Positive Public Financing Shocks could Increase Local Racial Disparity

Abstract

This paper shows that a positive local financial shock has heterogeneous effects on academic achievement. White students show meaningful improvement, but Black and Hispanic students do not. Consequently, the achievement racial gap widens following the shock. Changes in school funding are not responsible for this phenomenon; rather, it is explained by heterogeneous outcomes in household Socioeconomic Status (SES). Consistent with racial segregation hindering the even distribution of economic gains, the achievement racial gap widens more in more racially segregated areas. These results highlight the possibility that a credit shock induced increase in government spending could perversely increase local racial disparity.

Keywords: Public Financing, Household SES, Human Capital, Racial Gap
JEL Codes: I21, I24, H74

Tuesday 7th December, 2021

I. Introduction

Despite growing national support for civil rights movements, the socioeconomic status (SES) racial gap stubbornly persists in the US (Chetty, Hendren, Jones, and Porter, 2020b). One in five children in the US are growing up in poverty, and more than 60% of them are Black or Hispanic.[1] This persistent SES racial gap significantly affects children's human capital accumulation (Dahl and Lochner, 2012; Deckers, Falk, Kosse, Pinger, and Schildberg-Hörisch, 2021; Hanushek, 2001; Reardon, Kalogrides, and Shores, 2019a; Jang and Reardon, 2019) and has long-lasting impact on intergenerational mobility (Chetty et al., 2020b; Sylwester, 2002). As an important factor in the persistent racial disparity, the role of finance has attracted attention in both academic and policy works. Ample evidence suggest racial minorities do not have equal access to financial markets,[2] yet prior literature generally shows positive local financial shocks can benefit the underrepresented populations and reduce the racial gap.[3] Is this always the case?

This paper examines this question in the context of public financing conditions, the measure of local government's ability to raise funding through municipal bonds. Adelino, Cunha, and Ferreira (2017) provide causal evidence that improvement in public financing conditions positively affects labor market outcomes and household socioeconomic well-being.[4] This paper asks whether these positive effects have an impact on human capital accumulation. Further, do households and students from different race groups benefit equally from an improved public financing condition?

[1] According to National KIDS COUNT. See https://tinyurl.com/27hf863z.

[2] See for example Begley and Purnanandam (2021); Dougal, Gao, Mayew, and Parsons (2019); Haendler and Heimer (2021); Chu, Ma, and Zhang (2021); Stefan, Holzmeister, Müllauer, and Kirchler (2018); Giulietti, Tonin, and Vlassopoulos (2019); Bartlett, Morse, Stanton, and Wallace (2021); Bayer, Ferreira, and Ross (2016); Bayer, Casey, Ferreira, and McMillan (2017); Bayer, Ferreira, and Ross (2018); Fuster, Goldsmith-Pinkham, Ramadorai, and Walther (2020); Chu (2019); Charles and Hurst (2002); Butler, Mayer, and Weston (2020); Blanchflower, Levine, and Zimmerman (2003); Munnell, Tootell, Browne, and McEneaney (1996); Howell, Kuchler, Snitkof, Stroebel, and Wong (2021) among many others.

[3] See for example Levine, Rubinstein, and Levkov (2014); Beck, Levine, and Levkov (2010); Chatterji and Seamans (2012) and Stein and Yannelis (2020).

[4] Guiso, Sapienza, and Zingales (2004) provide a more general examination and show positive effect of local financial development on socioeconomic well-being.

To facilitate causal inference, I follow Adelino et al. (2017) and Cunha, Ferreira, and Silva (2019) and use Moody's municipal bond rating recalibration in 2010 as a positive exogenous shock to local public financing condition.[5] Adelino et al. (2017) show Moody's recalibration significantly increases local government spending and results in various positive spillover outcomes such as increased public and private employment as well as household income. These SES benefits came at a crucial time for the affected children. In 2010, most parts of the country were still recovering from the aftermath of the Great Recession. Parents who were enabled to earn more from the labor market by this recalibration can provide income and stability for the family (Hussam, Kelley, Lane, and Zahra, 2021). These economic and psychological benefits positively impact children's performance at school and the accumulation of their human capital. Indeed, a generalized difference-in-differences (DiD) estimation suggests students from treated counties show significant improvement in test scores. A back of the envelope calculation suggests the effect of the recalibration on education outcomes is comparable to an \$87 increase in school spending per pupil or a \$700 million increase in school spending nationwide.[6] Event-time analysis reveals similar pre-trends in test scores between treated and control counties. This parallel trend condition is particularly important in this setting because the pre-treatment period overlaps with the Great Recession. Jackson, Wigger, and Xiong (2021) and Shores and Steinberg (2019) show the Great Recession negatively affected student outcomes through the school funding channel. The parallel pre-trends between treated and control counties alleviates the concern that the effect on education outcomes documented in this paper is due to the heterogeneous impact of the Great Recession.

I then examine whether the improvement in academic achievement evenly spreads across students of different races. Results from re-estimating the generalized DiD test for each

[5]Corresponding to nearly 70,000 municipal bond issues that worth \$2.2 trillion in total par value. More discussion on the nature of the recalibration is provided in the next section.

[6]Based on the effect of school spending on test scores reported in Abott, Kogan, Lavertu, and Peskowitz (2020). Section III provides more detailed discussion on this economic magnitude.

race group suggest strong heterogeneity in the treatment effect. White students showed significant improvement in test scores. However, there is no significant improvement for Hispanic or Black students. This heterogeneity in treatment effect significantly widens the White-Minority academic achievement gap. A decomposition exercise shows this result is largely orthogonal to the recalibration's impact on achievement income gap (Di Maggio, Kermani, Keys, Piskorski, Ramcharan, Seru, and Yao, 2017). Robustness tests suggest this result is not driven by alternative explanations such as demographic shifts (Cornaggia, Gustafson, Israelsen, and Ye, 2019) or imprecise measurement.

Changes in school funding is not a factor in this heterogeneous treatment effect: There is on average no increase in spending per pupil for school districts in treated counties, and the main results continue to hold for treated school districts that do not increase spending. Consistent with the household's SES being the channel of this effect, there is economically and statistically significant improvement in SES for treated White households, but not Black or Hispanic households. For treated counties, more widening in the SES racial gap is associated with more widening in the test score racial gap. Corroborating this result, only White households in treated counties show an improvement in employment status and a reduction in reliance on the social safety net. Consistent with racial segregation facilitating the persistence of SES racial gaps (Ananat, 2011), the achievement racial gap widens more in more racially segregated areas. Taken together, the findings in this paper suggest even though better public financing can improve household socioeconomic status and human capital accumulation, minority households may benefit less in more segregated areas.

Understanding the effect of local financial shocks on education and SES racial gaps is especially important in light of the COVID-19 pandemic because research papers suggest minority households are more adversely affected in many aspects (Davydiuk and Gupta, 2020; An, Cordell, Geng, and Lee, 2021; Gerardi, Lambie-Hanson, and Willen, 2021). For instance, the Black-White unemployment gap more than doubled as a result of the pandemic.[7]

[7]See https://www.bloomberg.com/graphics/2021-covid-race-and-recovery/.

This unemployment effect can take a serious toll on children both physiologically and psychologically, due to the potential lack of food and shelter as well as stress about the uncertain future (Butler, Demirci, Gurun, and Tellez, 2021; Ananat, Gassman-Pines, Francis, and Gibson-Davis, 2017).

The findings in this paper are especially relevant as the economy recovers from the pandemic. Federal Reserve Chair Jerome Powell commented in May of 2021 that "while the U.S. economic recovery is making real progress, the gains have been uneven following a downturn that cut hard along the lines of race and income."[8] And Janelle Jones, the Chief Economist at the Department of Labor, commented similarly "while headline numbers touted recovery, many Black women and their families and communities continued to struggle. No economic recovery can be complete if some communities are left behind."[9] The results from this paper suggest racial segregation could impede the effort towards an nationwide economic recovery from the pandemic.

This paper contributes to the literature on the interaction between finance and race. There are two main themes in this topic. The first theme focuses on minority individuals' and households' access to financial markets. Chu et al. (2021), Bayer et al. (2017), Bayer et al. (2018), Butler et al. (2020), and Chu (2019) show racial minorities do not enjoy the same rate in the household financial markets (mortgage and auto loans). Minority entrepreneurs and HBCUs face higher costs of capital to fund their business (Blanchflower et al., 2003; Dougal et al., 2019; Howell et al., 2021). When trying to participate in the finance markets, racial minorities are more likely to be mistreated (Begley and Purnanandam, 2021; Stefan et al., 2018) and have a harder time settling a dispute (Haendler and Heimer, 2021). There is also evidence that the situation may not significantly improve even in the digital era: Bartlett et al. (2021) and Fuster et al. (2020) show that algorithms could continue to make biased decisions even without human interaction.

[8]See https://tinyurl.com/3suy8mjw.

[9]See https://blog.dol.gov/2021/02/09/a-more-inclusive-economy-is-key-to-recovery.

The literature's second theme, which this paper relates closely to, focuses on the effect of local financial shocks on racial disparities. Research on this topic generally shows a positive effect of finance on reducing racial inequality. For example, Stein and Yannelis (2020) show banking development can directly improve the SES of the historically marginalized populations by facilitating their financial inclusion. Levine et al. (2014) and Beck et al. (2010) show banking deregulation can indirectly improve the SES for minorities by increasing credit provision to local business and boosting labor demand. Chatterji and Seamans (2012) also provide positive evidence, showing state-level financial deregulation facilitates access to credit for racially underrepresented entrepreneurs. The evidence of positive effects, however, begs the question: If positive local financial shocks always reduce racial inequality, why does the achievement and SES racial gaps persist after centuries of financial development? This paper contributes to the literature by providing an important case highlighting that positive local financial shocks can also negatively affect racial disparity.

This paper also relates to the literature on how local financial conditions affect education outcomes. Using the shale boom in Texas as a positive economic shock, Marchand and Weber (2020) show improvement in local economic conditions could have negative effects on high school student outcomes. Using bank regulatory reform as positive credit shocks, Hu, Levine, Lin, and Tai (2020) also find a negative effect of local financial development on student outcomes. This paper contributes to this literature in two important aspects. First, in contradiction to Marchand and Weber (2020) and Hu et al. (2020), this paper show a positive effect of local financial conditions on education outcomes. This result differs from Marchand and Weber (2020) potentially because the two papers focus on different student groups. Marchand and Weber (2020) focus on high school students, who are drawn into local labor market due to the shale boom. The focus of this paper is on students in 3^{rd} to 8^{th} grades, an age group where the substitution channel from Marchand and Weber (2020) is largely muted. And this result differs from Hu et al. (2020) in part because the macroeconomic conditions are different for the sample periods used in this two papers. Moody's recalibration happened in the middle of 2010, immediately preceded by the Great Recession. On the other hand, bank regulatory reforms used in Hu et al. (2020) started around 1994, which is more economically stable than 2010 and had a significantly lower unemployment rate. In this situation, additional labor market participation by the parents may take away essential time spent with children, and this negative effect could outweigh the positive income effect. Further, Marchand and Weber (2020) and Hu et al. (2020) do not focus on the heterogeneous treatment effect of local financial conditions on children of different races. This paper helps fill in this gap and suggests a potential factor that contributed to the stubborn achievement racial gap.

