraz, and Frederico Finan. 2018. "Do government audits reduce corruption? Estimating the impacts of exposing corrupt politicians." *Journal of Political Economy*, 126(5): 1912–1964.

Betts, Julian R. 1995. "Does school quality matter? Evidence from the National Longitudinal Survey of Youth." *Review of Economics and Statistics*, 231–250.

Black, Sandra E. 1999. "Do better schools matter? Parental valuation of elementary education." *The Quarterly Journal of Economics*, 114(2): 577–599.

Black, Sandra E, and Stephen Machin. 2011. "Housing valuations of school performance." In *Handbook of the Economics of Education*. Vol. 3, 485–519. Elsevier.

Bowman, Kristine L, and Dirk F Zuschlag. 2022. "Balancing state and local power over school districts' finances." *Education Finance and Policy*, 17(3): 564–577.

Calonico, Sebastian, Matias D Cattaneo, Max H Farrell, and Rocio Titiunik. 2019. "Regression discontinuity designs using covariates." *Review of Economics and Statistics*, 101(3): 442–451.

Cattaneo, Matias D, Michael Jansson, and Xinwei Ma. 2018. "Manipulation testing based on density discontinuity." *The Stata Journal*, 18(1): 234–261.

Chung, Il Hwan, and Daniel Williams. 2021. "Local governments' responses to the fiscal stress label: the case of New York." *Local Government Studies*, 47(5): 808–835.

Dearden, Lorraine, Javier Ferri, and Costas Meghir. 2002. "The effect of school quality on educational attainment and wages." *Review of Economics and Statistics*, 84(1): 1–20.

Fack, Gabrielle, and Julien Grenet. 2010. "When do better schools raise housing prices? Evidence from Paris public and private schools." *Journal of Public Economics*, 94(1-2): 59–77.

Figlio, David N, and Maurice E Lucas. 2004. "What's in a grade? School report cards and the housing market." *American Economic Review*, 94(3): 591–604.

Gibbons, Stephen, Stephen Machin, and Olmo Silva. 2013. "Valuing school quality using boundary discontinuities." *Journal of Urban Economics*, 75: 15–28.

Gigerenzer, Gerd. 2004. "Fast and frugal heuristics: The tools of bounded rationality." *Blackwell handbook of judgment and decision making*, 62: 88.

Hasan, Sharique, and Anuj Kumar. 2019. "Digitization and divergence: Online school ratings and segregation in America." *Available at SSRN 3265316*.

Kane, Thomas J, Stephanie K Riegg, and Douglas O Staiger. 2006. "School quality, neighborhoods, and housing prices." *American Law and Economics Review*, 8(2): 183–212.

Lee, David S. 2008. "Randomized experiments from non-random selection in US House elections." *Journal of Econometrics*, 142(2): 675–697.

Lovenheim, Michael F, and C Lockwood Reynolds. 2013. "The effect of housing wealth on college choice: Evidence from the housing boom." *Journal of Human Resources*, 48(1): 1–35.

Marchand, Joseph, and Jeremy G Weber. 2020. "How local economic conditions affect school finances, teacher quality, and student achievement: evidence from the Texas shale boom." *Journal of Policy Analysis and Management*, 39(1): 36–63.

Mikulecky, Marga, and Kathy Christie. 2014. "Rating States, Grading Schools: What Parents and Experts Say States Should Consider to Make School Accountability Systems Meaningful." *Education Commission of the States.*

Nakhmurina, Anya. 2020. "Does fiscal monitoring make better governments? Evidence from US municipalities." *Available at SSRN 3293340.*

New York State Office of Real Property Tax Services. 2021. "Fair assessments: A guide for property owners." *https://www.tax.ny.gov/pdf/publications/orpts/pub1112.pdf*.

Nunes, Luis C, Ana Balcão Reis, and Carmo Seabra. 2015. "The publication of school rankings: A step toward increased accountability?" *Economics of Education Review*, 49: 15–23.

Office of the New York State Comptroller. 2017*a*. "Fiscal stress monitoring system results for municipalties: Four Year Review, 2013-2016."

Office of the New York State Comptroller. 2017*b*. "Fiscal Stress Monitoring System Manual." *www.osc.state.ny.us/files/local-government/fiscal-monitoring/pdf/system-manual.pdf*.

Olken, Benjamin A. 2007. "Monitoring corruption: evidence from a field experiment in Indonesia." *Journal of Political Economy*, 115(2): 200–249.

Owens, Ann, Sean F Reardon, and Christopher Jencks. 2016. "Income segregation between schools and school districts." *American Educational Research Journal*, 53(4): 1159–1197.

Pew Charitable Trusts. 2013. "The state role in local government fiscal distress." *www.pewtrusts.org/-/media/assets/2016/04/pew_state_role_in_local_government_financial_distress.pdf*.

