Who Benefits from Local Bond Elections?
Evidence from California's School Bond Reform

Max B. Norton
Vancouver School of Economics, UBC

JOB MARKET PAPER

Latest version (frequently updated)—click here

January 18, 2025

Abstract

This paper examines how increased local fiscal discretion affects the distribution of public goods and household sorting. I study California's 2001 reform that reduced the approval threshold for local school bonds from two-thirds to 55 percent of votes cast. Using the reform as a natural experiment, I show that affected districts experienced a capital spending increase of a magnitude equivalent to three extra years of capital investment not experienced elsewhere, increasing the total per-pupil value of their facility stock by 28%. This fiscal response is concentrated in districts with high income heterogeneity, leading to greater inequality in school facilities between districts. Despite this between-district divergence, within affected districts the reform increased income heterogeneity, particularly among families with children. These districts experienced a 10 percent increase in low-income families and greater income heterogeneity among households with children, while showing minimal effects on other household types. The main beneficiaries of the increased spending are low-income families with children who gain access to improved facilities, particularly in relatively affluent districts. These findings have direct implications for proposed school district and local government borrowing policies in several states.

JEL codes: H4, H7, I2, R2, R5

Keywords: school facilities, income sorting, real estate, property tax, local public finance, education finance, state and local policy, fiscal capacity, municipal debt

This paper supersedes previous work circulated as *School Bonds, Sorting, and the Distribution of School Capital: Evidence from California.* For insightful feedback, I thank Gorkem Bostanci, Victor Couture, Tom Davidoff, Don Davis, Jonathan Graves, David Green, John F. Helliwell, Kevin Milligan, Sam Norris, Victoria Prowse, M. Marit Rehavi, Paul Schrimpf, Silvia Vannutelli, Katherine Wagner, Guangli Zhang, and seminar participants at WHY, SEA, IIPF, Sauder CUERE, and UBC. I also thank my colleagues at VSE, in particular Pierre-Loup Beauregard, Joshua Catalano, Sam Gyetvay, Pascuel Plotkin, Federico Ricca, Amirhossein Takavoli, Pablo Valenzuela-Casasempere, and Catherine van der List. Victor Couture and Andrii Parkhomenko generously shared code, and EdSource/Education Data Partnership shared essential school bond election data. I gratefully acknowledge financial support from the Lincoln Institute of Land Policy's C. Lowell Harriss Dissertation Fellowship and the Centre for Innovative Data in Economics Research at UBC. All errors are mine alone. Email: max.norton@ubc.ca.

1 Introduction

Local control is an ingrained historical norm in American public education. This norm has long appeared to conflict with equal educational opportunity—a principle regarded as both essential economic policy and a constitutional mandate. Consequently, many reforms of the last 50 years tend toward more centralized provision of education. An outlier from the trend is *school capital investment*, that is, spending on building and renovating classrooms and school buildings, which remains decentralized. In fact, in 2001 California implemented a major reform that reduced the voter approval threshold for local school bonds from two-thirds to 55 percent, effectively increasing local school boards' discretion over capital investment. Reforms similar to this are regularly proposed in other states. This paper asks who benefits—a question that depends on how increased local discretion affects the distribution of capital spending and how household location decisions respond to these distributional changes.

Recent empirical research on school capital investment has mainly emphasized capitalization (effects on house prices)[1] and impacts on academic achievement.[2] Capitalization receives considerable attention because economists have traditionally viewed it as a signal of economic efficiency, based on the idea that households "vote with their feet" by moving to communities where the provision of public goods best matches their preferences (Tiebout, 1956; Oates, 1972). However, in a decentralized setting, focusing solely on efficiency can mask important distributional consequences, as sorting creates stratification across communities and may concentrate resources in wealthier areas. While the theoretical importance of sorting in these dynamics is well-established, the empirical magnitude of sorting responses to capital spending remains less well understood.

This paper addresses the empirical relationship between school facilities investment and household sorting by providing new estimates of sorting responses in the context of California's school bond reform. This reform, Proposition 39, represents the largest change in local school finance discretion since the 1970s. My main empirical strategy takes it as a natural experiment, comparing affected and unaffected districts in an event study framework. I show that the approval threshold reduction had a strong first-stage effect equivalent to three extra years of capital spending at typical levels. This shock to investment is economically large, totaling $31 billion between 2001 and 2017. I present a stylized model of household sorting across school districts, among whose predictions is that this investment shock should particularly attract low-income households with high taste for education. Consistent with this prediction, districts that use the approval threshold reduction to increase investment exhibit a greater than 10 percent increase in the number of low-income house-

[1] Capitalization of facilities spending is generally found to be neutral to positive. See Cellini et al. (2010); Neilson and Zimmerman (2014); Lafortune and Schönholzer (2022); and Biasi et al. (2024).

[2] Neilson and Zimmerman (2014) and Lafortune and Schönholzer (2022) find that building or rebuilding schools increases attendance and test scores. Park et al. (2020) and Biasi et al. (2024) show that air conditioning improves learning. The latter also identifies types of capital investments that do not, such as athletic facilities.

holds with children and a 12 percent increase in income heterogeneity among family households. Meanwhile, little change occurs among non-poverty families, among households without children, and in income heterogeneity among non-family households. Taken together, this evidence suggests that taste for school facility quality is relatively more important than income as a driver of household sorting. The main beneficiaries of increased local bond approvals are low-income families, who gain access to improved facilities through both direct investment in their existing districts and increased opportunities to move to districts with higher-quality facilities.

School districts across the United States rely heavily on long-term borrowing through bonds to finance facilities investments, with bonds accounting for the majority of school capital spending in California and a large share nationally. These bonds require voter approval in local elections that bundle borrowing authorization with a property tax increase for repayment.[3] Districts use bonds to finance classroom construction, building renovation, and other major facilities projects, creating a direct link between resident preferences and capital investment.

To analyze these dynamics through an economic lens, I develop a stylized model in the local public finance tradition, streamlining Epple and Platt (1998) into a form tailored to motivate empirical analysis. Households differing in income and taste for education sort across districts in a metropolitan area, evaluating a tradeoff between school facilities quality and its cost in terms of taxes and housing prices. This setup raises concerns similar to those of Buchanan (1965), who argued that such sorting processes can transform wealthier communities into exclusionary clubs, limiting access to high-quality public goods to those able to afford the associated costs. My model suggests an alternative perspective: it yields predictions that lowering the bond approval threshold will result in the largest fiscal response in districts with the greatest household heterogeneity, while the increased investment will attract households with a high taste for education relative to their income. This framework allows me to examine whether local discretion leads to exclusive investment patterns or a broader distribution of benefits.