II. Data and Institutional Background

A. Education, SES, Segregation, and School District Funding Data

Educational outcomes data in this paper is from the Stanford Education Data Archive (SEDA). Leveraging 45 million test scores each year, SEDA is the first dataset that comprehensively covers academic achievement and achievement racial gaps in school districts and counties across the United States from the 2008 to the 2017 school year (Reardon, Ho, Shear, Fahle, Kalogrides, Jang, and Chavez, 2021). County-level data is likely more appropriate for this study because neighbourhood and school segregation often occurs between school districts within a county.[10] In other words, school district level outcomes can mask the effect of within-county racial segregation. Nonetheless, Appendix Table A.7 shows the impact of recalibration on racial disparity continues to hold at the school district level. The dataset is a panel with observations at the county-year-grade-subject level.[11] Students are from 3^{rd} to 8^{th} grade and subjects include Math and English Literature. The measure of the test scores is in standard deviation units of the national distribution. This dataset is widely used in recent literature to analyze various factors that contribute to education outcomes: For

[10] See https://www.urban.org/urban-wire/why-segregation-between-school-districts-matters.

[11] To ensure the results are not driven by outliers, test score observations from counties with less than 100 total tested students are not included in the analysis. Results are unchanged when they are included.

examples, see Abott et al. (2020), Gilpin, Karger, and Nencka (2021), Reardon et al. (2019a), Reardon, Weathers, Fahle, Jang, and Kalogrides (2019b), and Torrats-Espinosa (2020). The advantage of this dataset (compared to other public use education datasets, such as the Early Childhood Longitudinal Study) is that it allows researchers to identify time-series variations in academic achievement at the county level as well as by student race. SEDA contains academic outcomes for all racial groups. However, this variable is not as well-populated outside of White, Black, and Hispanic students, which are the focus of this study.

The summary statistics for SEDA data are reported in Table 1. An average county has about 1,388 students in a given grade in a given year. The mean score for all students is close to zero (-0.04) because the measure is standardized at the national level. And it is not precisely zero because the standardization is based on observations at the school district level instead of the county level. White students (0.11) on average have higher test scores than Black (-0.48) and Hispanic (-0.28) students. There are fewer county-year level observations for Black (131,159) and Hispanic (139,241) students compared to White (262,845) students because, to protect student privacy, the test score is not reported if a race-grade-subject-year category of a given county includes less than 20 test takers.[12] In addition to test scores, SEDA reports an average household socioeconomic status for each county-year-race group. This SES measure is constructed through a principal component analysis using median household income, mother's education level, unemployment rate, poverty rate, SNAP eligibility rate, and single mother headed household rate. SEDA's technical documentation shows this measure is monotonically increasing in income and education level and decreasing in unemployment rate, poverty rate, SNAP eligibility rate, and single mother headed household rate.[13] An average White household (Mean SES = 0.07) has higher SES than average Black (-2.31) and Hispanic (-1.17) households. SEDA also provides a measure for levels of local racial segregation based on minority students' exposure to White students at school. For more

[12] See Reardon et al. (2021) page 37 and their footnote 24.
[13] See Reardon et al. (2021) Table 16 for more detail.

insights, Urban Institute provides a high-quality interactive visualization of this measure at the MSA-level.[14] Although available as a panel variable with both cross-sectional and time-series variations, this paper uses the segregation measure from the 2009 school year to avoid bad controls problem in DiD inference (Angrist and Pischke, 2008). The segregation variable varies little along the time-series, so the results are unchanged when using the time-varying measure instead. The last row of the table reports school district-year level spending per pupil data from the Public Elementary-Secondary Education Finance Dataset in the U.S. Census. Consistent with the number reported in Abott et al. (2020), average spending per pupil by a school district is around 11 thousand dollars in a given year.

B. Moody's Recalibration Background and Data

Moody's Investor Service is one of the big three credit rating agencies and offers bond ratings for a wide variety of financial products. It had two different rating systems before its rating recalibration in 2010. Its Municipal Rating Scale (for municipal bonds) measured the distance to distress, which calculates a municipality's likelihood of default in the absence of additional funding from the government. On the other hand, its Global Rating Scale[15] is designed to measure expected losses. Due to the more conservative ratings under the dual-standard system, Moody's share in the municipal bond market declined in years prior to the recalibration. In March 2010, as an attempt to increase its competitiveness in the municipal bond market, Moody's announced a recalibration of its Municipal Rating Scale to align it with the Global Rating Scale (Moody's, 2010). Shortly after the initial announcement, Moody's announced specific plans that resulted in upgrades of nearly 70,000 bond ratings in April and May 2010. This recalibration is at the local government unit level and not all municipal bond issues were upgraded in the recalibration. For instance, a local government is not affected by the recalibration if it receives the same rating under the Municipal Rating Scale and the Global Rating Scale. Municipal bonds with higher ratings were also less likely

[14]See `https://www.urban.org/features/segregated-neighborhoods-segregated-schools`.
[15]For sovereign bonds, corporate bonds, and structured finance products.

to be recalibrated than those with lower ratings. Finally, local governments without Moody's ratings or with no outstanding bonds were not subject to recalibration and can also be used in the control group. As demonstrated in Figure A.1, this diversity in treatment intensity provides abundant geographical variation to facilitate causal inference. Also importantly for the causal inference in the setting of this paper, Moody's (2010) specifically clarifies that the recalibration is intended solely to enhance the comparability of ratings across asset classes, and is not associated with a change in the credit quality of the issuer. Cornaggia et al. (2019) also provide evidence that this recalibration is not correlated with local economic indicators such as household income, poverty rate, population, and net Adjusted Gross Income (AGI) flow.[16] The list of recalibrated bond issues is from Moody's and includes the rating of each bond issue before and after the recalibration. The recalibration covered 69,657 municipal bonds, most of which had an investment-grade rating before the recalibration. The impact of this recalibration on local government financing is economically large: Adelino et al. (2017) show 13-14 bps lower issuance cost and 16%-20% increase in issuance amount following the recalibration. According to Adelino et al. (2017), an average treated county issues about $250 million. So the 16%-20% increase corresponds to around $40-$50 million additional issuance. The summary statistics for this recalibration are reported in Table 1: 31% of counties (962 out of 3,103) received recalibration in 2010. Following Cunha et al. (2019), this paper also includes the continuous recalibration intensity variable, measured as the fraction of the county's local government units that is recalibrated, to capture the intensive margin of the treatment effect. Unconditionally, about 3% of an average county's local government units are recalibrated.

[16]See Cornaggia et al. (2019) Figure 2.

III. The Effect of Public Financing on Test Score

A. Baseline Estimation

This section examines whether municipal financing condition affects academic achievement using the following generalized difference-in-differences estimation:

$$Test\ Score_{c,t,g,i} = \beta * (Post_t \times Recalibration_c) + \gamma_c + \gamma_{s\times t\times g\times i} + \epsilon_{c,t,g,i} \tag{1}$$

Where $Test\ Score_{c,t,g,i}$ is the standardized test score in county c, year t, grade g, and subject i. $Post_t = 1$ for the year 2011 and afterward because Moody's recalibration happened in the middle of 2010.[17] As discussed in the previous section, $Recalibration_c$ can be dichotomous (*Recal. Indicator*) or continuously (*Recal. Intensity*) measured and results for both specifications are reported to explore both the extensive and intensive margin effects. The β coefficient (DiD estimator) on the interaction term will capture the marginal response to municipal bond rating recalibration. County fixed effects (γ_c) are included to control for time-invariant local factors like culture and preference for education. State-year-grade-subject fixed effects ($\gamma_{s\times t\times g\times i}$) are also included to capture any source of state-specific trends and time-varying unobserved state-level heterogeneity, such as changes in transfers from state governments and difficulty of tests. County-level control variables are not included in the main specification to avoid bad controls problem in DiD setting (Angrist and Pischke, 2008).[18]

Results from estimating Eq. (1) are reported in Table 2. Panel A of the table reports the results using a dichotomous recalibration indicator. The indicator equals one if the county receives recalibration and equals zero otherwise. Column 1 is based on the universe of both Math and English tests. The result suggests test scores improved by 0.013 standard deviations after municipal bond rating recalibration. Columns 2 and 3 report the estimations for English Literature and Math, respectively. There are similar improvements for these two subjects. To put the economic magnitude of these results in context, using the same test

[17] Appendix Table A.1 show results are unchanged when excluding all observations from the 2010 school year.

[18] Appendix Table A.2 show results are unchanged after controlling for race-specific household income, mother's education, and family structure.

score data from SEDA, Abott et al. (2020) estimate a \$1,000 increase in direct investment per pupil is associated with a 0.15 standard deviation increase in test scores. Under the assumption of a linear relationship, the effect captured in this paper is comparable to a \$87 ($= 0.013/0.15 \times 1,000$) increase in spending per pupil. An average county-grade has 1,388 students, and there are six grades in the sample. So this \$87 per pupil translates into \$0.72 million ($= 87 \times 1,388 \times 6$) per treated county.[19] There are 962 treated counties, so the nationwide effect is approximately \$694 ($= 0.72 \times 962$) million. Panel B of the table reports estimations using a continuously measured recalibration intensity. By definition, this measure takes into account the intensive margin of the treatment. It is reasonable to expect an intensive margin effect because the labor market impact of recalibration is stronger when more units of local government are upgraded and more jobs are created. Compared to students from a county that received no upgrades, those from a county that has all of its local government units upgraded experience a 0.065 standard deviation increase in their test scores. This intensive margin result suggests the effect on test scores is considerably larger for counties that received more boost in public financing condition through the recalibration. To illustrate this result is not driven by any specific outlier state, Appendix Figure A.2 reports a robustness check exercise for the DiD estimator by dropping one state from the sample at a time. All the permutations in the figure are significant at the 5% level, indicating the result is not likely to be driven by outliers.

B. Event-time Estimation

Event-time results are reported in Panel (a) of Figure 1. Specifically, the recalibration indicator variable is interacted with the indicator variables for each school year and the resulting coefficients are tabulated in the figure. The baseline is the 2008 school year, the first year of SEDA data. Effects post 2013 are aggregated to the "2013+" indicator for brevity and symmetry; alternatively, full time-series results are tabulated in Figure A.3. By construction,

[19] As a benchmark, calculated using estimates from Adelino et al. (2017), a treated county on average increases issuance amount by about \$40 to \$50 million.

these coefficients capture the time-series of differences in academic achievements between treated and control counties. Importantly, there is no significant difference in academic achievement between the recalibrated and control counties in 2008 and 2009, indicating that the parallel trend condition is likely satisfied for the DiD setting. This condition is especially important considering the timing of the Great Recession, which overlaps with the pre-treatment period. Negative economic conditions like the Great Recession hinder academic performance (Jackson et al., 2021; Shores and Steinberg, 2019). The parallel trend suggests students from treated and control counties are similarly affected, indicating the results captured in this paper are not a simple manifestation of heterogeneous treatment effect of the Great Recession. Although Moody's recalibration happened in the middle of 2010, Figure 1 suggests there is no effect in that year. This lack of immediate response indicates there could be a short delay between the county's increased ability to raise funding and the materialization of the positive effect. There is a 0.009 standard deviation improvement in 2011 and a similar magnitude persists through 2012. Finally, there is a strong and persistent long-run effect in the 2013-2017 period. This long-run effect is sensible because students are in the sample for consecutive years through their 3^{rd} to 8^{th} grades and an impact in one year would have meaningful chain effect in the following years. For example, if a student falls behind in their 5^{th} grade, it is likely to also negatively impact their performance in 6^{th} through 8^{th} grades.