Pew Charitable Trusts. 2016. "State strategies to detect local fiscal distress." *www.pewtrusts.org/en/research-and-analysis/reports/2016/09/state-strategies-to-detect-local-fiscal-distress*.

Pope, Devin G. 2009. "Reacting to rankings: evidence from "America's Best Hospitals"." *Journal of Health Economics*, 28(6): 1154–1165.

Reardon, Sean F, Kendra Bischoff, Ann Owens, and Joseph B Townsend. 2018. "Has income segregation really increased? Bias and bias correction in sample-based segregation estimates." *Demography*, 55(6): 2129–2160.

Simon, Herbert A. 1990. "Bounded rationality." In *Utility and probability*. 15–18. Springer.

Spreen, Thomas Luke, and Caitlin M Cheek. 2016. "Does monitoring local government fiscal conditions affect outcomes? Evidence from Michigan." *Public Finance Review*, 44(6): 722–745.

Thompson, Paul N. 2016. "School district and housing price responses to fiscal stress labels: Evidence from Ohio." *Journal of Urban Economics*, 94: 54–72.

Tiebout, Charles M. 1956. "A pure theory of local expenditures." *Journal of Political Economy*, 64(5): 416–424.

Yang, Lang Kate. 2022. "Fiscal transparency or fiscal illusion? Housing and credit market responses to fiscal monitoring." *International Tax and Public Finance*, 29(1): 1–29.

Figure 1: The Effects of a Fiscal Stress Label

Figure 2a: Log Property Values, t+1

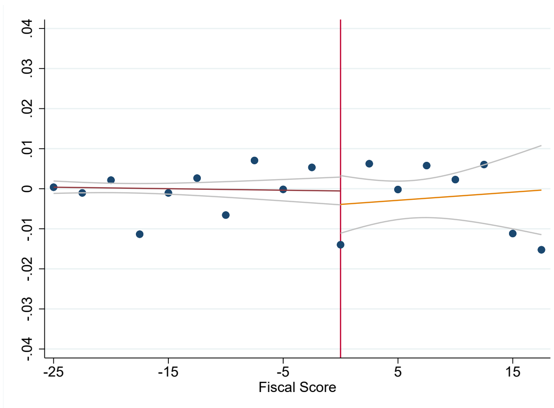


Figure 2b: Log Enrollment, t+1

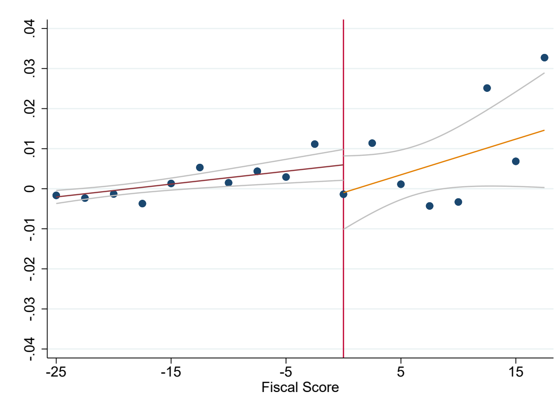


Figure 2c: Percent Disadvantaged, t+1

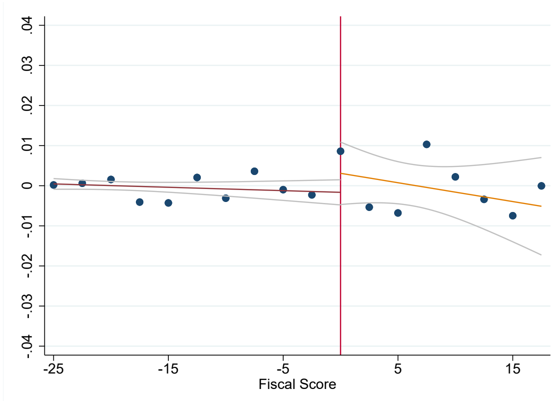
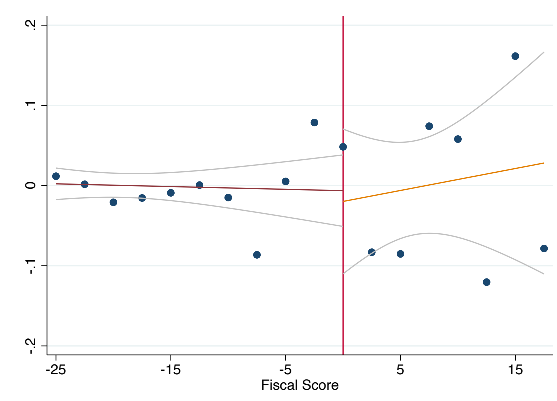


Figure 2d: Grade 3 Math Score, t+1



Note: The figures show school district outcomes on both sides of the fiscal stress label cutoff. The circles represent local sample means based on bins of size 2. The y axis shows residuals from a regression of the outcome variable in year t+1 on district and year fixed effects. The x-axis measures the distance from the cut-off, i.e. the running variable recentered at zero. The grey lines show the confidence intervals for the fitted regression lines.