To evaluate the model's predictions empirically, I develop a strategy that identifies affected districts and establishes a credible comparison group—a challenge since Proposition 39 was implemented simultaneously across California. My approach leverages the model's prediction that different districts face varying exposure to the reform. The strategy is intuitive: directly affected school districts are those passing bonds with 55-67% support—an outcome only possible under the reform. Comparison districts either pass bonds regardless of the threshold or pass no bonds at all. This naturally maps to the response-type terminology developed by (Imbens and Angrist, 1994), with districts categorized as compliers, always-takers, or never-takers.[4] This design is particularly

[3]The academic literature is divided on whether to refer to these contests as "elections" or the perhaps more precise "referenda." However, the term "bond election" is widely used and understood among voters and policymakers, and "referendum" has a distinct meaning in California election law. For clarity, I use "bond election" throughout this paper.

[4]While my strategy does not employ instrumental variables as is common in literature using these terms, the analogy can be made more formal by thinking of the approval threshold reduction as a statewide assignment to treatment.

valuable since it invites a natural comparison between compliers and non-compliers, with both groups operating in similar environments and facing similar macro-shocks. I estimate event studies comparing compliers to non-compliers, weighted by the inverse probability of being directly affected—estimated from pre-reform characteristics—to identify the reform's effects on fiscal, sorting, housing market, and academic outcomes. Crucially, since the reform affects these outcomes only through its impact on bond approvals, I can interpret the reduced-form policy effects as causal effects of the bond approvals themselves. The results support the model's predictions and reveal that increased local fiscal discretion can promote economic diversity rather than exclusion.

The empirical analysis is enabled by a newly compiled panel of California school districts in the years 1990-2017. The foundation of this panel is district and school data from the Common Core of Data (CCD; National Center for Education Statistics, 2024), covering all California public schools and their parent districts. The CCD provides information on district characteristics, enrollment, student characteristics, teaching staff, and district finances. I merge in data on school bond elections dating back to 1987 provided by EdSource/Education Data Partnership. Since districts split and combine over time, I use the school-level data to trace out the history of district changes during this period. This allows me to generate a balanced panel over time and trace the impacts of a bond passed in one district on its successor. From Census and American Community Survey (ACS) data, I merge in additional Census tract-level data for the years 1990, 2000, and 2008-2017, notably households by income, home-ownership status, and presence of children. Finally, I merge in house price data from Zillow (2023) and Contat and Larson (2021). This comprehensive panel allows me to examine how school facilities and population composition vary across California's school districts, and to track changes in these patterns following Proposition 39.

The panel data illuminates patterns in how Proposition 39 reshaped school capital investment. Cross-district variation in capital spending increased dramatically after 2001, with the coefficient of variation rising from about 20 percent to over 60 percent by 2009. This increased variation reflects the reform's differential impact across districts—districts with greater income heterogeneity were substantially more likely to approve bonds and approved larger amounts. I illustrate the broader distributional implications by plotting concentration curves showing the cumulative share of capital as a function of the cumulative share of students, ordered by district property wealth. These curves reveal that the distribution of school facilities became increasingly unequal from 2000 to 2017. Moreover, by isolating the shift in these curves attributable to complier districts, I show that essentially the entire increase in inequality came through the reform's effects.

This paper provides new evidence related to several strands of economic literature. My empirical estimates are closely related to the active literature on the efficiency and effectiveness of school facility investment (Cellini et al., 2010; Neilson and Zimmerman, 2014; Martorell et al., 2016; Lafortune and Schönholzer, 2022; Biasi et al., 2024). In particular, my empirical strategy offers an alternative means of identifying the causal effects of capital spending on household sorting. Exploiting Proposition 39's approval threshold reduction emphasizes a permanent change in the fiscal capacity of a subset of school districts. This differs in nature from research designs that exploit random variation in the passage of individual bond proposals. In the short run, some districts affected by Proposition 39 may pass more than one bond, yielding a larger first stage. In the long run, the heterogeneity of the increase in local fiscal capacity may produce large divergences between districts, providing a test of Buchanan's club theory and a more compelling basis for households to change their long-term location decision. The large expected first stage also adds interest to my estimates of the effects of facilities spending on academic outcomes and capitalization.

I also contribute to a literature on the underlying causes of disparities in school facilities. This literature encompasses a variety of themes, including fiscal rules (Biasi et al., 2021); competition from charter schools (Cook, 2018; Kofoed and Fawson, 2021); competition from other districts (Balsdon and Brunner, 2005); the agenda-setting power of school boards (Balsdon et al., 2003; Grosz and Milton, 2023); and variation in property wealth (Brunner et al., 2023). The first-stage fiscal effects I document align with those by Grosz and Milton (2023), who study the mechanics of the fiscal response to Proposition 39. My paper provides, to the best of my knowledge, the first estimates of the sorting effects that followed.[5] My distributional analysis adds to the evidence on the relevance of property wealth, while also revealing the importance of school boards as agenda setters. In a striking parallel, my findings complement recent work by Krimmel (2022) in exploring how the loss and gain of local fiscal discretion shapes communities. Krimmel shows that when California localities suffered reduced fiscal discretion in the 1970s, they responded by adopting restrictive land use policies that made their populations less diverse. I find that when districts partially regained fiscal discretion through the easier bond passage threshold, they used this flexibility in ways that increased economic diversity. Together, these results challenge conventional wisdom about local discretion—rather than inevitably leading to exclusion as predicted by club theory, increased local fiscal discretion can potentially promote heterogeneous communities.

More broadly, this paper contributes to an extensive literature on segregation and sorting in schools that spans urban economics, sociology, and education. School segregation by income is widespread in the U.S., with substantial variation across regions (Reardon and Owens, 2014). The composition of a student's peer group has significant impacts: higher-achieving classmates improve academic performance and behaviors (Sacerdote, 2011; Carrell et al., 2018), concentration of disadvantage can amplify negative outcomes (Billings et al., 2014; Damm and Dustmann, 2014), and exposure to higher-income peers is a particularly important determinant of economic mobility (Chetty et al., 2022a,b). Parents actively respond to these peer effects—Caetano and Maheshri (2017) find evidence that parents' preferences for peer composition drive tipping point behavior in school

[5]In an upcoming working paper, Biasi, Lafortune, and Schönholzer take a structural approach to estimating the aggregate effects of supermajority school bond approval thresholds on spending and academic outcomes.

segregation. The mechanisms behind economic stratification are thus relevant to policymakers seeking to improve student outcomes and prepare students for the labor force. A substantial literature examines the various forces shaping school composition. While some studies focus on direct education policies like court-ordered desegregation (Billings et al., 2014; Guryan, 2004) and school choice (Epple and Romano, 2008), others highlight the crucial role of neighborhood factors and amenities (Caetano and Macartney, 2021; Baum-Snow and Lutz, 2011). However, relatively little research explores how school finance policies affect economic integration. A notable exception is Chakrabarti and Roy (2015), who find that equalization of current expenditures in Michigan increased income mixing in low-spending districts. My results complement theirs by showing that capital investment policies can similarly promote economic diversity within districts when facilities improvements are valued primarily by households with children.