C. Accounting for Multiple Hypothesis Testing

Heath, Ringgenberg, Samadi, and Werner (2020) suggest researchers need to be careful with statistical inference when reusing an experimental setting. To address this concern, a total of 13 outcome variables on Moody's recalibration are gathered to conduct three widely used adjustments for multiple hypothesis testing: Bonferroni, Holm, and Benjamini, Hochberg, and Yekutieli (BHY) following Harvey, Liu, and Zhu (2016).[20] Results are reported

[20]The 13 variables include the academic achievement from this paper and the 12 variables from other academic papers using the same experimental design.

in Appendix Table A.3. Common parameters in these three tests are the 5% unadjusted significance level ($\alpha_\omega = 5\%$) and the 13 outcome variables ($M = 13$). The Bonferroni method rejects any hypothesis with *p-value* $\leq \frac{\alpha_\omega}{M}$. Here the adjusted *p-value* for 5% statistical significance is 0.38% (= 5% / 13). The Holm method rejects any hypothesis with *p-value* $\leq \frac{\alpha_\omega}{M+1-b}$. Here the t-stat for the academic achievement estimation ranks 8^{th} on the ordered list. Thus, the adjusted *p-value* for 5% statistical significance is 0.83%. Finally, the BHY method rejects any hypothesis with *p-value* $\leq \frac{\alpha_\omega \times b}{M \times c(M)}$. Following Benjamini and Yekutieli (2001), $c(M)$ is set to equal $\sum_{j=1}^{M} \frac{1}{j}$. The t-stat on the academic achievement variable ranks 8^{th} on the ordered list. Hence, the BHY method adjusted *p-value* for 5% statistical significance is 0.97%. The p-value from column (1) of Table 2 is 0.20%, comfortably below the thresholds for all three methods. The null hypothesis that municipal financing does not affect education outcomes can be rejected at the 5% level even after adjusting for multiple hypothesis testing.

IV. The Effect of Recalibration on Achievement Racial Gaps

A. Baseline Estimation

Did the recalibration benefit students of all races equally? To analyze the heterogeneous treatment effect of public financing on student outcomes, this section repeats the DiD analysis described in Eq. (1) but replaces the dependent variable with test scores by each racial group with the subscript r denoting race:

$$Test\ Score_{c,t,g,i,r} = \beta * (Post_t \times Recalibration_c) + \gamma_c + \gamma_{s \times t \times g \times i} + \epsilon_{c,t,g,i,r} \tag{2}$$

As well as the White-Minority achievement gap:

$$Score\ Racial\ Gap_{c,t,g,i,r} = \beta * (Post_t \times Recalibration_c) + \gamma_c + \gamma_{s \times t \times g \times i} + \epsilon_{c,t,g,i,r} \tag{3}$$

The results are reported in Table 3. Column (1) of Panel A estimates a 0.022 standard deviation increase in test scores for White students. Columns (2) and (3) indicate the coefficients for Black and Hispanic students are slightly positive but not statistically significant. In other words, it can not be concluded that there is a reliable improvement in academic achievement for racially underrepresented students. Consequently, columns (4) and (5) show the White-Black and White-Hispanic academic achievement gaps widen by 0.017 and 0.018 standard deviations, respectively. The inference is unchanged when the continuous measure of recalibration intensity is used in Panel B. The intensive margin effect suggests the achievement racial gap widens more in counties that have more units of local government recalibrated.

B. Event-time Estimation

Event-time results by student race are reported in Panel (b) of Figure 1. For White students, there is no pre-recalibration trend in the 2009 and 2010 school year. Similar to the full sample result in Panel (a), there is a significant and persistent improvement after recalibration for White students. Hispanic students in treated counties show a statistically insignificant improvement in pre-recalibration period and a slight decline in post-recalibration period, but none of the estimates are statistically significant. Black students also display a mild improvement after the recalibration, but the effect is economically small compared to White students and statistically indistinguishable from zero. Full time-series plot in Panel (b) of Appendix Figure A.3 reveals more detail on these dynamics. For White students, the effect stabilizes after 2014 and persists through the end of the sample period. For Black and Hispanic students, the small initial positive effect diminishes over the sample period and becomes economically indistinguishable from zero. The positive public financing shock had a persistent effect on White students, but not racially underrepresented students.

C. Alternative Explanations

C.1. Is it the Achievement Race Gap or the Achievement Income Gap?

On average, students from more affluent households do better at school (Dahl and Lochner, 2012). There is also well documented correlation between achievement racial gap and achievement income gap due to the persistent difference in income between White and minority households (Reardon et al., 2019b; Card and Rothstein, 2007; Chetty et al., 2020a). If the recalibration increased the achievement income gap, there could be a mechanical effect on the test score racial gap by its correlation with the income gap. Is the achievement racial gap captured in this paper just a simple manifestation of the test score differences between poor and affluent students? To address this possibility, in the spirit of Reardon et al. (2019b), I decompose the β coefficient from Eq. (3) into a race component (β_1) and an income component (β_2) using the following system:

$$\begin{cases} Score\ Racial\ Gap_{c,t,g,i,r} = \beta_1 * (Post_t \times Recalibration_c) + Score\ Income\ Gap_{c,t,g,i} + \gamma_c + \gamma_{s\times t\times g\times i} + \epsilon_{c,t,g,i,r} \\ Score\ Income\ Gap_{c,t,g,i} = \beta_2 * (Post_t \times Recalibration_c) + Score\ Racial\ Gap_{c,t,g,i,r} + \gamma_c + \gamma_{s\times t\times g\times i} + \epsilon_{c,t,g,i,r} \end{cases} \tag{4}$$

Where the "Score Income Gap" captures the within-county test score difference between "non-poor" and "poor" students, regardless of student race.[21] The intuition is that if the β coefficient from Eq. (3) captures the sum of both the effect on racial gap and the effect on income gap, then the first model of Eq. (4) partials out the income effect and the β_1 coefficient represents the effect of recalibration on the test score racial gap. Similarly, the second model of Eq. (4) partials out the race effect and the β_2 coefficient represents the effect of recalibration on the test score income gap. And the sum of β_1 & β_2 should be close to the β from Table 3. Results for these estimations are reported in Table A.4. The "Test Score N (Non-poor) - P (Poor) Gap" represents the "Score Income Gap". Panel A (B) reports results from decomposing the White-Black (White-Hispanic) test score gap. Columns (1) & (2) report results using the dichotomous measure of recalibration while columns (3) & (4) use the continuous measure. The results from this table are interpreted in the following way: Take columns (3) & (4) from Panel A for example, β_1 is 0.032 and β_2 is 0.012 and they sum up closely to the β coefficient (0.046). This result suggests that, in this model, about three quarters $[= 0.032/(0.032 + 0.012)]$ of the effect captured in Table 3 is attributable to the impact of recalibration on the test score racial gap. Across the different models from Table A.4, β_1 is about 50% - 75% the size of β coefficient from Table 3. The results from this decomposition suggest Moody's recalibration had a significant impact on the test score racial gap even after accounting for its impact on the test score income gap.

[21] According to SEDA, the categorization of "poor" is done by ED*Facts* and the cutoff varies by state.

C.2. Migration/Demographic Change?

Cornaggia et al. (2019) suggest better public financing induced by Moody's recalibration could attract low wealth households.[22] Are the results captured in this paper driven by migration? The following evidence suggest this is not the case. First, the migration effects captured in Cornaggia et al. (2019) are concentrated in populations nearing retirement age. In other words, the people that the recalibration attracts are not likely to have kids in the 3^{rd} to 8^{th} grade. Indeed, Cornaggia et al. (2019) also find there is no effect on migration in the [0,19] age group. In this case, the change in the test score is unlikely to be driven by migration. Second, in Table A.5, I re-estimate the DiD effect from Table 3 using only the subsample of counties that do not experience any increase in the fraction of Black or Hispanic population between 2009 (pre-recalibration) and 2011 (post-recalibration). Results from this table show that even in counties without migration, there is still a significant widening of White-Minority achievement gap and the magnitude is similar to the full sample results. These two pieces of evidence indicate the heterogeneous effect captured in this paper is not likely to be driven by migration or demographic shift.

C.3. Imprecise Measurement for Minority Students?

There are more observations for White students in the sample than for minority students. Could the insignificant improvement within minority students be driven by a lack of precision in measurement? The following evidence suggest this is not the case. First, aside from statistical insignificance, the economic magnitudes of the coefficients reported in Table 3 for Black and Hispanic students are small. This insignificance in magnitude suggests there are no economically meaningful improvements for Black or Hispanic students compared to White students. Second, Table A.6 shows the inference is unchanged (compared to Table 3) when re-estimating the impact of recalibration on White students using only the subsample

[22] Yi (2021) arrives at a similar conclusion using a different exogenous shock to public finance.

of counties that also have a substantial amount of minority students.[23] Furthermore, the standard error for White students are quantitatively similar to those for Black and Hispanic students. These results suggest the widening of the achievement gap is not driven by imprecise measurements for minority students.

V. What are the Channels for this Heterogeneity in Treatment?

A. The Household Socioeconomic Well-being Channel

If the various SES benefits documented in Adelino et al. (2017) were primarily captured by White households, it may not be surprising that only White students show improvement on their tests. This section studies the heterogeneous treatment effect of Moody's recalibtation on household socioeconomic well-being as a potential explanation utilizing the county-year-race level SES variable provided by SEDA. The test is in a generalized DiD setting similar to Eq. (1), but now the unit of observation is at the county-year-race (c,t,r) level:

$$SES_{c,t,r} = \beta * (Post_t \times Recalibration_c) + \gamma_c + \gamma_{s\times t} + \epsilon_{c,t,r} \tag{5}$$

Results are reported in Table 4. Column 1 shows there is a significant improvement in White households' socioeconomic status. The improvement is also economically meaningful: Based on the coefficient from Panel A, White households in treated counties show 22% (= 0.02/0.09) improvement from the mean. The recalibration does not have a statistically significant effect on Black households' SES. Consistent with the downward trend in test scores in post-recalibration period, the socioeconomic status of Hispanic households are negatively affected, although the economic magnitude is small compared to the mean of -1.17.

To further tighten the relationship between SES and test scores, Table 5 examines the correlation between changes in SES and test score (gap) for recalibrated counties between 2009 and 2017. Consistent with SES being the channel for the effect of recalibration on test

[23] Specifically, columns (1) & (2) only use the subsample of counties that have non-missing estimates for Black students and columns (3) & (4) only use the subsample of counties that have non-missing estimates for Hispanic students.

scores, column (1) of Table 5 shows more increase in SES is associated with more improvement in test scores. Columns (2) & (3) of Table 5 suggest for the treated counties, more increase in SES racial gap is associated with more widening in the test score racial gap.

Adelino et al. (2017) suggest employment is an important aspect of socioeconomic well-being that is improved by the recalibration. Is there heterogeneous treatment effects for White and minority employment status? Table 6 provides corollary evidence that there is a significant reduction in the unemployment rate for White labor force, but not for Black or Hispanic labor forces. For White labor force, column (1) in Panel A of Table 6 suggests there is 0.12 percentage points decrease in the unemployment rate, or a 1.7% (= 0.12%/7%) decrease from the mean. For Black labor force, although the magnitude of the coefficient in column (2) is close to that from column (1), it is small relative to the mean unemployment rate (13%) and is statistically insignificant. For Hispanic labor force, the coefficient in column (3) is also small and insignificant. In short, the recalibration only resulted in positive employment outcomes for White households and not minority households. This employment effect allows the treated White parents to provide a sound family environment for their children to performance well at school.

Importantly, the changes in the extensive margin of labor market participation shown in Table 6 do not paint the full picture of the socioeconomic benefits. For instance, an individual who was never unemployed throughout the sample period could still be positively affected by the recalibration if it enabled them to work for longer hours or secure an additional part-time job. To this extent, reliance on the social safety net (such as the food stamp) may be a better measure for the intensive margin effect of labor market outcomes. Table 7 examines this effect. Results from this test suggest some White households show reduced reliance on SNAP, but the effects are muted for Black and Hispanic households. Taken together, the tests in this subsection show household SES is likely an important channel for this heterogeneous treatment effect.