Figure 2: The Effects of an Environmental Stress Label

Figure 2a: Log Property Values, t+1

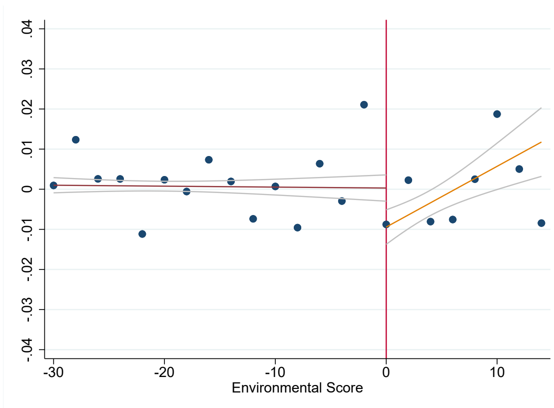


Figure 2b: Log Enrollment, t+1

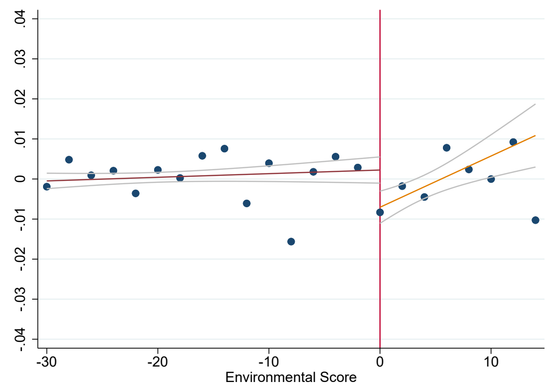


Figure 2c: Percent Disadvantaged, t+1

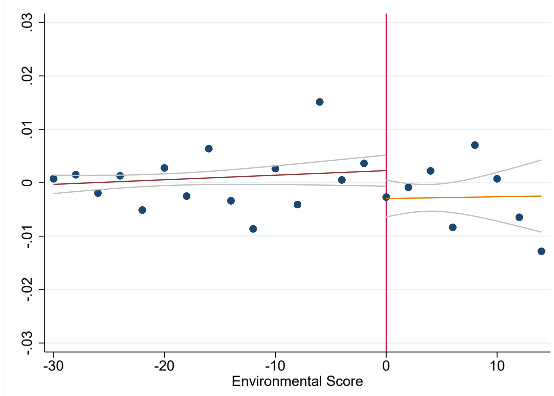
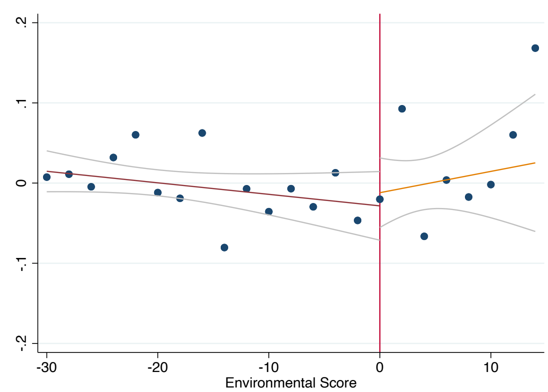


Figure 2d: Grade 3 Math Score, t+1



Note: The figures show school district outcomes on both sides of the environmental stress label cutoff. The circles represent local sample means based on bins of size 2.5. The y axis shows residuals from a regression of the outcome variable in year t+1 on district and year fixed effects. The x-axis measures the distance from the cut-off, i.e. the running variable recentered at zero. The grey lines show the confidence intervals for the fitted regression lines.

Table 1: Financial and Environmental Indicators (2021)

Category		**Financial Indicators**	Max Points
Fund Balance	**1.**	General Fund's Unassigned Fund Balance / (Gross Expenditures - Transfers to Capital Projects Fund)	25
	2.	General Fund's Total Fund Balance / (Gross Expenditures - Transfers to Capital Projects Fund)	25
Operating Deficit	**3.**	Surplus (or Deficit) / (Gross Expenditures - Transfers to Capital Projects Fund)	20
Cash Position	**4.**	General Fund's Total Cash and Short Term Investments / Current Liabilities	10
	5.	Cash as a Percent of Monthly Expenditures	10
Short-Term Debt	**6.**	Percent Change in Short-Term Cash-Flow Debt Issued	10

Category		**Environmental Indicators**	Max Points
Poverty	**1.**	Percent of Economically Disadvantaged Students	25
Class Size	**2.**	Common Branch Class Size	15
Teacher Turnover	**3.**	Turnover Rate of All Teachers	15
Tax Base	**4.**	Percent Change in Property Value	15
Budget Support	**5.**	Budget Vote Approval Percent	15
English Language Learners	**6.**	Percent of English Language Learners	15