The theoretical section of this paper draws on and contributes to the literature on what Kuminoff et al. (2013) classify as "equilibrium sorting models." This subfield highlights the endogenous connection between heterogeneous households, location choices, voting, and variation in the provision of public goods, building on foundations laid by (Samuelson, 1954; Tiebout, 1956; Musgrave, 1959; Oates, 1972; Brueckner, 1979). My approach draws on and relates to the approach of Epple and Platt (1998); Epple and Sieg (1999) and Calabrese et al. (2006).

Roadmap The remainder of the paper proceeds as follows. Section 2 gives a detailed overview of the policy setting. In Section 3, I provide a theoretical framework in the local public finance tradition, using it to make predictions about the reform's effects and to discuss the measurement of key outcomes. Section 4 introduces the data and provides summary statistics along with a descriptive analysis of reform take-up. Section 5 outlines the empirical strategy used to estimate causal effects. In Section 6, I present the main empirical findings and robustness tests. Section 8 concludes.

2 Setting

California's school system is in many ways typical of those in other U.S. states, making it a good setting for studying the provision of school facilities. The state also has a number of distinctive features. Some of these, like its large size, geographic and demographic diversity, and abundance of school bond proposals, contribute to its desirability as an empirical setting. Other distinctions are anomalies that need some discussion in advance of the empirical work. This section surveys the essential policy features relevant to my empirical analysis and the inferences drawn from it.

I proceed in two subsections. The first describes the structure of California's public school system and its similarities to those of other states in the era before Proposition 39. The second discusses Proposition 39, as well as other recent changes to California's school system that might complicate interpretation of the reform's effects.

2.1 California School Provision and Finance, 1971-2000

California is home to nearly 1,000 school districts—independent local governments that operate the school system in a particular geographic area. School districts are created under the authority of the state and are usually independent of other local governments such as cities. In California, there are three major classes of school district: elementary districts, which mainly offer pre-Kindergarten (PK) through 8th grade instruction; high school districts, mainly offering 9th through 12th grade; and unified districts, offering PK through 12th grade. Any given residential address will typically be served by either a unified district or both an elementary district and a high school district. Districts are governed by a locally elected Board of Education (the "school board"), whose members are granted specific fiduciary and oversight responsibilities. They oversee expenditures on instructional items like teacher salaries, support services such as routine building maintenance and operations, and capital spending on renovation, construction, and land purchases.

A key power of school boards in California and most other states is their authority to propose school bonds, a form of municipal debt used widely in the United States to raise school capital funds. By enabling school districts to take long-term loans, bonds finance the majority of school capital spending in California and likely in the country as a whole. Typically, before this borrowing takes place, bonds must be approved by the school district's electorate in a referendum or 'bond election' that bundles the borrowing with a property tax increase to finance its repayment. Districts are able borrow at competitive rates due to a long-standing principle of American federalism that exempts interest income from municipal debt in an investor's home state from income taxation.

Until the 1970s, California school boards had broad powers to determine the level of funding available for all categories of school spending. Boards could set and adjust property tax rates to generate additional revenue for operational costs, and could propose local bond measures to finance capital spending, which could be approved by a simple majority of the district's voters. In California, the state government has historically also issued bonds of its own to fund school construction, resulting in a variable balance of state and local funding for school capital investment over time. In general, California funds about one-third of total school capital spending at the state level, with the remainder coming from local governments. This places it approximately in the middle of the distribution of states. The federal government has very little fiscal involvement in school facilities.[6]

Due to high inequality generated under the traditional degree of local discretion, however, the

[6]Even with the federal government's unusual foray into facilities funding via the American Recovery and Reinvestment Act (2009), federal funds amounted to only 0.2 percent of school capital outlays nationwide in the years 1994-2013 (Filardo, 2016).

California Supreme Court intervened on equal protection grounds through a series of court rulings in 1971, 1976, and 1977.[7] California voters' approval of 1978's Proposition 13 also disrupted the old system by capping local property tax rates at 1% and raising the approval threshold for local bonds from a simple majority to a two-thirds supermajority. These events prompted legislative actions that imposed caps on school district property tax revenues, redistributed these revenues across districts and other local governments, and in general, significantly increased the share of instructional and operational school spending provided directly from the state's general fund.[8] Collectively, the events of the 1970s imposed strict new limits on the fiscal capacity and independence of California's school districts.

The California Supreme Court's 1970s rulings were the first in a wave of school finance equalization mandates that eventually spread to every state. These reforms reduced the degree of inequality in instructional spending across American school districts, and there is some empirical evidence from Michigan that this has reduced the degree of sorting across districts (Chakrabarti and Roy, 2015). Yet reducing instructional spending disparities increases the relative importance of capital spending disparities, which helps explain the increased attention to facilities from education economists. The implications of Proposition 13 for sorting on the basis of facility provision are more ambiguous. Its passage imposed strict limits on the power of school districts to raise taxes. This contributes to holding facilities disparities fixed, and amplifies the exceptions generated by successful bond proposals. But it is not unusual for school districts in other states to face strict limits on their authority to raise taxes. By contrast, Proposition 13 also strictly limits the growth rate of property tax assessments between property transfers–which remains an unusual feature of California law. This adds an incentive for homeowners to stay in their homes even when their district raises the property tax rate in a way that they may not favor. As a result, this rule may slow down the sorting response of existing homeowners. To the extent that this rule shifts the burden of new spending onto movers, it may accelerate sorting among new movers.

In the wake of the changes of the 1970s, the use of school bonds at the local level became rare for several years. In the late 1980s, however, classroom conditions and overcrowding became issues of public concern. Local bonds re-emerged as a strategy to address these problems. From the 1980s through 2000, the use and importance of local school bonds grew steadily. This trend is illustrated in Figure 1. The dashed black and grey series show that from 1987-88 through 1999-2000, the number of passing and failing bond proposals grew substantially every cycle, peaking in 1997-98 with well over 100 passing proposals.

[7] *Serrano v. Priest I*, 5 Cal.3d 584; *Serrano v. Priest II*, 18 Cal.3d 728; *Serrano v. Priest III*, 20 Cal.3d 25.
[8] Cal. Assemb. B. 65 (1977); Cal. Assemb. B. 8 (1979).

2.2 Ongoing Reform and Challenges to Policy Evaluation, 2001-2017

In spite of growth in local bond usage and a series of eight statewide school bonds passed between 1986 and 1998, public concern over classroom overcrowding continued. (Table C1 lists all statewide school bond proposals between 1986 and 2020.) In 2000, California voters faced two ballot initiatives to loosen restrictions on school districts' authority to issue bonds. In the March primary election, the state electorate rejected Proposition 26, which would have reverted the two-thirds supermajority requirement back to a simple majority for school districts. In November, however, voters approved Proposition 39, reducing the supermajority threshold to a more forgiving standard of 55 percent, conditional on certain new requirements regarding oversight, noticing, election timing, and capital funding for charter schools, which were nascent in the state at the time.