B. The School Funding Channel

Abott et al. (2020) show increase in school district funding can positively affect student outcomes. As an important component of local government units, school districts' funding could be boosted by the recalibration. Is the school funding a channel for the impact of Moody's recalibration? To explore this possibility, I gather school district level public school funding data from the Public Elementary-Secondary Education Finance Dataset in the U.S. Census. As a sanity check, I first confirm the main findings on the test score and test score racial gaps continue to hold using school district (SD) level data and report results in Panel A of Appendix Table A.7. I then examine whether recalibration resulted in a change in school funding using the following model at the school district-year level:[24]

$$ln(Spending\ per\ Pupil)_{sd,t} = \beta * (Post_t \times Recalibration_c) + \gamma_c + \gamma_{s\times t} + \epsilon_{sd,t} \quad (6)$$

Following Abott et al. (2020), I use the natural logarithm of the dollar amount of spending per pupil per school district year as the outcome variable. Results are reported in Table 8. The first two columns show there is no significant effect of recalibration on school spending in the full universe of school districts. However, it could still be possible that the statistically "predominantly White" school districts reacted differently to the recalibration. Columns (3) & (4) explore this possibility by focusing on school districts with more than 95% White students and still find no effect on spending. Finally, Panel B of Appendix Table A.7 replicates column (1) from Table 2 and columns (4) and (5) from Table 3 for school districts that do not increase spending from the 2009 to 2011 school year. Estimates for this subsample are quantitatively similar to the full sample results reported in Panel A. Results from the first two columns in this panel show there is significant improvement in test scores even for school districts that do not show any increase in spending. Results from the last four columns show that achievement racial gaps widen even for school districts that do not show any increase in

[24] See https://www.census.gov/programs-surveys/school-finances.html. The "Individual Unit Tables" for each year contains the school-district level data used in this paper. The merging identifiers are "NCESID" from Census and "sedalea" from SEDA. The Education Finance dataset includes county identifier (CONUM) that allows merging with recalibration data.

spending. Taken together, the results in this section suggest change in school funding is not likely a channel for the main results.

VI. The Role of Segregation

Prior literature suggests residential racial segregation could be an important source of inequality in income distribution (Ananat, 2011). Without specific incentives that promote a diverse workforce, racial segregation could cause an uneven distribution of employment opportunities even when the employers are not racially biased: When new jobs are created, employers are more likely to hire from their own social network. Similarly, job seekers are more likely to search for jobs within their own network. In highly segregated counties, this type of preference means hiring is mechanically more likely to be done within-race.

Is segregation facilitating the heterogeneous treatment effect from Moody's recalibration? The following triple-diff test analyzes whether racial segregation plays a role in the heterogeneous treatment effect on education outcomes:

$$Score\ Gap_{c,t,g,i} = \beta * (Post_t \times Recalibration_c \times Segregation_c) + \gamma_c + \gamma_{s\times t\times g\times i} + \epsilon_{c,t,g,i} \quad (7)$$

And the results are reported in Table 9. Consistent with the hypothesis that racial segregation exacerbates the uneven distribution of economic gain, the achievement racial gap widens more in treated counties with a higher level of racial segregation. For a one-standard-deviation increase in White-Black (White-Hispanic) segregation, the recalibration results in 0.009 (0.007) standard deviation wider test score gap.[25] For robustness, Appendix Table A.8 report consistent results using an alternative segregation measure from Opportunity Insights (OI).[26] The segregation measure from OI is based on residential, instead of school segregation. Although, according to the Urban Institute, these two types of segregation are highly correlated.[27] The downside of this alternative measure is it does not differentiate

[25]Calculated using the dichotomous measure of recalibration. $0.009 = 0.13 \times 0.068$; $0.007 = 0.10 \times 0.073$.
[26]See https://opportunityinsights.org/data/ for more detail.
[27]See https://tinyurl.com/mywxkbfz and https://tinyurl.com/9ttp4ye3.

between White-Black and White-Hispanic segregation. Consequently, although the point estimates in Table A.8 are similar to those in Table 9, the statistical inference is not as precise for Hispanic students. Finally, Appendix Table A.9 provides supporting evidence that SES gaps also widen more in more segregated areas. Taken together, the results from this section show that racial segregation facilitated the heterogeneous treatment effect of the recalibration.

VII. Conclusion

This paper analyzes the heterogeneous effect of local financial shock on students of different races. Following Moody's municipal bond rating recalibration in 2010, only White households in treated counties show significant improvement in their socioeconomic well-being. This heterogeneous treatment effect has important intergenerational impact on human capital accumulation: White students from treated counties show significant improvement in their test scores but there is no effect for Black or Hispanic students. Consequently, the achievement racial gap widens as a result of a positive county-wide financial shock. The phenomenon is concentrated in racially segregated areas, suggesting that racial segregation could hinder the even distribution of economic benefits from an improved public financing condition.

References

Abott, Carolyn, Vladimir Kogan, Stéphane Lavertu, Zachary Peskowitz. 2020. School district operational spending and student outcomes: Evidence from tax elections in seven states. *Journal of Public Economics*, **183** 104–142.

Adelino, Manuel, Igor Cunha, Miguel A Ferreira. 2017. The economic effects of public financing: Evidence from municipal bond ratings recalibration. *The Review of Financial Studies*, **30** 3223–3268.

An, Xudong, Larry Cordell, Liang Geng, Keyoung Lee. 2021. Inequality in the time of covid-19: Evidence from mortgage delinquency and forbearance. *Working Paper*.

Ananat, Elizabeth O, Anna Gassman-Pines, Dania V Francis, Christina M Gibson-Davis. 2017. Linking job loss, inequality, mental health, and education. *Science*, **356**(6343) 1127–1128.

Ananat, Elizabeth Oltmans. 2011. The wrong side(s) of the tracks: The causal effects of racial segregation on urban poverty and inequality. *American Economic Journal: Applied Economics*, **3**(2) 34–66.

Angrist, Joshua D, Jörn-Steffen Pischke. 2008. *Mostly harmless econometrics: An empiricist's companion*. Princeton University Press.

Bartlett, Robert, Adair Morse, Richard Stanton, Nancy Wallace. 2021. Consumer-lending discrimination in the FinTech era. *Journal of Financial Economics*.

Bayer, Patrick, Marcus Casey, Fernando Ferreira, Robert McMillan. 2017. Racial and ethnic price differentials in the housing market. *Journal of Urban Economics*, **102** 91–105.

Bayer, Patrick, Fernando Ferreira, Stephen L Ross. 2016. The vulnerability of minority homeowners in the housing boom and bust. *American Economic Journal: Economic Policy*, **8**(1) 1–27.

Bayer, Patrick, Fernando Ferreira, Stephen L Ross. 2018. What drives racial and ethnic differences in high-cost mortgages? the role of high-risk lenders. *The Review of Financial Studies*, **31**(1) 175–205.

Beck, Thorsten, Ross Levine, Alexey Levkov. 2010. Big bad banks? the winners and losers from bank deregulation in the united states. *The Journal of Finance*, **65**(5) 1637–1667.

Begley, Taylor A, Amiyatosh Purnanandam. 2021. Color and credit: Race, regulation, and the quality of financial services. *Journal of Financial Economics*.

Benjamini, Yoav, Daniel Yekutieli. 2001. The control of the false discovery rate in multiple testing under dependency. *Annals of Statistics* 1165–1188.

Blanchflower, David G, Phillip B Levine, David J Zimmerman. 2003. Discrimination in the small-business credit market. *Review of Economics and Statistics*, **85**(4) 930–943.

Butler, Alexander W, Irem Demirci, Umit G Gurun, Yessenia Tellez. 2021. Hand to mouth: Financial stress and food insecurity. *Working Paper*.

Butler, Alexander W, Erik J Mayer, James Weston. 2020. Racial discrimination in the auto loan market. *Working Paper*.

Card, David, Jesse Rothstein. 2007. Racial segregation and the Black–White test score gap. *Journal of Public Economics*, **91** 2158–2184.

Charles, Kerwin Kofi, Erik Hurst. 2002. The transition to home ownership and the black-white wealth gap. *Review of Economics and Statistics*, **84**(2) 281–297.

Chatterji, Aaron K, Robert C Seamans. 2012. Entrepreneurial finance, credit cards, and race. *Journal of Financial Economics*, **106**(1) 182–195.

Chetty, Raj, John N Friedman, Emmanuel Saez, Nicholas Turner, Danny Yagan. 2020a. Income segregation and intergenerational mobility across colleges in the United States. *The Quarterly Journal of Economics*, **135**(3) 1567–1633.

Chetty, Raj, Nathaniel Hendren, Maggie R Jones, Sonya R Porter. 2020b. Race and economic opportunity in the United States: An intergenerational perspective. *The Quarterly Journal of Economics*, **135**(2) 711–783.

Chu, Yongqiang. 2019. Banking deregulation and discrimination in mortgage lending. *Working Paper*.

Chu, Yongqiang, Xiaonan Flora Ma, Tim Zhang. 2021. Bank public status and mortgage lending discrimination. *Working Paper*.

Cornaggia, Jess, Kimberly J Cornaggia, Ryan D Israelsen. 2018. Credit ratings and the cost of municipal financing. *The Review of Financial Studies*, **31**(6) 2038–2079.

Cornaggia, Jess, Matthew Gustafson, Ryan D Israelsen, Zihan Ye. 2019. Government spending and local demographics: Evidence from Moody's municipal ratings recalibration. *Working Paper*.

Cunha, Igor, Miguel A Ferreira, Rui Silva. 2019. Do credit rating agencies influence elections? *Working Paper*.

Dahl, Gordon B, Lance Lochner. 2012. The impact of family income on child achievement: Evidence from the earned income tax credit. *American Economic Review*, **102** 1927–56.

Davydiuk, Tetiana, Deeksha Gupta. 2020. Income inequality, debt burden and covid-19. *Working Paper*.

Deckers, Thomas, Armin Falk, Fabian Kosse, Pia R Pinger, Hannah Schildberg-Hörisch. 2021. Socio-economic status and inequalities in children's IQ and economic preferences. *Journal of Political Economy*.

Di Maggio, Marco, Amir Kermani, Benjamin J Keys, Tomasz Piskorski, Rodney Ramcharan, Amit Seru, Vincent Yao. 2017. Interest rate pass-through: Mortgage rates, household consumption, and voluntary deleveraging. *American Economic Review*, **107**(11) 3550–88.

Dougal, Casey, Pengjie Gao, William J Mayew, Christopher A Parsons. 2019. What's in a (school) name? racial discrimination in higher education bond markets. *Journal of Financial Economics*, **134**(3) 570–590.

Fuster, Andreas, Paul Goldsmith-Pinkham, Tarun Ramadorai, Ansgar Walther. 2020. Predictably unequal? The effects of machine learning on credit markets. *Journal of Finance*.

Gerardi, Kristopher, Lauren Lambie-Hanson, Paul Willen. 2021. Racial differences in mortgage refinancing, distress, and housing wealth accumulation during covid-19. *Working Paper*.

Gilpin, Gregory, Ezra Karger, Peter Nencka. 2021. The returns to public library investment. *Working Paper*.

Giulietti, Corrado, Mirco Tonin, Michael Vlassopoulos. 2019. Racial discrimination in local public services: A field experiment in the united states. *Journal of the European Economic Association*, **17**(1) 165–204.

Guiso, Luigi, Paola Sapienza, Luigi Zingales. 2004. Does local financial development matter? *The Quarterly Journal of Economics*, **119**(3) 929–969.

Haendler, Charlotte, Rawley Heimer. 2021. The financial restitution gap in consumer finance: Insights from complaints filed with the CFPB. *Working Paper*.

Hanushek, Eric A. 2001. Black-White achievement differences and governmental interventions. *American Economic Review*, **91**(2) 24–28.