Note: Source: Office of the New York State Comptroller. The table shows the method by which the Office of the State Comptroller (OSC) assigns fiscal and environmental stress labels to school districts. For each indicator, the OSC uses a scoring rubric to allocate a certain number of points up to a maximum value. These points are then summed to arrive at a fiscal/environmental score. Districts with fiscal scores between 25-44.9 points are labeled "susceptible fiscal stress," 45-65.9 points "moderate fiscal stress," and 65 points or above "significant fiscal stress." Districts with environmental scores between 30-44.9 are labeled "susceptible environmental stress," 45-59.9 points "moderate environmental stress," and 60 points or above "significant fiscal stress." Office of the New York State Comptroller (2017*b*) provides a more detailed breakdown of the scoring rubrics for the individual indicators.

Table 2: Summary Statistics

VARIABLES	(1) mean	(2) sd	(3) min	(4) max
Fiscal Score	8.5	12.1	0	98.3
Environmental Score	15.8	14.0	0	80
Property Values (2013 $)	1.7×10^9	2.1×10^9	6.6×10^7	2.6×10^{10}
Enrollment	2,219	2,295	8	19,969
Grade 3 Math Scores	0	1	-3.9	4.3
Grade 3 ELA Scores	0	1	-3.3	4.2
Grade 8 Math Scores	0	1	-3.9	3.4
Grade 8 ELA Scores	0	1	-4.7	3.2
4 year H.S. Graduation Rate	0.89	0.09	0	1
Percent Disadvantaged	0.41	0.19	0	1

Note: The sample includes 678 school over the period 2013-2021. Fiscal scores, environmental scores, and property values come from New York state's fiscal monitoring program. Enrollment and school demographic variables come from the New York department of education. Test scores have been standardized by year. 2020 test scores are missing because the state cancelled standardized testing for the 2019-2020 year.

Table 3: Falsification Tests - Fiscal Stress Label

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Log Property Values		Log Enrollment		% Disadvantaged		Grade 3 Math Score	
	t-1	t	t-1	t	t-1	t	t-1	t
	-0.007	-0.006	0.001	0.001	0.004	0.000	-0.057	-0.103
	(0.007)	(0.009)	(0.008)	(0.007)	(0.008)	(0.007)	(0.073)	(0.067)
N	5,323	5,329	5,348	6,034	5,348	6,034	4,538	5,099

Note: ** p > 0.01, * p< 0.05. The table presents regression discontinuity results for the effect of a fiscal stress label in year t-1 and year t. All specifications use district and year fixed effects as well as a quadratic polynomial, estimated separately on both sides of the cut-off. Standard errors are clustered by district.

Table 4: Falsification Tests - Environmental Stress Label

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Log Property Values		Log Enrollment		% Disadvantaged		Grade 3 Math Score	
	t-1	t	t-1	t	t-1	t	t-1	t
	-0.004	-0.004	-0.005	-0.008	-0.006	-0.007	0.033	-0.010
	(0.006)	(0.006)	(0.005)	(0.005)	(0.005)	(0.004)	(0.060)	(0.056)
N	5,323	5,323	5,348	6,032	5,348	6,032	4,538	5,099

Note: ** p > 0.01, * p< 0.05. The table presents regression discontinuity results for the effect of an environmental stress label in year t-1 and year t. All specifications use district and year fixed effects as well as a quadratic polynomial, estimated separately on both sides of the cut-off. Standard errors are clustered by district.

Table 5: The Effects of a Fiscal Stress Label

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
	Log Property Values			Log Enrollment			% Disadvantaged			Grade 3 Math Score		
	t+1	t+2	t+3	t+1	t+2	t+3	t+1	t+2	t+3	t+1	t+2	t+3
	-0.007	-0.006	0.004	-0.007	-0.004	-0.000	0.004	-0.002	-0.007	-0.121	-0.019	-0.065
	(0.007)	(0.006)	(0.006)	(0.007)	(0.006)	(0.005)	(0.006)	(0.006)	(0.005)	(0.078)	(0.072)	(0.083)
N	4,653	3,985	3,315	5,348	4,679	4,009	5,348	4,679	4,009	4,439	3,785	3,140

Note: ** p > 0.01, * p< 0.05. The table presents regression discontinuity results for the effect of a fiscal stress label. All specifications use district and year fixed effects as well as a quadratic polynomial, estimated separately on both sides of the cut-off. Standard errors are clustered by district.