The plot of bond proposals (Figure 1) shows the initial take-up of this option in the 2001-02 elections. 55 percent threshold proposals are represented by solid lines and fall into three categories: those approved with over two-thirds support (blue), those approved with between 55 percent and two-thirds support (green), and those that failed (maroon). The continuation of the dashed lines reflects the fact that school districts retained the option to propose bonds under the two-thirds threshold, allowing them to bypass the oversight and charter school requirements of Proposition 39. The empirical strategy in this paper is based on the idea that for districts where bond proposals received 55 percent to two-thirds support (green), the threshold change is an exogenous shock that re-assigns the proposals from failing to passing. I treat this as a natural experiment that quasi-randomly increases capital spending in a subset of California districts.

To develop an empirical strategy in this vein, it is helpful to be aware of three contemporaneous developments in California's school finance system that could pose potential confounders. First, the passage of Proposition 39 coincided with a surge in state support for capital spending. These funds, often distributed as matching grants, could have amplified the heterogeneity of Proposition 39's effects. Second, this period corresponds with the rapid growth of charter schools throughout the state, a development encouraged by both federal and state policy. This includes Proposition 39, which required districts that use 55 percent threshold bonds to provide, for any charter schools in the district with at least 80 enrolled students, "facilities sufficient to accommodate the charter school's students ... reasonably equivalent to the district schools that these students would otherwise attend" (Legislative Analyst's Office, 2000). Third, between 2014 and 2021, California's school finance equalization approach changed again with the phase-in of the Local Control Funding Formula (LCFF). The LCFF granted districts greater autonomy over how operational funds are spent and restructured state aid to provide more funding to districts with high concentrations of low-income, English language learner, and foster youth students. Hair (2021) provides evidence that the LCFF caused capitalization and population growth in districts receiving higher funding. I address these three developments as potential threats to my empirical strategy in Section 5.

3 Theoretical Framework

How should economists think about the institutional evolution recapped in the preceding section? There exists a long tradition of local public finance models, stretching back to Tiebout (1956), that characterize the equilibrium co-determination of local public good provision and households' residential location decisions. A fundamental insight from this tradition is that the capitalization of local public spending into house prices can be interpreted as a signal of economic efficiency, which is why many empirical studies of school capital spending emphasize capitalization effects. In theory, these effects are driven by sorting, and in empirical practice, estimates are based on the decisions of movers through the use of sales-based house-price indices. Thus, increasing our knowledge about the direct effects of school capital spending on sorting is an open avenue of research.

To provide a conceptual framework for the effects of reducing the approval threshold for school bonds, this section presents a stylized model grounded in the local public finance tradition. This model serves several purposes. First, it provides a structured way to analyze how relaxing this constraint on local discretion shifts equilibrium school capital investment. This leads to specific predictions about the types of districts that will leverage the reform to increase school spending. Additionally, the model identifies mechanisms by which increased spending might influence household sorting. Changes in school funding are likely to have stratifying effects, attracting certain household types to specific districts while driving away others, based on variation across households in income and preference for public goods. Finally, the model informs my empirical strategy by guiding the measurement of key outcomes. It helps to translate abstract concepts about household sorting into measurable constructs that can be analyzed empirically. Thus, the model helps me to bridge the gap between the reform's policy intent and the observable patterns in district composition and public finance outcomes.

Specifically, I develop a streamlined version of the model by Epple and Platt (1998, henceforth EP). The model represents public good provision by three districts constituting a metropolitan area. Households are characterized by their income I and a taste for education parameter α. The provision level of the public good is determined by a vote of resident households, which sort across districts based on their valuation of the good. This sorting drives housing demand and capitalizes public good spending into house prices.

In the equilibrium that emerges, households are stratified across districts by I and α. To better connect this framework to the real-world setting of this paper, I add to EP's analysis of the equilibrium in a few novel ways. First, I demonstrate a straightforward extension to the model accommodating voting thresholds higher than a simple majority, and I derive a prediction concerning the distribution of expenditure changes when the voting threshold changes. *Second, I explore how the voting equilibrium of this model relates to a canonical median voter equilibrium when the public good in* *question is a long-lived stock of capital. Third, I build on EP's illustration of the equilibrium to discuss how stratification within districts can be effectively measured. These extensions link the theoretical framework to practical policy settings by highlighting the dynamics of decision-making under varying voting rules and the implications for how households sort across districts.*

3.1 Setup

The model represents a metropolitan area home to a continuum of households differing in income I and a continuous taste for education parameter $\alpha \in [0, \infty)$, which is not directly observable. Income and taste jointly determine how households value a public good K, i.e. school facilities. Households consume h units of housing at a market rate of p per unit, and pay an ad valorem property tax with rate τ to fund the public good, allocating the rest of their income to a numeraire b. Their preferences are represented by the utility function $U(h, K, b; \alpha, I)$, subject to the budget constraint $I = b + (1 + \tau)ph$.

Within the metropolitan area, let there be three districts, $j = \{A, B, C\}$. District j is characterized by aggregate housing stock H_j, housing unit price p_j, and public good provision at level K_j. The level of K_j is set by a vote of the district's resident households, with the tax rate τ_j adjusting to satisfy the district's government budget constraint $K_j = \tau_j \cdot p_j \cdot H_j$. The vote is governed by an approval threshold d: the support of $100 \times d$ percent of local voters is required to proceed with any given level of spending. Households sort costlessly across districts, taking their incomes with them. They rationally anticipate the effects of public good provision and taxation on their own and other households' location decisions.

3.2 Initial equilibrium

The equilibrium characteristics of this model are characterized by EP for a simple majority approval threshold ($d = 0.5 + \varepsilon$, requiring at least one vote beyond 50%). In equilibrium, households sort themselves across jurisdictions to maximize their indirect utility $V(K_j, p_j; \alpha, I)$; all government budgets are balanced; each jurisdiction's public good-tax bundle (K_j, τ_j) is a voting equilibrium of the residents; and housing prices adjust so that markets clear in each jurisdiction. A key assumption, shown by EP to be sufficient to ensure the existence, uniqueness, and several interesting properties of an equilibrium, is that the indifference curves relating public good spending and net-of-tax housing costs have slopes that decrease with both income and taste for the public good. Holding taste constant, higher-income households are less willing to trade additional housing costs for public spending than lower-income households; holding income constant, higher-taste households are more willing to trade additional housing costs for public spending than lower-taste households. In the equilibrium that follows, districts A, B, and C are ordered by their investment in the public good K_j, and this ordering holds for housing prices p_j; for income conditional on taste, $(I_j|\alpha)$; and for taste conditional on income, $(\alpha_j|I)$. The ordering generates a partition over the space of income and taste, illustrated in Figure 2.[9] Households are sorted across districts by both income and taste, but not entirely by either factor. The taste parameter (α) ensures that districts are not strictly sorted by income alone, resulting in overlap in income distributions across districts. Similarly, income prevents sorting strictly by taste. Each district has a diverse but bounded range of household types, with households at the boundaries indifferent between districts. The population of district B (in households) is obtained by integrating over the area between its boundary with A and its boundary with C.