Harvey, Campbell R, Yan Liu, Heqing Zhu. 2016. ... and the cross-section of expected returns. *The Review of Financial Studies*, **29**(1) 5–68.

Heath, Davidson, Matthew Ringgenberg, Mehrdad Samadi, Ingrid M Werner. 2020. Reusing natural experiments. *Working Paper*.

Howell, Sabrina T, Theresa Kuchler, David Snitkof, Johannes Stroebel, Jun Wong. 2021. Racial disparities in access to small business credit: Evidence from the paycheck protection program. *Working Paper*.

Hu, Qing, Ross Levine, Chen Lin, Mingzhu Tai. 2020. Finance and children's academic performance. *Working Paper*.

Hussam, Reshmaan, Erin M Kelley, Gregory Lane, Fatima Zahra. 2021. The psychosocial value of employment. *Working Paper*.

Jackson, C Kirabo, Cora Wigger, Heyu Xiong. 2021. Do school spending cuts matter? Evidence from the Great Recession. *American Economic Journal: Economic Policy*.

Jang, Heewon, Sean F Reardon. 2019. States as sites of educational (in)equality: State contexts and the socioeconomic achievement gradient. *AERA Open*, **5**(3).

Levine, Ross, Yona Rubinstein, Alexey Levkov. 2014. Bank deregulation and racial inequality in america. *Critical Finance Review*, **3**(1) 1–48.

Marchand, Joseph, 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.

Moody's. 2010. Recalibration of Moody's U.S. municipal ratings to its global rating scale. New York: Moody's Investor Services.

Munnell, Alicia H, Geoffrey MB Tootell, Lynn E Browne, James McEneaney. 1996. Mortgage lending in boston: Interpreting hmda data. *The American Economic Review* 25–53.

Reardon, Sean F, Andrew D Ho, Benjamin R Shear, Erin M Fahle, Demetra Kalogrides, H Jang, B Chavez. 2021. Stanford Education Data Archive (Version 4.0).

Reardon, Sean F, Demetra Kalogrides, Kenneth Shores. 2019a. The geography of racial/ethnic test score gaps. *American Journal of Sociology*, **124**(4) 1164–1221.

Reardon, Sean F, Ericka Weathers, Erin Fahle, Heewon Jang, Demetra Kalogrides. 2019b. Is separate still unequal? New evidence on school segregation and racial academic achievement gaps. *Working Paper*.

Shores, Kenneth, Matthew P Steinberg. 2019. Schooling during the Great Recession: Patterns of school spending and student achievement using population data. *AERA Open*, **5**(3).

Stefan, Matthias, Felix Holzmeister, Alexander Müllauer, Michael Kirchler. 2018. Ethnical discrimination in Europe: Field evidence from the finance industry. *PloS one*, **13**(1).

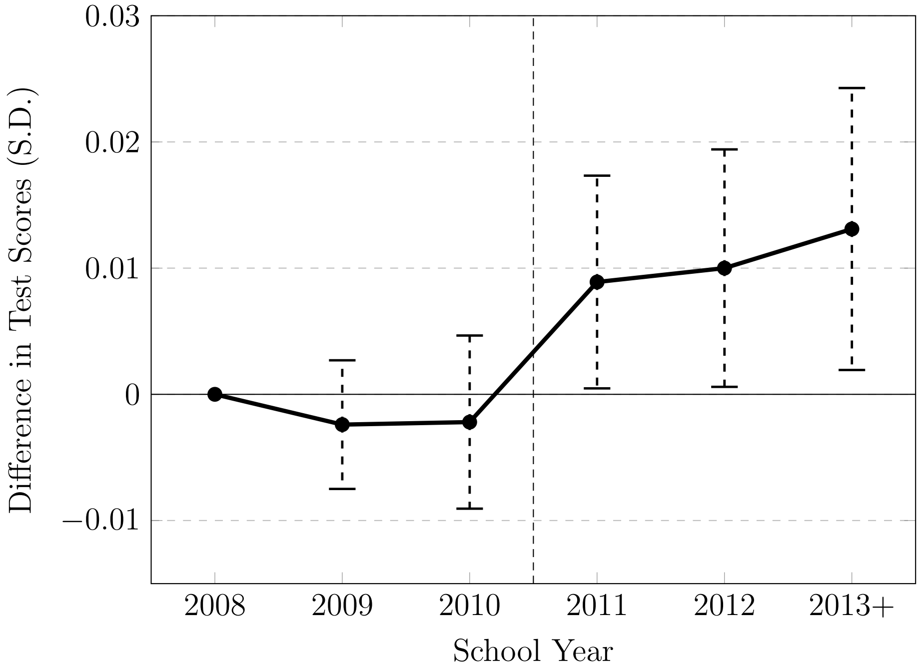
Stein, Luke CD, Constantine Yannelis. 2020. Financial inclusion, human capital, and wealth accumulation: Evidence from the freedman's savings bank. *The Review of Financial Studies*, **33**(11) 5333–5377.

Sylwester, Kevin. 2002. A model of public education and income inequality with a subsistence constraint. *Southern Economic Journal* 144–158.

Torrats-Espinosa, Gerard. 2020. Crime and inequality in academic achievement across school districts in the united states. *Demography*, **57**(1) 123–145.

Yi, Hanyi. 2021. Finance, public goods, and migration. *Working Paper*.

(a) Full Sample



(b) By Race Group

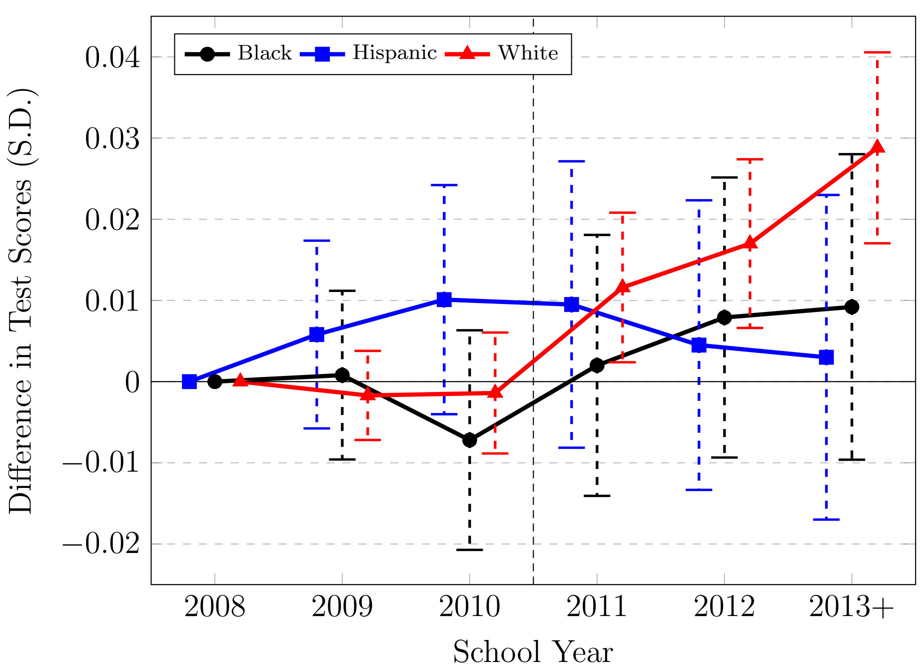


Fig. 1
Effect of Recalibration on Test Scores

This figure plots coefficients of β_j from the following regression of education achievement on the interaction of recalibration in event time. The sample includes counties with full time-series data. County and state-year-grade-subject fixed effects are included. Standard errors are clustered at the county level. Dashed lines represent 95% CI. The point estimates are staggered for ease of reading.

$$Test\ Score_{c,t,g,i} = \sum_j \beta_j(Recal\ Inicator \times Year\ Indicator) + \gamma_c + \gamma_{s\times t\times g\times i} + \epsilon_{c,t,g,i}$$

Table 1
Summary Statistics

Variable	N	Mean	S.D.	P5	P95
Academic Achievement	(County-Year-Grade-Subject level observations)				
Number of Tested Students	272,290	1,388	4,056.08	120	5,585
All Students	272,290	-0.04	0.28	-0.52	0.39
White Students	262,845	0.11	0.25	-0.30	0.53
Black Students	131,159	-0.48	0.27	-0.91	-0.03
Hispanic Students	139,241	-0.28	0.26	-0.69	0.17
White - Black Gap	127,554	0.62	0.25	0.22	1.05
White - Hispanic Gap	137,126	0.45	0.26	0.05	0.89
Non-poor - Poor Gap	255,332	0.55	0.20	0.24	0.90
Socioeconomic Status	(County-Year level observations)				
White Household	29,092	0.09	0.56	-0.83	0.97
Black Household	22,732	-2.32	0.89	-3.74	-0.65
Hispanic Household	26,570	-1.17	0.63	-2.15	-0.09
White-Black Gap	22,661	2.44	0.67	1.26	3.59
White-Hispanic Gap	26,540	1.35	0.44	0.64	2.04
Unemployment	(County-Year level observations)				
White Household	29,092	0.07	0.02	0.03	0.11
Black Household	22,732	0.13	0.05	0.04	0.21
Hispanic Household	26,570	0.08	0.03	0.02	0.13
SNAP Rate	(County-Year level observations)				
White Household	29,092	0.10	0.05	0.03	0.18
Black Household	22,732	0.27	0.08	0.12	0.39
Hispanic Household	26,570	0.18	0.07	0.07	0.29
Moody's Recalibration	(County level observations)				
Recal. Indicator	3,103	0.31	0.46	0.00	1.00
Recal. Intensity	3,103	0.03	0.08	0.00	0.18
Racial Segregation	(County level observations)				
White-Black Segregation	2,888	0.09	0.13	0.00	0.38
White-Hispanic Segregation	2,936	0.07	0.10	0.00	0.30
Spending per Pupil	(School District-Year level observations)				
Dollar Amount ($)	104,192	10,932	5,848.30	7,336	17,872

Note: This table reports summary statistics for academic achievement, SES, unemployment rate, SNAP rate, Moody's recalibration, and racial segregation. Measures of achievement are in standard deviation units of the national distribution.