Table 6: The Effect of an Environmental Stress Label

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
	Log Property Values			Log Enrollment			% Disadvantaged			Grade 3 Math Score		
	t+1	t+2	t+3	t+1	t+2	t+3	t+1	t+2	t+3	t+1	t+2	t+3
	-0.015**	-0.014*	-0.027**	-0.007	-0.007	-0.009*	-0.008	-0.006	-0.009	0.049	-0.068	0.014
	(0.005)	(0.006)	(0.005)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.005)	(0.059)	(0.058)	(0.070)
N	4,651	3,983	3,313	5,346	4,677	4,007	5,346	4,677	4,007	4,439	3,785	3,140

Note: ** p > 0.01, * p< 0.05. The table presents regression discontinuity results for the effect of an environmental stress label. All specifications use district and year fixed effects as well as a quadratic polynomial, estimated separately on both sides of the cut-off. Standard errors are clustered by district.

Table 7: Alternative Specifications

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
	Log Property Vales						Log Enrollment					
	t+1		t+2		t+3		t+1		t+2		t+3	
	Linear Parametric	LLR	Linear Parametric	LLR	Linear Parametric	LLR	Linear Parametric	LLR	Linear Parametric	LLR	Linear Parametric	LLR
	-0.009*	-0.022	-0.014**	-0.001	-0.023**	-0.016	-0.007	-0.003	-0.007	-0.008	-0.009**	-0.009
	(0.004)	(0.014)	(0.004)	(0.015)	(0.004)	(0.011)	(0.004)	(0.009)	(0.004)	(0.012)	(0.004)	(0.017)
N	4,651	1,063	3,983	9,62	3,313	859	5,346	1,210	4,677	1,062	4,007	963

Note: ** p > 0.01, * p< 0.05. The table presents regression discontinuity results for the effect of a environmental stress label under alternative specifications. All specifications use district and year fixed effects. The linear parametric regressions include a linear polynomial, estimated separately on both sides of the cut-off. The LLR (local linear regressions) use a linear polynomial estimated separately on both sides of the cut-off with a bandwidth of 10 and a triangular kernel. Standard errors are clustered by district.

Table 8: Estimating the Effects of Both Labels Jointly

	(1)	(2)	(3)	(4)	(5)	(6)
	Log Property Values			Log Enrollment		
	t+1	t+2	t+3	t+1	t+2	t+3
Fiscal Stress	-0.008	-0.006	0.004	-0.006	-0.002	0.000
	(0.007)	(0.006)	(0.005)	(0.007)	(0.006)	(0.006)
Environmental Stress	-0.015**	-0.014*	-0.027**	-0.006	-0.006	-0.009*
	(0.005)	(0.006)	(0.005)	(0.005)	(0.005)	(0.005)
N	4,651	3,983	3,313	5,346	4,677	4,007

Note: ** p > 0.01, * p< 0.05. The table presents the results from RD specifications that jointly estimate the effect of fiscal and environmental stress labels. All specifications use district and year fixed effects as well as quadratic polynomials estimated separately on both sides of the cut-off. Standard errors are clustered by district.

Table 9: Alternative Specification with Lagged Treatment

	(1)	(2)	(3)	(4)	(5)	(6)
	Log Property Values			Log Enrollment		
	t+1	t+2	t+3	t+1	t+2	t+3
	-0.005	-0.011*	-0.024**	-0.004	-0.005	-0.008*
	(0.004)	(0.004)	(0.003)	(0.004)	(0.004)	(0.004)
N	3,307	2,642	1,980	3,995	3,330	2,661

Note: ** p > 0.01, * p< 0.05. The table presents RD specifications for the effect of an environmental stress label that include two years of lagged treatment and running variables. All specifications use school and year fixed effects as well as linear polynomials estimated separately on both sides of the cut-off. Standard errors are clustered by district.

Table 10: Fiscal Stress - Heterogeneity by Wealth of District

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
	Log Property Values						Log Enrollment					
	t+1		t+2		t+3		t+1		t+2		t+3	
	Above Median	Below Median	Above Median	Below Median	Above Median	Below Median	Above Median	Below Median	Above Median	Below Median	Above Median	Below Median
	-0.008	-0.006	-0.008	0.000	0.010	0.004	-0.011	-0.002	-0.002	-0.005	-0.001	0.000
	(0.007)	(0.012)	(0.006)	(0.010)	(0.007)	(0.008)	(0.008)	(0.010)	(0.007)	(0.010)	(0.005)	(0.009)
N	2,334	2,319	2,000	1,985	1,666	1,649	2,661	2,663	2,328	2,330	1,996	1,995

Note: ** p > 0.01, * p< 0.05. The table presents heterogeneity results by wealth of district. District wealth is measured by aggregate property values in 2013. All specifications use district and year fixed effects as well as a quadratic polynomial, estimated separately on both sides of the cut-off. Standard errors are clustered by district.