3.3 Equilibrium effects of a voting rule change

This section introduces several extensions of EP's analysis of the equilibrium, beginning with a straightforward but essential generalization that makes the theory applicable to the real-world dynamics of my empirical setting: I show that the equilibrium properties established by EP for simple majority approval thresholds also hold for supermajority thresholds. Intuitively, in a traditional median-voter model, a single factor (e.g., income) directly determines a voter's ranking over all potential spending policies. The center of the distribution of this underlying factor aligns with the center of the policy preference distribution, making the median voter pivotal in assembling a majority coalition. In contrast, in the present model, a household's ranking over provision-taxation bundles is jointly determined by multiple household characteristics and the local housing price. This ranking is based on whether each bundle improves the household's utility relative to the status quo. It is this ranking that drives voting behavior.[10] Under simple majority rule, the equilibrium provision level is determined by the set of households whose support tips the group of supporters into a majority. These pivotal voters are heterogeneous in their characteristics but share the feature that the proposed provision-taxation bundle is the largest that they would support. Any larger bundle would fail to improve their utility relative to the status quo, leading them to withdraw their support. Appendix A (Subsection A.1) formally extends EP's proof of equilibrium existence and uniqueness to an arbitrary supermajority threshold, showing that these principles apply even when higher consensus is required.

Nest, suppose that a supermajority approval threshold is reduced, such that a different set of households with rank $d' < d$ becomes pivotal. Holding fixed the composition of local households, the partial-equilibrium change in public spending is $\Delta K_j = K^*_{d'} - K^*_d$, where K^*_d is the largest

[9]The figure is based directly on Figure 3 in EP.

[10]Traditional median-voter models typically assume these preferences over bundles are single-peaked with respect to spending levels. This assumption is not required here because the set of feasible provision-taxation bundles is constrained by general equilibrium effects, as demonstrated in EP. These constraints define a "redistribution possibility frontier," within which preferences over the feasible bundles are effectively single-peaked, regardless of the shape of the entire preference distribution.

acceptable provision-taxation bundle for households with rank d. This shift reflects the within-district distribution of preferred spending levels across resident households, which is determined by the local joint distribution of (α, I). This partial-equilibrium effect also alters the attractiveness of districts, prompting households near the previous boundaries in income and taste to re-sort. The model yields a number of predictions about these effects.

Proposition 1. *Public spending response to approval threshold reduction depends on within-district preference heterogeneity.* Districts with greater internal heterogeneity in (α, I) will experience a weakly larger increase in public spending following a reduction in the approval threshold. When the joint distribution of taste and income is more dispersed, the preferences of pivotal voters at higher and lower thresholds differ more significantly. Consequently, the fiscal effects of an approval threshold change are concentrated in more heterogeneous districts. If the fiscal response is non-zero, its magnitude is proportional to the spread in voter preferences over spending. While subsequent re-sorting drives a wedge between the partial-equilibrium and general-equilibrium spending effects of an approval threshold change, households that re-sort come from the tails of the local distribution of spending preferences. General equilibrium effects will not overpower the partial equilibrium effect. Subsection 4.3 tests this proposition empirically.

Next, suppose that district B experiences a relatively large increase in spending ΔK_B due to the approval threshold reduction, in comparison to the fiscal impact on districts A and C.

Proposition 2. *Change in relative spending leads to rotation of indifference boundaries.* A relative spending increase in B changes the equilibrium stratification depicted in Figure 2 by flattening the indifference boundary between districts A and B while steepening the boundary between B and C. Consider a household on the left-hand side of the boundary between A and B. Among households on this indifference boundary, this household is relatively high in taste for education and low in income. It spends relatively little on housing. Consequently, after the change ΔK_B, this household can choose district B and enjoy the increased public expenditure while paying less than a proportional share of the associated tax increase. A household on the right-hand side of this same indifference boundary pays a disproportionately large share of the tax increase while placing relatively little value on education, so it will now choose district A. Thus the boundary "below" a district that experiences increased spending rotates counter-clockwise.

The reverse logic holds along the indifference boundary between districts B and C, which rotates clockwise. A household on the left-hand side of this boundary is relatively income-constrained. When the educational disparity between B and C is reduced, this previously indifferent household will no longer find it worthwhile to pay a premium to reside in C. On the right-hand side of this boundary, households that place relatively little value on education are asked to pay a large cost for increased spending in B. By reducing the discount that B offered over C, they no longer find it worthwhile to remain in B.

Appendix A, Subsection A.2 provides a formal proof of Proposition 2.

These indifference boundaries are not directly observable, so the prediction of boundary rotation is not directly testable. However, it generates two corollary predictions, described below, which I subsequently test in Section 5 and Section 6.

Corollary 2A. *The concentration of high-α, low-income households in B will increase.* The widening of the area between the left-hand sides of the indifference boundaries represents an increase in the population of households with high taste for education and low income. Given a fixed population of households, this is both an increase in level and an increase in the share of this population that chooses district B—a net increase in this group's presence in B.

Corollary 2B. *The within-district income heterogeneity of high-α households in B will increase.* Since the area between the left-hand sides of the indifference boundaries widens, for a fixed range of high values of α, district B will exhibit greater heterogeneity of incomes.

Corollaries 2A and 2B can both be plausibly reversed for the low-α, high-income end side of the household distribution. In this region of the distribution, however, capitalization effects play a much more significant role, whereas low-income households are relatively underexposed to capitalization effects due to their low housing consumption and incomplete pass-through of taxes and capitalization to rental housing. Thus, while the theoretical predictions for effects on low-income households are relatively strong, predictions for effects on high-income households are ambiguous, and I focus my empirical analysis on the former.

3.4 How Stratified is the Equilibrium?

A prerequisite to testing this model's predictions with respect to residential sorting is considering appropriate measures. This subsection considers the various possible measurement concepts for residential sorting in the context of the stylized model just discussed, presenting the intuition and merits of each. As indicated by the formulation of the corollaries just described, in the empirical section I focus on two particular sorting measurement approaches: the *concentration* and *within-district heterogeneity* of particular groups in specific districts.

The equilibrium concept illustrated in Figure 2 shows how, when households are differentiated by multiple dimensions (income, taste for education), the resulting equilibrium sorting is complex and non-degenerate. How can we characterize the degree of stratification, or "sortedness," present in this equilibrium? I outline four categories of measure that capture different aspects of the sorting patterns that arise in the model.