Table 2
Effect of Municipal Financing on Academic Achievement

Panel A: Dichotomous Recalibration Measure

	(1) Score **All**	(2) Score **English**	(3) Score **Math**
Post × Recal. Indicator	0.013***	0.014***	0.012***
	(0.003)	(0.003)	(0.004)
County FE	Yes	Yes	Yes
State-Year-Grade-Subject FE	Yes	-	-
State-Year-Grade FE	-	Yes	Yes
R^2	0.804	0.828	0.804
Observations	272,179	140,744	131,396

Panel B: Continuous Recalibration Measure

	(1) Score **All**	(2) Score **English**	(3) Score **Math**
Post × Recal. Intensity	0.067***	0.069***	0.066***
	(0.023)	(0.022)	(0.028)
County FE	Yes	Yes	Yes
State-Year-Grade-Subject FE	Yes	-	-
State-Year-Grade FE	-	Yes	Yes
R^2	0.804	0.828	0.804
Observations	272,179	140,744	131,396

Note: This table reports regression results from estimating equation (1). The unit of observation is a county-year-grade-subject and county and state-year-grade-subject fixed effects are included. Column (1) in both panels use full sample that combines English and Math test scores. Columns (2) and (3) analyze English and Math test scores separately. Standard errors reported in parentheses are clustered at the county level. *** p<0.01, ** p<0.05, * p<0.1

Table 3
Effect of Municipal Financing on Racial Gap in Academic Achievement

Panel A: Dichotomous Recalibration Measure

	(1) Score **White**	(2) Score **Black**	(3) Score **Hispanic**	(4) Gap **W - B**	(5) Gap **W - H**
Post × Recal. Indicator	0.022***	0.010	0.001	0.017***	0.018***
	(0.004)	(0.006)	(0.006)	(0.005)	(0.004)
County FE	Yes	Yes	Yes	Yes	Yes
State-Year-Grade-Subject FE	Yes	Yes	Yes	Yes	Yes
R^2	0.735	0.641	0.631	0.636	0.654
Observations	262,728	130,686	139,029	127,075	136,911

Panel B: Continuous Recalibration Measure

	(1) Score **White**	(2) Score **Black**	(3) Score **Hispanic**	(4) Gap **W - B**	(5) Gap **W - H**
Post × Recal. Intensity	0.112***	0.044	-0.005	0.046**	0.099***
	(0.020)	(0.032)	(0.031)	(0.020)	(0.021)
County FE	Yes	Yes	Yes	Yes	Yes
State-Year-Grade-Subject FE	Yes	Yes	Yes	Yes	Yes
R^2	0.735	0.641	0.631	0.636	0.654
Observations	262,728	130,686	139,029	127,075	136,911

Note: This table reports the changes in academic achievement by race and changes in White-Minority achievement gaps. "W - B"("W - H") stands for "White - Black"("White - Hispanic"). The unit of observation is a county-year-grade-subject and county and state-year-grade-subject fixed effects are included. Standard errors reported in parentheses are clustered at the county level. *** p<0.01, ** p<0.05, * p<0.1

Table 4
Effect of Municipal Financing on Household's Socioeconomic Well-being

Panel A: Dichotomous Recalibration Measure

	(1) SES **White**	(2) SES **Black**	(3) SES **Hispanic**	(4) SES Gap **W - B**	(5) SES Gap **W - H**
Post × Recal. Indicator	0.020***	-0.005	-0.069***	0.037**	0.080***
	(0.006)	(0.019)	(0.017)	(0.017)	(0.014)
County FE	Yes	Yes	Yes	Yes	Yes
State-Year FE	Yes	Yes	Yes	Yes	Yes
R^2	0.959	0.845	0.733	0.797	0.704
Observations	29,078	22,722	26,558	22,651	26,528

Panel B: Continuous Recalibration Measure

	(1) SES **White**	(2) SES **Black**	(3) SES **Hispanic**	(4) SES Gap **W - B**	(5) SES Gap **W - H**
Post × Recal. Intensity	0.105***	-0.075	-0.221*	0.217***	0.322***
	(0.035)	(0.080)	(0.119)	(0.073)	(0.103)
County FE	Yes	Yes	Yes	Yes	Yes
State-Year FE	Yes	Yes	Yes	Yes	Yes
R^2	0.959	0.845	0.733	0.797	0.704
Observations	29,078	22,722	26,558	22,651	26,528

Note: This table reports the change in socioeconomic status (SES) by race. The unit of observation is a county-year and county and state-by-year fixed effects are included. Standard errors reported in parentheses are clustered at the county level. *** p<0.01, ** p<0.05, * p<0.1

Table 5
The Relationship between Changes in SES and Test Score (Gap)

	(1) $\Delta Score_{(2017-2009)}$ **All**	(2) $\Delta Gap_{(2017-2009)}$ **W - B**	(3) $\Delta Gap_{(2017-2009)}$ **W - H**
$\Delta SES_{(All,\,2017-2009)}$	0.071***		
	(0.026)		
$\Delta SES\ Gap_{(W-B,\,2017-2009)}$		0.026**	
		(0.013)	
$\Delta SES\ Gap_{(W-H,\,2017-2009)}$			0.026**
			(0.011)
Grade-Subject FE	Yes	Yes	Yes
R^2	0.014	0.020	0.039
Observations	9,734	5,838	5,987

Note: This table analyzes the relationship between changes in SES and test score (gap) for recalibrated counties between 2009 and 2017. Column (1) examines the relationship between changes in SES for all households in the county and the change in test score for all students. Columns (2) & (3) examine the relationship between changes in White-Minority SES gap and changes in White-Minority test score gap. The unit of observation is a county-grade-subject and grade-subject fixed effects are included. Standard errors reported in parentheses are clustered at the county level. *** p<0.01, ** p<0.05, * p<0.1

Table 6
Effect of Municipal Financing on Unemployment

Panel A: Dichotomous Recalibration Measure

	(1) Unemployment **White**	(2) Unemployment **Black**	(3) Unemployment **Hispanic**
Post × Recal. Indicator	-0.0012***	-0.0010	-0.0007
	(0.0004)	(0.0018)	(0.0011)
County FE	Yes	Yes	Yes
State-by-Year FE	Yes	Yes	Yes
R^2	0.863	0.626	0.629
Observations	29,078	22,722	26,558

Panel B: Continuous Recalibration Measure

	(1) Unemployment **White**	(2) Unemployment **Black**	(3) Unemployment **Hispanic**
Post × Recal. Intensity	-0.0048**	0.0045	0.0025
	(0.0023)	(0.0067)	(0.0058)
County FE	Yes	Yes	Yes
State-by-Year FE	Yes	Yes	Yes
R^2	0.863	0.626	0.629
Observations	29,078	22,722	26,558

Note: This table reports the change in unemployment rate by race. The unit of observation is a county-year. County and state-by-year fixed effects are included. Standard errors reported in parentheses are clustered at the county level. The results in this table are reported to four decimal places in order to avoid reporting 0.000 for the standard error in column(1) of Panel A due to rounding. *** p<0.01, ** p<0.05, * p<0.1

Table 7
Effect of Municipal Financing on SNAP Rate

Panel A: Dichotomous Recalibration Measure

	(1) SNAP Rate **White**	(2) SNAP Rate **Black**	(3) SNAP Rate **Hispanic**
Post × Recal. Indicator	-0.003***	0.003*	0.009***
	(0.001)	(0.002)	(0.002)
County FE	Yes	Yes	Yes
State-by-Year FE	Yes	Yes	Yes
R^2	0.937	0.785	0.699
Observations	29,078	22,722	26,558

Panel B: Continuous Recalibration Measure

	(1) SNAP Rate **White**	(2) SNAP Rate **Black**	(3) SNAP Rate **Hispanic**
Post × Recal. Intensity	-0.015***	0.014	0.029**
	(0.003)	(0.009)	(0.015)
County FE	Yes	Yes	Yes
State-by-Year FE	Yes	Yes	Yes
R^2	0.937	0.785	0.699
Observations	29,078	22,722	26,558

Note: This table reports the change in SNAP rate by race. The unit of observation is a county year. County and state-by-year fixed effects are included. Standard errors reported in parentheses are clustered at the county level. *** p<0.01, ** p<0.05, * p<0.1

Table 8
The (Lack of) Response in School District Spending

	(1) ln(Spending) **All School Dist.**	(2) ln(Spending) **All School Dist.**	(3) ln(Spending) **>95% White**	(4) ln(Spending) **>95% White**
Post × Recal. Indicator	0.000		0.004	
	(0.002)		(0.003)	
Post × Recal. Intensity		-0.012		0.019
		(0.010)		(0.031)
School District FE	Yes	Yes	Yes	Yes
State-Year FE	Yes	Yes	Yes	Yes
R^2	0.938	0.938	0.929	0.929
Observations	103,710	103,710	27,436	27,436

Note: This table studies the change in school district funding as a potential channel for the effect of recalibration on test scores. Columns (1) and (2) use the full universe of school districts. Columns (3) and (4) uses the subsample of districts with more than 95% White students. The unit of observation is a school district-year and school district and state-year fixed effects are included. Standard errors reported in parentheses are clustered at the school district level. *** p<0.01, ** p<0.05, * p<0.1

Table 9
The Role of Racial Segregation

	(1) Score Gap **W - B**	(2) Score Gap **W - H**	(3) Score Gap **W - B**	(4) Score Gap **W - H**
Post × Recal. Indicator	0.004	0.007		
	(0.007)	(0.006)		
Post × Recal. Intensity			0.009	0.050*
			(0.028)	(0.028)
Post × Recal. Indicator × W-B Seg.	0.068***			
	(0.019)			
Post × Recal. Indicator × W-H Seg.		0.073***		
		(0.020)		
Post × Recal. Intensity × W-B Seg.			0.201**	
			(0.097)	
Post × Recal. Intensity × W-H Seg.				0.299***
				(0.116)
County FE	Yes	Yes	Yes	Yes
State-Year-Grade-Subject FE	Yes	Yes	Yes	Yes
R^2	0.637	0.654	0.637	0.654
Observations	127,075	136,897	127,075	136,897

Note: This table reports the role of local racial segregation in the heterogeneous treatment effect on test score racial gaps. Segregation (*Seg.*) is measured at the county level in the 2009 school year. The unit of observation is a county-year-grade-subject and county and state-year-grade-subject fixed effects are included. Standard errors reported in parentheses are clustered at the county level. *** p<0.01, ** p<0.05, * p<0.1

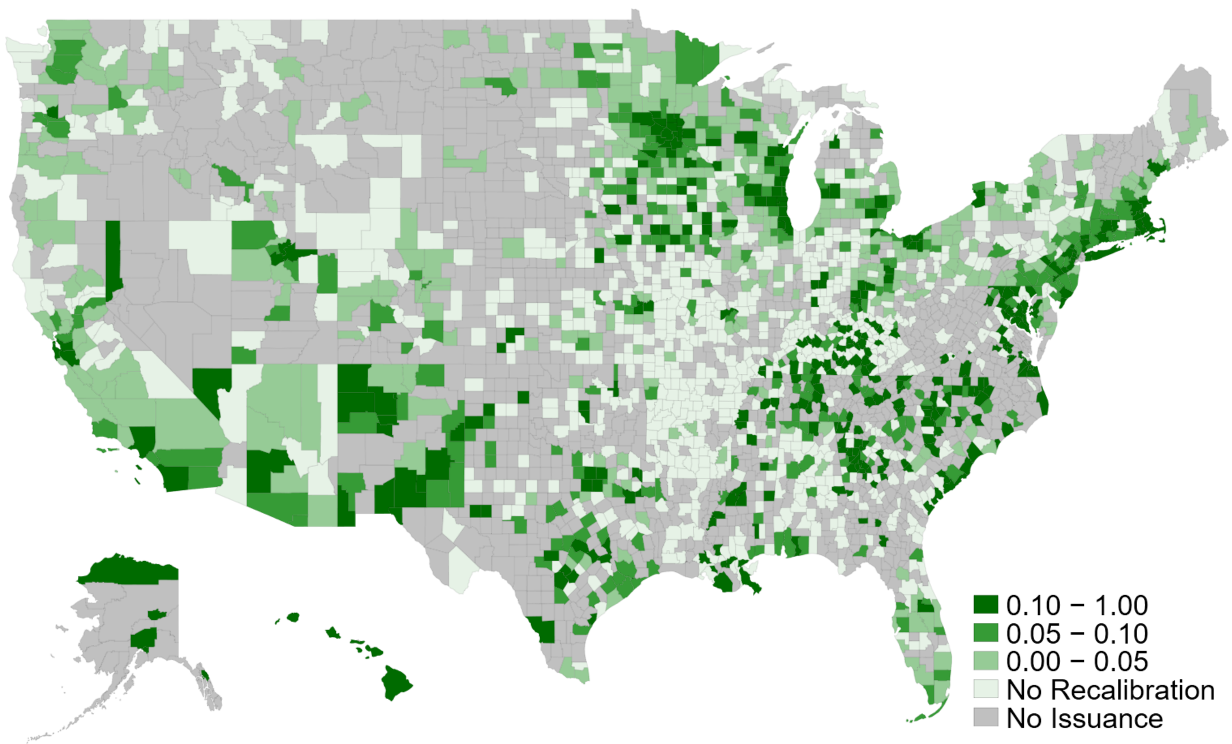
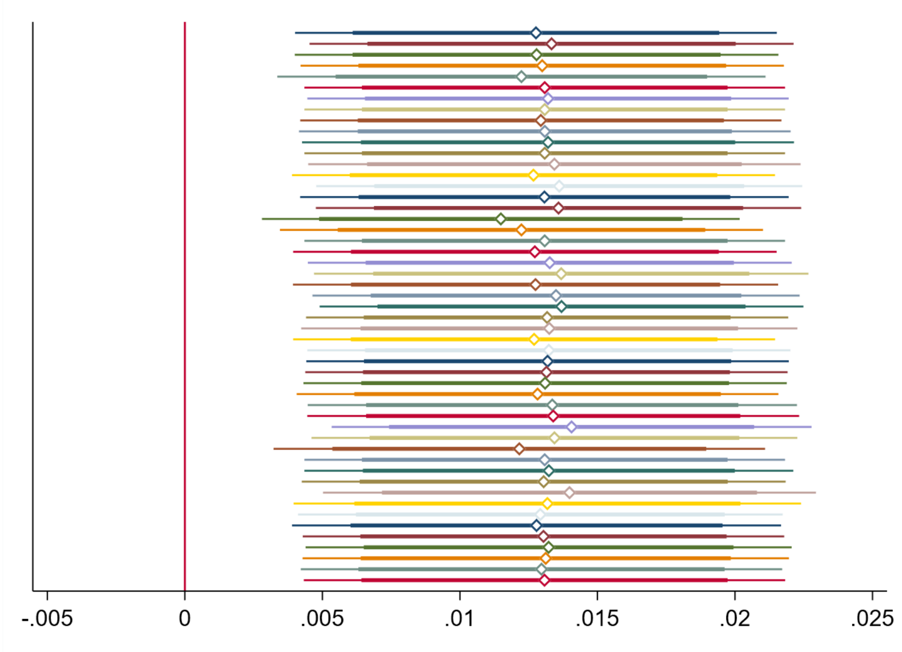