Table 11: Environmental Stress - Heterogeneity by Wealth of District

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
	Log Property Values						Log Enrollment					
	t+1		t+2		t+3		t+1		t+2		t+3	
	Above Median	Below Median	Above Median	Below Median	Above Median	Below Median	Above Median	Below Median	Above Median	Below Median	Above Median	Below Median
	-0.013*	-0.001	-0.007	-0.002	-0.021**	-0.016**	-0.010	0.009	-0.017**	0.016	-0.012*	-0.003
	(0.006)	(0.009)	(0.009)	(0.008)	(0.007)	(0.006)	(0.006)	(0.009)	(0.005)	(0.010)	(0.005)	(0.010)
N	2,334	2,317	2,000	1,983	1,666	1,647	2,661	2,661	2,328	2,328	1,996	1,993

Note: ** p > 0.01, * p< 0.05. The table presents heterogeneity results by wealth of district. District wealth is measured by aggregate property values in 2013. All specifications use district and year fixed effects as well as a quadratic polynomial, estimated separately on both sides of the cut-off. Standard errors are clustered by district.

Table 12: Fiscal Stress - Heterogeneity by High School Graduation Rate

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
	Log Property Values						Log Enrollment					
	t+1		t+2		t+3		t+1		t+2		t+3	
	Above Median	Below Median	Above Median	Below Median	Above Median	Below Median	Above Median	Below Median	Above Median	Below Median	Above Median	Below Median
	-0.005	-0.004	-0.005	-0.003	0.002	0.008	-0.017	0.001	-0.008	0.000	0.001	-0.008
	(0.006)	(0.012)	(0.007)	(0.010)	(0.008)	(0.008)	(0.009)	(0.009)	(0.007)	(0.008)	(0.007)	(0.007)
N	2,187	2,189	1,874	1,875	1,561	1,558	2,526	2,504	2,210	2,191	1,894	1,876

Note: ** p > 0.01, * p< 0.05. The table presents heterogeneity results by high school graduation rate. Graduation rates are based on 4-year graduation rates in 2014. All specifications use district and year fixed effects as well as a quadratic polynomial, estimated separately on both sides of the cut-off. Standard errors are clustered by district.

Table 13: Environmental Stress - Heterogeneity by High School Graduation Rate

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
	Log Property Values						Log Enrollment					
	t+1		t+2		t+3		t+1		t+2		t+3	
	Above Median	Below Median	Above Median	Below Median	Above Median	Below Median	Above Median	Below Median	Above Median	Below Median	Above Median	Below Median
	-0.002	-0.017*	-0.011	-0.008	-0.022**	-0.024**	-0.025*	0.001	-0.015*	-0.005	-0.011	-0.010
	(0.007)	(0.008)	(0.006)	(0.010)	(0.006)	(0.008)	(0.008)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)
N	2,187	2,187	1,874	1,873	1,561	1,556	2,526	2,502	2,210	2,189	1,894	1,874

Note: ** p > 0.01, * p< 0.05. The table presents heterogeneity results by high school graduation rate. Graduation rates are based on 4-year graduation rates in 2014. All specifications use district and year fixed effects as well as a quadratic polynomial, estimated separately on both sides of the cut-off. Standard errors are clustered by district.

Table 14: Enrollment Effects by Grade

	(1)	(2)	(3)	(4)	(5)	(6)
	Log Enrollment					
	t+1		t+2		t+3	
	Elem	Sec	Elem	Sec	Elem	Sec
	0.009	-0.032**	0.008	-0.025**	-0.001	0.010
	(0.007)	(0.009)	(0.006)	(0.007)	(0.006)	(0.009)
N	5,327	5,171	4,661	4,526	3,993	3,877

Note: ** $p > 0.01$, * $p < 0.05$. The table presents the effects of an environmental stress label on elementary (grades 1-6) and secondary (7-12) enrollment. All specifications use district and year fixed effects as well as a quadratic polynomial, estimated separately on both sides of the cut-off. Standard errors are clustered by district.

Appendix Figure 1: Google Search Trend Analysis

Figure A1a: Google Searches for the Term "ny school rating"