One approach to describing how sorted households are in this equilibrium is to measure the representation of distinct groups within districts. Such measures include the share of a district's population that belongs to a specific group, e.g., the share of households below the poverty line. Closely related are average socioeconomic characteristics of the district. Measures from this category are frequently emphasized in existing literature.[11] These measures can offer insights into the relative attractiveness of a district to a particular demographic group, compared to its attractiveness to other demographic groups. However, it is difficult to resolve certain ambiguities of interpretation. Consider an increase in the low-income share of a district that is already 90% low-income, which suggests increased segregation. In contrast, an increase of the same magnitude in a district that is only 10% low-income implies decreased segregation. Ambiguities like these motivate the need for alternative measures.

A related but distinct concept is the concentration of a group in a given district, such as the share of low-income households in a metropolitan area that reside in a given district. Like representation measures, this family of measures captures the relative attractiveness of a district to a particular group. Unlike representation, though, these measures abstract away from attractiveness to other groups. For example, my empirical estimates strongly suggest that the observed spending increases induced by Proposition 39 were desirable to low-income families with children.

A third category encompasses sorting measures reflecting heterogeneity within districts. Such measures captures the degree of diversity or inequality within a district along a given dimension, offering insights into how mixed or homogeneous the district's population is. In one of the few relevant empirical studies of which I am aware, Chakrabarti and Roy (2015) find that a school finance equalization reform in Michigan increased income heterogeneity in low-spending districts and decreased heterogeneity in high-spending districts. This finding suggests that equalization led to less sorting, in the sense that districts became more similar to each other. I extend their logic by estimating effects on conditional measures of heterogeneity, such as the heterogeneity of income among families with children, where having children is a proxy for high taste for education. This idea is illustrated in Figure 2, where $T_B(\alpha_1)$ represents the district B income dispersion among households with taste α_1. These conditional measures reflect how much diversity a district accommodates for a specific socioeconomic group. High conditional heterogeneity suggests that a district is appealing to a broad range of households in the group, making it similar to concentration in that both reflect the relative desirability of a district for that group.

A final category consists of overall inequality measures that abstract away from individual districts. These measures are valuable because they characterize the overall degree of stratification across a region. However, their downside is that they can only be observed at an aggregate level,

[11] For instance, Cellini et al. (2010) examine the effects of capital spending on the proportion of white or Asian residents and students, and on average income and educational attainment. Similarly, Biasi et al. (2024) estimate the impact of spending on the share of students who qualify for free or reduced-price lunch.

making them less suitable for district-level causal identification strategies.

The empirical sections of this paper focus on testing the effects of Proposition 39 on the concentration and heterogeneity of high-α households in directly impacted school districts. In the model, the structural relationship of interest is a straightforward function $Y_B = f(\Delta K_B - \Delta \overline{K_{-B}})$, where Y is an empirical measure of the concentration of heterogeneity of district B households, and $\Delta K_B - \Delta \overline{K_{-B}}$ represents the increased investment in Proposition 39-impacted districts (ΔK_B) not also experienced by non-impacted districts ($\overline{K_{-B}}$). Regrettably, Y_B, ΔK_B, and $\Delta \overline{K_{-B}}$ are all equilibrium objects, rendering direct estimation of this equation unreliable.

In Section 5 I explain the empirical strategy I use to overcome this endogeneity and estimate a causal relationship between ΔK and Y. First, though, I introduce the data and present some descriptive statistics, including an exploration demonstrating the validity Proposition 1.

4 Data and Descriptive Statistics

To evaluate the dynamic consequences of Proposition 39, I need longitudinal data that relates district-level bond elections and capital spending to district-, school-, and Census tract-level covariates. I assemble a new data set of three linked panels to meet this need. This data set is a panel covering all California school districts from 1990 through 2017. In the next subsection, I summarize the sources and structure of each panel. Following that, I briefly describe the construction of key measures of sorting and how they relate to the theoretical discussion in Section 3. A concluding subsection provides summary statistics and a descriptive analysis of the take-up of Proposition 39.

4.1 Construction of the California School District Panel

4.1.1 District-Level Data

The main data set used in this paper is a panel consisting of all school districts in California from 1990 to 2017. I construct the panel by starting with a comprehensive list of agencies in the Local Education Agency Universe Survey administered by the National Center for Education Statistics (NCES). This survey provides annual data on location, governance structure, type and operational status of schools, employment, enrollment, and grade levels offered. I restrict the list to traditional school districts operating 'regular' schools—schools that offer comprehensive instruction to any student residing in a geographic attendance zone. This restriction excludes irrelevant types of educational agencies, such as county offices of education and transportation authorities, focusing the analysis on locations where the link between residence and school enrollment is meaningful. The resulting panel includes the main school districts across nearly all of California's geography, with the exception of a few areas served exclusively by all-charter school districts.

I merge in additional district-level data from several sources: the Local Education Agency Finance Survey of the NCES, containing data on revenues, debt, and expenditures; the Census Bureau's Small Area Income and Poverty Estimates of total and school-age population by poverty status; and a comprehensive list of local school bond elections in California since 1986, provided to me by EdSource/Education Data Partnership.[12]

I also attach a measure of total capital stock computed using historical data from the Census of Governments. Intuitively, it should be capital stock, not spending, that enters the household utility function. Unfortunately, measures of the current value of facilities are not consistently available for a broad selection of school districts. I construct such a measure by aggregating total capital outlays from the Census of Governments dating back to 1972, depreciated at a rate of 5% per year.[13]

Since there is no centralized data source on property tax rates at the school district level, I estimate the effective rate in each district-year. My estimate is the district's total property tax revenues divided by the aggregate home value in the district reported in Census/ACS data. For years between Census/ACS observations, I use linearly interpolated aggregate home values. This method does not account for commercial real estate, which is taxed at the same rate, so the measure is an upper bound on the effective aggregate rate. However, it does not account for California's limitation on increases in assessed property values. Since this limitation shifts the burden of taxation toward more recent purchasers, the individual tax rate will exceed the effective aggregate rate for some newer purchasers.

This panel also contains measures constructed by aggregating data from more granular levels described below.

4.1.2 School-Level Data

To aggregate data available at the school level up to the district level, I begin by constructing a secondary panel of all California public schools in the years 1990-2017. The comprehensive list of schools comes from the NCES Public Elementary/Secondary School Universe Survey, from which I also use data on enrollment by grade, race/ethnicity, and eligibility for free or reduced-price lunch (a measure of household income). I retain all schools that are (i) operated by a traditional school district, as described above, and (ii) either are regular schools or directly compete with regular schools by providing comprehensive instruction to potentially any student in the area. This restriction retains charter, magnet, alternative, and continuation schools, while excluding specialized schools such as special education schools, schools for the blind or the deaf, schools for incarcerated minors, and adult education programs.