Fig. A.1
Geographic Distribution of Moody's Recalibration

This map demonstrates the geographic distribution of Moody's municipal bond rating recalibration. Variations in color shades represent the fraction of treated local government units in a county. Grey colored counties either do not have local government bonds issued in the three years prior to recalibration or do not have a rating from Moody's.

(a) Dichotomous Recalibration Measure



(b) Continuous Recalibration Measure

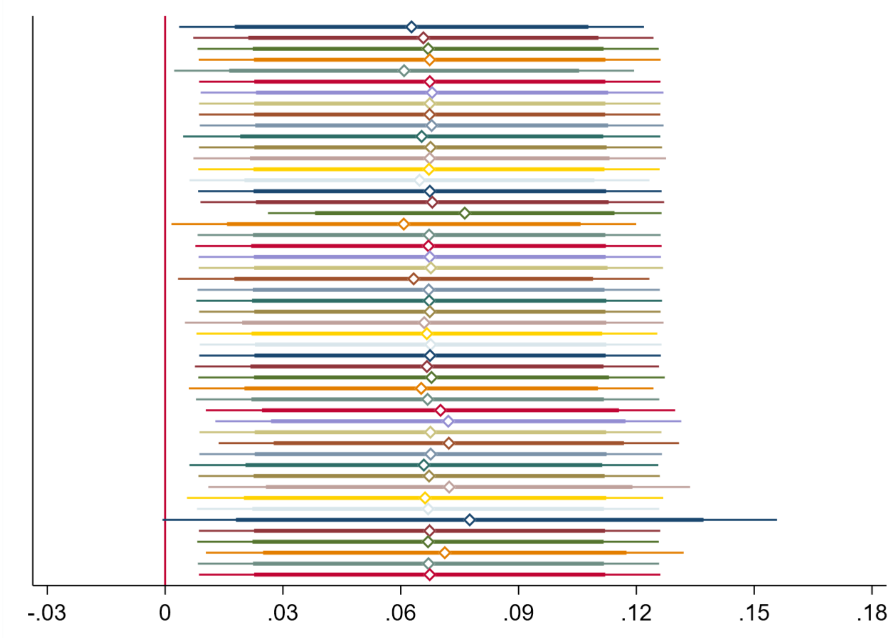
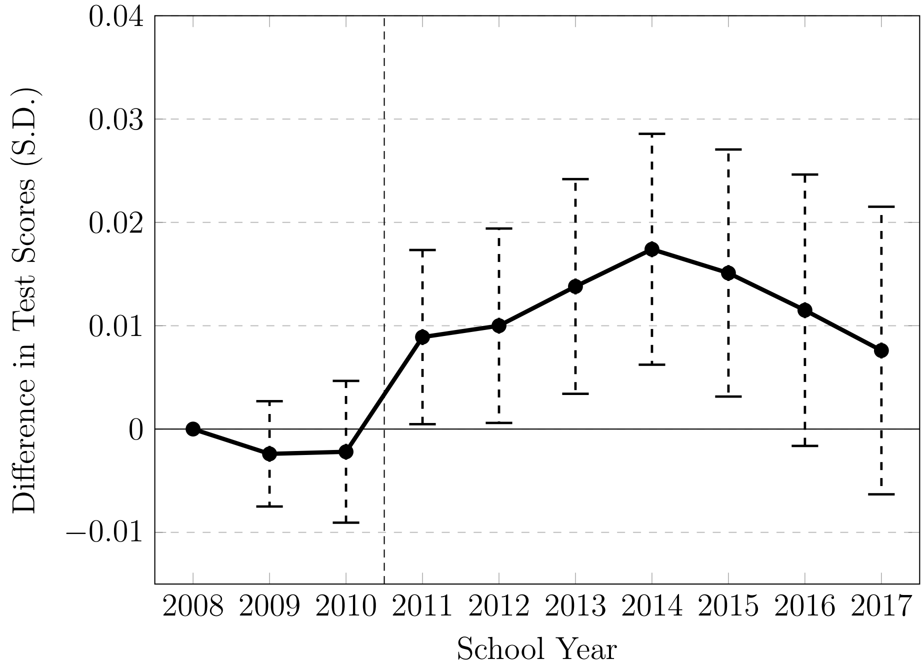


Fig. A.2
Robustness of Result to Removing One State at a Time

This figure plots coefficients of β_j from column (1) of Table 2, but remove one state from the sample at a time. The coefficients are alphabetically ordered based on state name abbreviation and legends are omitted due to space constraints. The Diamonds are point estimates and the thicker (thinner) lines represent the 95% (99%) CIs.

(a) Full Sample



(b) By Race Group

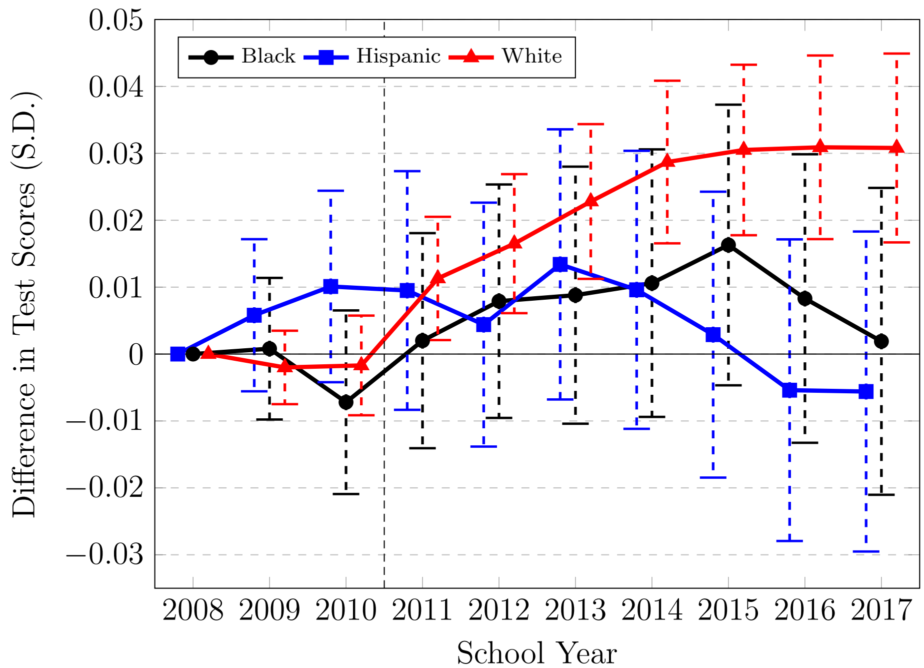


Fig. A.3
Full Time-Series Event-Time Estimation

This figure replicates the estimation from Figure 1 using full time-series indicators instead of aggregating post-2013 effects to the "2013+" indicator. The sample includes counties with full time-series data. County and state-year-grade-subject fixed effects are included. Standard errors are clustered at the county level. Dashed lines represent 95% CI. The point estimates are staggered for ease of reading.

$$Test\ Score_{c,t,g,i} = \sum_j \beta_j (Recal\ Inicator \times Year\ Indicator) + \gamma_c + \gamma_{s\times t\times g\times i} + \epsilon_{c,t,g,i}$$

Internet Appendix Table A.1
Robustness to Excluding the 2010 School Year

	(1) Score **All**	(2) Score **All**	(3) Score **White**	(4) Score **White**	(5) Score **Black**	(6) Score **Black**	(7) Score **Hispanic**	(8) Score **Hispanic**
Post × Recal. Indicator	0.014***		0.023***		0.007		0.003	
	(0.004)		(0.004)		(0.007)		(0.006)	
Post × Recal. Intensity		0.070***		0.114***		0.034		0.013
		(0.025)		(0.023)		(0.034)		(0.036)
County FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
State-Year-Grade-Subject FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
R^2	0.805	0.805	0.736	0.736	0.645	0.645	0.630	0.630
Observations	242,827	242,827	234,590	234,590	116,559	116,559	124,525	124,525

Note: This table replicates results from Table 2 and Table 3 but exclude the 2010 school year. The unit of observation is a county-year-grade-subject and county and state-year-grade-subject fixed effects are included. Standard errors reported in parentheses are clustered at the county level. *** p<0.01, ** p<0.05, * p<0.1

Internet Appendix Table A.2

Robustness to Controlling for Local Demographics

	(1) Score **All**	(2) Score **All**	(3) Score **White**	(4) Score **White**	(5) Score **Black**	(6) Score **Black**	(7) Score **Hispanic**	(8) Score **Hispanic**
Post × Recal. Indicator	0.011***		0.020***		0.009		0.001	
	(0.003)		(0.003)		(0.006)		(0.006)	
Post × Recal. Intensity		0.058**		0.101***		0.038		-0.006
		(0.023)		(0.020)		(0.032)		(0.031)
ln(Median Income) (All)	-0.026	-0.028						
	(0.023)	(0.023)						
Bachelor Education Rate (All)	0.370***	0.375***						
	(0.074)	(0.073)						
Single Mom Rate (All)	-0.042	-0.043						
	(0.065)	(0.065)						
ln(Median Income) (White)			-0.046*	-0.047*				
			(0.024)	(0.024)				
Bachelor Education Rate (White)			0.340***	0.349***				
			(0.071)	(0.072)				
Single Mom Rate (White)			-0.099	-0.097				
			(0.081)	(0.081)				
ln(Median Income) (Black)					0.015	0.015		
					(0.012)	(0.012)		
Bachelor Education Rate (Black)					0.127**	0.128**		
					(0.054)	(0.054)		
Single Mom Rate (Black)					0.002	0.002		
					(0.036)	(0.036)		
ln(Median Income) (Hispanic)							-0.013	-0.013
							(0.011)	(0.011)
Bachelor Education Rate (Hispanic)							-0.000	0.001
							(0.047)	(0.047)
Single Mom Rate (Hispanic)							-0.025	-0.025
							(0.033)	(0.033)
County FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
State-Year-Grade-Subject FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
R^2	0.802	0.802	0.736	0.736	0.641	0.641	0.631	0.631
Observations	271,990	271,990	262,728	262,728	130,686	130,686	139,029	139,029