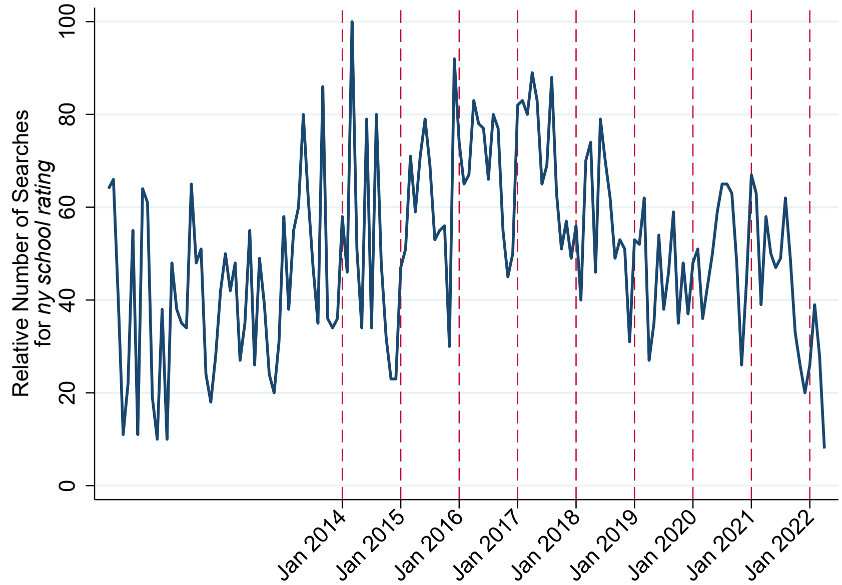
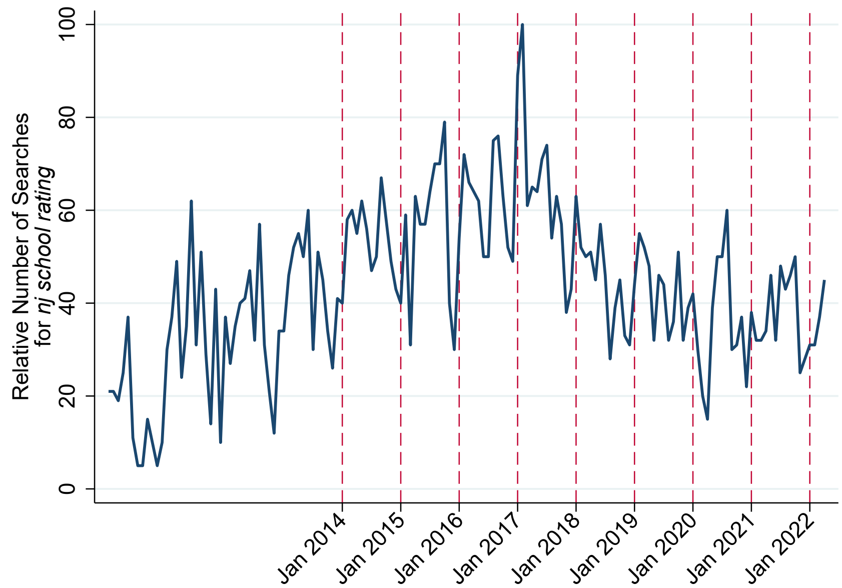


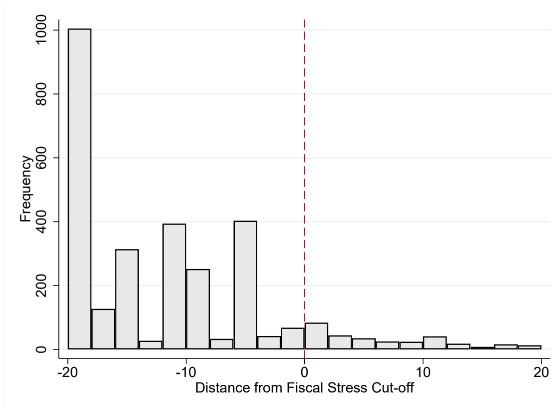
Figure A1b: Google Searches for the Term "nj school rating"

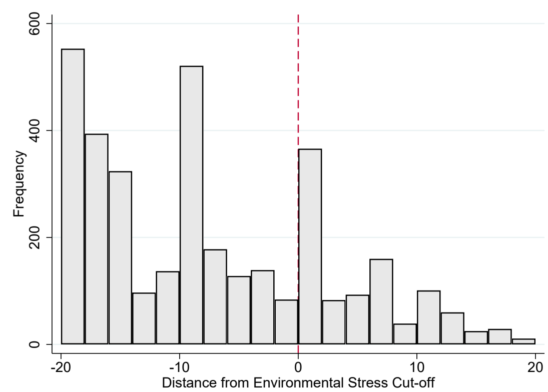


Note: The figure presents evidence of popular interest in New York's school district stress labels using data from Google Trends. Figure A1a shows the relative interest over time in the term "ny school rating." The data covers searches in New York state between January 2010 - April 2022. Interest peaks shortly after the Office of the State Comptroller assigned the labels to school districts for the first time in late January of 2014. For the most part, the use of the term shows a cyclical increase around the same time each year, with the exception of 2016. Figure A1b shows the relative interest over time in the term "nj school rating" in New Jersey over the same time frame. The New Jersey results do not show the same pattern as the New York results, with searches peaking in February 2017. An alternative search using the term "fiscal stress" does not yield sufficient data for an analysis.

Appendix Figure 2: Density Distributions Around the Cut-off

Figure A2a: Fiscal Stress **Figure A2b: Environmental Stress**





Note: The figures show the density distribution of school districts around the cut-offs for a stress label, 2013-2021.