[12] I use the Urban Institute's excellent Education Data API and R package to access several of these data sets: the Local Education Agency Universe Survey; the Small Area Income and Poverty Estimates; plus the Public Elementary/Secondary School Universe Survey and EdFacts, which are discussed in the subsection on school-level data (Urban Institute, 2024).

[13] This measure is similar to the one constructed by Biasi et al. (2024), although they zero out the value of investments after 30 years.

I merge in additional school-level data from two sources: the Department of Education's EdFacts data covering reading and math proficiency by grade level; and project records from the California Office of Public School Construction that report information on project type, funding, and completion status.

Each school is uniquely associated with a school district in the NCES data. Using this relationship, I aggregate the following variables up from school to district level and merge them with the main panel: enrollment by grade by race/ethnicity, enrollment by grade by household income, and reading and math proficiency by grade. Additionally, I use enrollment by free or reduced-price lunch eligibility to calculate within-district measures of student sorting by income, and merge these with the main panel.

4.1.3 Data at Other Levels

To add to the main panel with data provide at Census tract, block group, block, or ZIP code levels, I assemble a second supplementary panel consisting that observes all California Census tracts, using 2010 Census tract definitions, from 1990 through 2017. I retain all tracts served by any regular school. For 1990, 2000, and 2008-2017, Census Bureau data provides numerous measures of local demographics and housing market conditions. 1990 and 2000 data come from the Decennial Census; the source for 2008-2017 data is the American Community Survey (ACS) 5-year sample-based estimates. I treat the ACS estimates as representative of their central year: e.g., 2017 refers to the 2015-2019 ACS sample.[14] The demographic features I observe in this way are: high school completion, households by poverty status and presence of children, mobility across and within metropolitan areas and by home-ownership status, population by race/ethnicity and educational attainment, and distributional data on household income. The housing market features that I observe are: housing units by occupancy status and by occupant ownership, and distributional data on house values and residential rents. I merge in additional tract-level housing data from two sources: Zillow (1996-2017)[15] and the Federal Housing Finance Agency (1990-2017) (Zillow, 2023; Contat and Larson, 2021). I also attach the results of historical presidential, gubernatorial, Congressional, and statewide school bond elections retrieved from the election history database maintained by the state for use in the redistricting process (Statewide Database, 2023).

All of the above data is also aggregated up from tract to district level and merged with the main panel. To relate tracts to school districts, I rely on the NCES School District Geographic Relationship Files for 2017 (National Center for Education Statistics, 2019). This ensures that regardless of variation in district boundaries over time, tract-level outcomes are precisely associated with the

[14] I collect Census and ACS data from IPUMS NHGIS (Manson et al., 2023), using the ipumsr package for retrieval. I convert data to 2010 Census boundaries at the lowest available geographic level using the appropriate crosswalk weights before aggregating up to the 2010 tract level.

[15] Zillow's home value indices are crosswalked from ZIP code to Census tract using code kindly provided by Victor Couture.

relevant district-level bond histories in the final outcome period. Additionally, I use measures of income to calculate within-district measures of household sorting, and merge these with the main panel, as explained in the following subsection.

4.1.4 Measuring Population Composition and Income Heterogeneity

To understand sorting, that is, how households group themselves with like households across districts, we would like to select measure(s) that capture the degree of homophilic grouping versus inequality within and between districts. The class of additively decomposable inequality measures known as generalized entropy offers an elegant answer to this need (Shorrocks, 1980). For every level of observation where I observe an income distribution—that is, for Census tracts and school districts—I calculate three such measures. Specifically, I calculate generalized entropy with an inequality aversion parameter of 0 (a measure also referred to as the mean log deviation); with a parameter of 1 (Theil's T statistic); and parameter 2 (half the squared coefficient of variation).[16] The exact construction of these measures is discussed in Appendix B.

Other than Theil's H, a generalization of Theil's T for categorical data, the generalized entropy indices are relatively little utilized in related literature. Chakrabarti and Roy (2015), for example, relies on similar data as that used in this paper and measures income inequality across the 15 Census income brackets using an inverse Herfindahl index. Since the brackets are nominal and time-invariant, however, the inverse Herfindahl index is susceptible to bias due to top-coding in the upper bracket ($150,000 and above) as nominal incomes grow over time. The generalized entropy indices are not as sensitive to this concern.

4.1.5 Balancing the District Panel

A challenge to constructing a long panel of districts is handling district mergers and splits. Since there is no comprehensive history of these events publicly available, I use the school-level supplementary panel to trace a history of district changes during my study period. As an example, if a school affiliated with district d in 1996 is affiliated with district d' in 1997 and district d no longer exists, I infer that d was merged or split into d'. This history is critical to understanding each district's exposure to bonds: if d passed a bond in 1996, its successor d' is subsequently exposed to potential causal effects of the bond.

To cleanly account for these district histories, I start with the 2017 cross-section of districts, and collapse all predecessors in previous years to a single observation. This approach results in a balanced panel for the years 1996-2017 and attrition of no more than 6 districts in earlier years.

[16]These measures are jointly analyzed as a class by (Shorrocks, 1980). Cowell (2011) includes a thorough review.

4.2 Summary Statistics

What do these data reveal about California's public schools during my study period? Table 1 reports average characteristics of California's nearly 1,000 school districts in 1996, 2000, and 2017. Mean enrollment is approximately stable across time, while the number of schools per thousand pupils grows, driven in part by the growth of the charter school sector. Capital outlays per student in 2017 are 267% of their 1995 rate, while current expenditures, including instructional spending, stand at 181% of their 1996 rate. Growth in current expenditures is driven by balanced growth in federal, state, and local revenues. In contrast, growth in capital outlays is mostly locally financed. As shown by the plot of bond proposals over time (Figure 1), growth in local capital outlays includes both a general pattern of school bond usage and the increase in school bond usage directly caused by Proposition 39. These phenomena are reflected in the table in a 1,550% increase in the cumulative value of approved bonds per student and a 1,440% increase in long-term debt per student from 1996 to 2017.

4.3 Differential Impacts of Reducing the School Bond Approval Threshold: Descriptive Analysis

In order to assess how Proposition 39's reduced approval threshold has been leveraged, this subsection provides a detailed descriptive analysis of the take-up of the reform. First I test my model's prediction that resident heterogeneity predicts the degree to which districts are affected by the threshold reduction. Second, I present 2000 summary statistics broken down by response types to illuminate the typical characteristics of districts directly affected by the reform. Finally, I show how the distribution of school capital stock by property wealth shifts from 2000 to 2017, revealing the distributional impact of the reform.