Note: This table replicates results from Table 2 and Table 3 but adds county-year-race level control variables including the natural logarithm of median household income, the proportion of adults with a bachelor's degree or higher, and the proportion of households with children that are headed by a single mother. The unit of observation is a county-year-grade-subject and county and state-year-grade-subject fixed effects are included. Standard errors reported in parentheses are clustered at the county level. *** p<0.01, ** p<0.05, * p<0.1

Internet Appendix Table A.3
Robustness to Adjusting for Multiple Hypothesis Testing

Variable	Source	t-stat	Order	Holm threshold	BHY threshold
Trading volume	Cornaggia et al. (2018)	11.52	1	0.38%	0.12%
Tax return	Cornaggia et al. (2019)	6.20	2	0.42%	0.24%
Household income	Adelino et al. (2017), Cornaggia et al. (2019)	5.52	3	0.45%	0.36%
Moody's rating	Adelino et al. (2017), Cornaggia et al. (2018)	4.32	4	0.50%	0.48%
Issuance	Adelino et al. (2017), Cornaggia et al. (2018)	3.59	5	0.56%	0.60%
Pvt. employment	Adelino et al. (2017)	3.47	6	0.63%	0.73%
Yield	Adelino et al. (2017), Cornaggia et al. (2018)	3.24	7	0.71%	0.85%
Education outcome	**This paper**	**3.13**	**8**	**0.83%**	**0.97%**
Election outcome	Cunha et al. (2019)	3.07	9	1.25%	1.09%
Gini	Cornaggia et al. (2019)	2.41	10	1.67%	1.21%
AGI net flow	Cornaggia et al. (2019)	2.18	11	2.50%	1.33%
Gov. spending	Adelino et al. (2017)	2.10	12	1.00%	1.45%
Gov. employment	Adelino et al. (2017)	1.70	13	5.00%	1.57%

Note: This table reports results for two adjustments for multiple hypothesis testing. Outcome variables are collected from Adelino et al. (2017), Cornaggia, Cornaggia, and Israelsen (2018), Cornaggia et al. (2019), and Cunha et al. (2019). In cases that an outcome variable is tested in two papers, the larger t-stat is used.

Internet Appendix Table A.4

Decomposition of the Score Racial Gap and the Score Income Gap

Panel A: Test Score White-Black Gap

	(1) Score Gap **W - B**	(2) Score Gap **Income**	(3) Score Gap **W - B**	(4) Score Gap **Income**
Post × Recal. Indicator	0.009** (0.005)	0.010*** (0.004)		
Post × Recal. Intensity			0.032** (0.016)	0.012 (0.020)
Test Score Income Gap	0.450*** (0.010)		0.450*** (0.010)	
Test Score W - B Gap		0.267*** (0.007)		0.268*** (0.007)
β from Table 3	**0.017**		**0.046**	
County FE	Yes	Yes	Yes	Yes
State-Year-Grade-Subject FE	Yes	Yes	Yes	Yes
R^2	0.685	0.686	0.685	0.686
Observations	119,457	119,457	119,457	119,457

Panel B: Test Score White-Hispanic Gap

	(1) Score Gap **W - H**	(2) Score Gap **Income**	(3) Score Gap **W - H**	(4) Score Gap **Income**
Post × Recal. Indicator	0.013*** (0.004)	0.007** (0.003)		
Post × Recal. Intensity			0.073*** (0.019)	0.028 (0.019)
Test Score Income Gap	0.424*** (0.009)		0.424*** (0.009)	
Test Score W - H Gap		0.237*** (0.006)		0.237*** (0.006)
β from Table 3	**0.018**		**0.099**	
County FE	Yes	Yes	Yes	Yes
State-Year-Grade-Subject FE	Yes	Yes	Yes	Yes
R^2	0.688	0.697	0.688	0.697
Observations	130,652	130,652	130,652	130,652

Note: This table reports results for decomposing the DiD coefficient from Table 3 into a race effect and an income effect. "Test Score Income Gap" captures the difference in test scores between non-poor and poor students. "β from Table 3" reports the point estimates from columns (4) & (5) from Table 3. Panel A (B) reports results related to White-Black (White-Hispanic) test score gap. The unit of observation is a county-year-grade-subject and county and state-year-grade-subject fixed effects are included. Standard errors reported in parentheses are clustered at the county level. *** p<0.01, ** p<0.05, * p<0.1

Internet Appendix Table A.5
Robustness to Controlling for Demographic Change

Panel A: Dichotomous Recalibration Measure

	(1) Score **White**	(2) Score **Black**	(3) Score **Hispanic**	(4) Gap **W - B**	(5) Gap **W - H**
Post × Recal. Indicator	0.031***	0.003	-0.003	0.033***	0.023**
	(0.008)	(0.011)	(0.012)	(0.012)	(0.010)
County FE	Yes	Yes	Yes	Yes	Yes
State-Year-Grade-Subject FE	Yes	Yes	Yes	Yes	Yes
R^2	0.741	0.618	0.623	0.641	0.702
Observations	80,727	39,576	31,275	37,529	30,564

Panel B: Continuous Recalibration Measure

	(1) Score **White**	(2) Score **Black**	(3) Score **Hispanic**	(4) Gap **W - B**	(5) Gap **W - H**
Post × Recal. Intensity	0.178***	0.041	0.024	0.119**	0.144**
	(0.049)	(0.062)	(0.072)	(0.048)	(0.056)
County FE	Yes	Yes	Yes	Yes	Yes
State-Year-Grade-Subject FE	Yes	Yes	Yes	Yes	Yes
R^2	0.741	0.618	0.623	0.640	0.702
Observations	80,727	39,576	31,275	37,529	30,564

Note: This table reports the changes in academic achievement by race and changes in White-Underrepresented achievement gaps for subsample of counties that do not have increase in fraction of Black or Hispanic population between 2009 and 2011. The unit of observation is a county-year-grade-subject and county and state-year-grade-subject fixed effects are included. Standard errors reported in parentheses are clustered at the county level. *** p<0.01, ** p<0.05, * p<0.1

Internet Appendix Table A.6
Robustness to Controlling for Imprecise Measurement

	(1) Score **White**	(2) Score **White**	(3) Score **White**	(4) Score **White**
Post × Recal. Indicator	0.022***		0.016***	
	(0.005)		(0.004)	
Post × Recal. Intensity		0.071***		0.075***
		(0.022)		(0.023)
County FE	Yes	Yes	Yes	Yes
State-Year-Grade-Subject FE	Yes	Yes	Yes	Yes
R^2	0.732	0.732	0.726	0.726
Observations	127,075	127,075	136,911	136,911

Note: This table reports the change in academic achievement for White students, but only focus on counties with available estimates for Black (Hispanic) students. The unit of observation is a county-year-grade-subject and county and state-year-grade-subject fixed effects are included. Standard errors reported in parentheses are clustered at the county level. *** p<0.01, ** p<0.05, * p<0.1

Internet Appendix Table A.7

School District Level Results on Test Score

Panel A: All School Districts

	(1) Score **All**	(2) Score **All**	(3) Gap **W - B**	(4) Gap **W - H**	(5) Gap **W - B**	(6) Gap **W - H**
Post × Recal. Indicator	0.014***		0.038***	0.024***		
	(0.003)		(0.005)	(0.005)		
Post × Recal. Intensity		0.090***			0.133***	0.120***
		(0.016)			(0.024)	(0.022)
School District FE	Yes	Yes	Yes	Yes	Yes	Yes
State-Year-Grade-Subject FE	Yes	Yes	Yes	Yes	Yes	Yes
R^2	0.738	0.738	0.298	0.317	0.298	0.317
Observations	1,117,508	1,117,508	473,891	524,077	473,891	524,077

Panel B: School Districts with No Increase in Spending

	(1) Score **All**	(2) Score **All**	(3) Gap **W - B**	(4) Gap **W - H**	(5) Gap **W - B**	(6) Gap **W - H**
Post × Recal. Indicator	0.009**		0.039***	0.026***		
	(0.004)		(0.008)	(0.007)		
Post × Recal. Intensity		0.092***			0.115***	0.113***
		(0.022)			(0.031)	(0.029)
School District FE	Yes	Yes	Yes	Yes	Yes	Yes
State-Year-Grade-Subject FE	Yes	Yes	Yes	Yes	Yes	Yes
R^2	0.752	0.752	0.337	0.389	0.337	0.389
Observations	408,676	408,676	184,889	204,513	184,889	204,513

Note: This table reports the test score results at the school district (SD) level. Panel A uses the full universe of school districts. Panel B uses the subsample of districts that do not increase spending per pupil between 2009 and 2011. For both Panels, the first two columns replicate column (1) from Table 2 while columns (3) through (6) replicate columns (4) & (5) from Table 3. The unit of observation is a SD-year-grade-subject and SD and state-year-grade-subject fixed effects are included. Standard errors reported in parentheses are clustered at the county level. *** p<0.01, ** p<0.05, * p<0.1

Internet Appendix Table A.8

Robustness to an Alternative Measure of Racial Segregation

	(1) Score Gap **W - B**	(2) Score Gap **W - H**	(3) Score Gap **W - B**	(4) Score Gap **W - H**
Post × Recal. Indicator	0.002	0.002		
	(0.008)	(0.006)		
Post × Recal. Intensity			-0.009	0.068*
			(0.031)	(0.035)
Post × Recal. Indicator × Segregation	0.107***	0.125***		
	(0.039)	(0.029)		
Post × Recal. Intensity × Segregation			0.352**	0.201
			(0.162)	(0.158)
County FE	Yes	Yes	Yes	Yes
State-Year-Grade-Subject FE	Yes	Yes	Yes	Yes
R^2	0.636	0.654	0.636	0.654
Observations	126,964	136,747	126,964	136,747

Note: This table replicates the test from Table 9 using an alternative measure of racial segregation from Opportunity Insights. The unit of observation is a county-year-grade-subject and county and state-year-grade-subject fixed effects are included. Standard errors reported in parentheses are clustered at the county level. *** p<0.01, ** p<0.05, * p<0.1

Internet Appendix Table A.9
The Effect of Racial Segregation on SES

	(1) SES Gap **W - B**	(2) SES Gap **W - H**	(3) SES Gap **W - B**	(4) SES Gap **W - H**
Post × Recal. Indicator	-0.023	0.039**		
	(0.021)	(0.017)		
Post × Recal. Intensity			-0.099	0.084
			(0.099)	(0.125)
Post × Recal. Indicator × W-B Seg.	0.399***			
	(0.061)			
Post × Recal. Indicator × W-H Seg.		0.358***		
		(0.066)		
Post × Recal. Intensity × W-B Seg.			1.967***	
			(0.347)	
Post × Recal. Intensity × W-H Seg.				2.043***
				(0.453)
County FE	Yes	Yes	Yes	Yes
State-Year FE	Yes	Yes	Yes	Yes
R^2	0.798	0.705	0.798	0.704
Observations	22,651	26,528	22,651	26,528

Note: This table reports the role of local racial segregation in the heterogeneous treatment effect on SES racial gaps. Segregation (*Seg.*) is measured at the county level in the 2009 school year. The unit of observation is a county-year and county and state-year fixed effects are included. Standard errors reported in parentheses are clustered at the county level. *** p<0.01, ** p<0.05, * p<0.1