Table A1: Additional Outcomes - Effects of a Fiscal Stress Label

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)
	Graduation Rate			% LEP			Grade 3 ELA			Grade 8 Math			Grade 8 ELA		
	t+1	t+2	t+3	t+1	t+2	t+3	t+1	t+2	t+3	t+1	t+2	t+3	t+1	t+2	t+3
	-0.005	0.001	0.001	0.001	0.005	0.006	-0.038	0.067	-0.031	0.157*	-0.018	0.005	0.139	-0.197**	-0.091
	(0.007)	(0.009)	(0.007)	(0.003)	(0.005)	(0.007)	(0.071)	(0.072)	(0.076)	(0.080)	(0.082)	(0.086)	(0.081)	(0.074)	(0.085)
N	5,025	4,395	3,765	5,348	4,679	4,009	4,436	3,783	3,138	4,172	3,546	2,932	4,330	3,691	3,059

Note: ** $p > 0.01$, * $p< 0.05$. The table presents regression discontinuity results for the effect of a fiscal stress label. All specifications use district and year fixed effects as well as a quadratic polynomial, estimated separately on both sides of the cut-off. Standard errors are clustered by district.

Table A2: Additional Outcomes - Effects of an Environmental Stress Label

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)
	Graduation Rate			% LEP			Grade 3 ELA			Grade 8 Math			Grade 8 ELA		
	t+1	t+2	t+3	t+1	t+2	t+3	t+1	t+2	t+3	t+1	t+2	t+3	t+1	t+2	t+3
	0.005	-0.003	-0.006	-0.003	-0.002	-0.002	0.009	0.012	0.078	0.032	0.046	0.052	-0.086	0.053	-0.019
	(0.005)	(0.005)	(0.006)	(0.001)	(0.001)	(0.002)	(0.055)	(0.058)	(0.063)	(0.068)	(0.067)	(0.076)	(0.060)	(0.062)	(0.066)
N	5,023	4,393	3,763	5,346	4,677	4,007	4,436	3,783	3,138	4,172	3,546	2,932	4,330	3,691	3,059

Note: ** $p > 0.01$, * $p< 0.05$. The table presents regression discontinuity results for the effect of an environmental stress label. All specifications use district and year fixed effects as well as a quadratic polynomial, estimated separately on both sides of the cut-off. Standard errors are clustered by district.

Table A3: RD Results at Susceptiple/Moderate and Moderate/Significant Thresholds for Fiscal Stress

	(1)	(2)	(3)	(4)	(5)	(6)
	Log Property Values			Log Enrollment		
	t+1	t+2	t+3	t+1	t+2	t+3
Moderate Fiscal Stress	0.029	0.011	0.003	-0.014	-0.006	-0.004
	(0.022)	(0.026)	(0.011)	(0.021)	(0.015)	(0.017)
N	385	354	328	414	386	355
Significant Fiscal Stress	0.017	0.008	-0.023	0.015	0.015	0.007
	(0.022)	(0.018)	(0.023)	(0.023)	(0.032)	(0.022)
N	125	116	111	133	125	116

Note: ** $p > 0.01$, * $p < 0.05$. The table presents regression discontinuity results at 1) the cut-off between susceptiple and moderate fiscal stress (labeled "moderate" above) and 2) the cut-off between moderate and significant fiscal stress (labeled "significant" above). All specifications use district and year fixed effects as well as a quadratic polynomial, estimated separately on both sides of the cut-off. The sample range is limited to observations above the cut-off for susceptible fiscal stress. Standard errors are clustered by district.

Table A4: RD Results at Susceptiple/Moderate and Moderate/Significant Thresholds for Environmental Stress

	(1)	(2)	(3)	(4)	(5)	(6)
	Log Property Values			Log Enrollment		
	t+1	t+2	t+3	t+1	t+2	t+3
Moderate Environmental Stress	-0.013	-0.023*	0.001	-0.007	-0.006	-0.002
	(0.010)	(0.011)	(0.008)	(0.008)	(0.009)	(0.007)
N	921	852	786	999	918	851
Significant Environmental Stress	-0.036	-0.044	0.003	-0.009	-0.001	-0.008
	(0.021)	(0.023)	(0.021)	(0.018)	(0.016)	(0.012)
N	163	143	123	187	163	142

Note: ** $p > 0.01$, * $p < 0.05$. The table presents regression discontinuity results at 1) the cut-off between susceptiple and moderate environmental stress (labeled "moderate" above) and the cut-off between moderate and significant environmental stress (labeled "significant" above). All specifications use district and year fixed effects as well as a quadratic polynomial, estimated separately on both sides of the cut-off. The sample range is limited to observations above the cut-off for susceptible environmental stress. Standard errors are clustered by district.