The model presented in Section 3 yields a prediction that when the voting rule is changed from $d = \frac{2}{3}$ to $d = 0.55$, the change in equilibrium spending levels for district j depends on the difference between the optimal spending levels of the 33rd and 45th percentile voters when ranked by optimal spending levels. Whether this bears any relationship to the actual response to Proposition 39 is a helpful initial plausibility check for the theory. However, the true distribution of preferred spending levels is not observable. As a proxy, I use the interquartile range of the household income distribution.[17] I regress two measures of post-Proposition 39 bond usage on the IQR (measured in standard deviations). The results are presented in Table 2, reported separately for districts that do

[17]There are several reasons why this approximation might fail. First, households can consist of more than one voter. Different voters may have different preferences over personal versus household income. Further, since taste parameterizes a voter's preferred spending level, the voter at rank p in the marginal income distribution may not be at the same rank in the preference distribution. To account for these measurement issues, I run the same tests using the P33-P45 range of household income. Results are substantially the same. I prefer the IQR as a measure because captures a wider slice of the population, allowing for the effects of variable household size, and because it does not pretend to be as close of an approximation to the distribution of preferred spending.

not have a history of passing bonds before 2001 and for those that do, as these groups of districts are hypothetically affected by the reform on different margins. The first dependent variable is an indicator equal to 1 if a school district passes a bond in the first ten years after the reform. This regression represents the extensive margin effect of the income spread. Column (1) shows that among districts without a history of bond passage, a one-standard-deviation increase in the IQR is associated with a 10% increase in the probability of passing a bond after the reform, and an 8% increase among districts with a pre-established history of bond usage. Column (2) shows that controlling for the school board's proposal decisions absorbs most of the explanatory power of the IQR, suggesting that school board are well-informed about the likelihood of a proposal's success. In column (3), the dependent variable is the log of total approved bond value per student in the first ten years after the reform. In both cases, a one standard deviation increase in the IQR of income is associated with significantly greater bond values: a 38 percent increase in value in the case of districts with no history, and a 21 percent increase in value in the case of districts with bond history. These results strongly support the first prediction of the model, and also reveal that school board decisions are a crucial mediating channel.

The differential impact of Proposition 39 across school districts is the variation I would like to use in this paper to estimate the causal effects of facilities spending. This suggests that we divide California districts into an impacted group—those whose fiscal behaviors are directly changed by the approval threshold reduction—and a comparison group consisting of those not directly affected. For the main empirical analysis, I use the observed behavior of districts to define this classification. The impacted group includes any district that passes a bond in the first ten years following the reform, 2001-2010, with 55 percent to two-thirds of the vote. These bond proposals only passed because of the reduced approval threshold. The comparison group consists of two types of districts. The first is districts that simply do not pass any bonds during the 2001-2010 period. Switching the approval threshold between two-thirds and 55 percent is irrelevant for these districts. The second subset of the comparison group is districts whose passing bond(s) in the 2001-2010 period receive greater than two-thirds support. These districts are considered unaffected because their bonds would have passed even if the approval threshold had remained high.[18] By way of analogy to familiar applied microeconomic terminology, the affected group consists of *compliers*—districts that pass bonds because of the intervention represented by Proposition 39. Comparison group districts that do not pass bonds are *never-takers*: they would not have passed bonds in 2001-2010 in either state of the world. Comparison group districts that pass bonds with a high degree of support are *always-takers*: these districts would have increased their capital spending in either state of the world.

As elaborated in Section 5, I estimate the treatment effect of the reform by comparing outcomes

[18] Arguably, districts in this group are spurred to invest in facilities due to competition with nearby districts that fall into the impacted group. If this is the case, my research design only leverages some of the fiscal effect of Proposition 39, and therefore offers a lower bound estimate of the reform's effects.

among the compliers to the non-complier comparison group using an event study framework. Recent work in the difference-in-differences literature, such as Borusyak et al. (2024), highlights the risk of bias in designs where outcomes in affected units are compared to controls drawn from the group of units that are affected later. To avoid such "forbidden comparisons", I follow the example of a series of recent empirical papers and impose a "clean controls" restriction excluding districts that leverage the reform for the first time in the period 2011-2017.[19]

Table 3 reports summary statistics for each of these three groups of districts in 2000, the year that Proposition 39 was adopted, excluding the contaminated controls. Of the remaining districts, 28% are compliers, 18% are always-takers, and 54% are never-takers. Compliers and always-takers are strikingly similar on pre-reform observables. The major difference among the never-takers is the greater representation of rural districts. This group also appears to enjoy substantially greater capital stock, as measured by schools per 1,000 students—a difference likely driven by the diseconomies of small-scale provision of education in rural areas.

How did reducing the school bond approval threshold affect inequality in the distribution of school capital? I illustrate descriptive evidence on this front in two ways. First, in Figure 3, I plot the cross-district coefficient of variation in three categories of expenditure among California school districts from 1994 through 2012. The plotted categories are instructional spending, support services and other current expenditures, and facilities spending. The coefficient of variation (the ratio of the standard deviation to the mean) is used because it provides a measure of variation that is directly comparable across categories. For both instructional and other current expenditures, variance is relatively small—the standard deviation is well below 10 percent of the average spending level. Facilities spending ranges between 10 and 20 percent of average spending in the 1990s, and following the 2001 implementation of Proposition 39 explodes to 40 percent by 2003 and over 60 percent by 2009.

Figure 3 shows that variance in capital spending increased dramatically following Proposition 39, but it does not show us the aggregate effect on inequality in the distribution of school facilities. In principle, if the sudden increase in capital spending was concentrated in previously under-investing areas, this episode could have led to a more equal distribution of capital stock across districts. Alternatively, this surge of spending could have exacerbated existing disparities. To take a closer look at this question, in Figure 4 I plot the 2000 and 2017 concentration curves for cumulative capital outlays (panel (a)) and capital stock value (panel (b)). Specifically, I rank the school districts in my sample by median home value and plot the cumulative share of school capital against the cumulative share of students. Under a perfectly equal distribution of school capital, all districts would fall along the green 45 degree line. Districts that bend the empirical curve below 45

[19] Cengiz et al. (2019); Deshpande and Li (2019) use a clean controls restriction in event studies in the context of studies of the minimum wage and disability programs. Biasi et al. (2024) applies the clean controls restriction in a dynamic regression discontinuity study of school bond elections.

degrees spend less than a proportional share on facilities, while those that bend the curve above 45 degrees spending more than a proportional share. During the decade before the reform, middle-value school districts spend a less on facilities relative to the size of their student populations, which appears directly in the capital stock distribution. In the post-reform period, low-valued and moderately high-valued districts also spend less relative to very high-value districts, shifting the overall distribution further away from equity by 2017.

How much of the increasing deviation from equality from 2001-2017 is attributable to the reduction of the school bond approval threshold? To quantify the contribution of the reform, I employ a Shapley decomposition approach. This method allows me to attribute the overall shift in the global concentration curve to the share from the complier districts and the share from the non-compliers. I begin by smoothing the 2000 empirical capital stock concentration curve by fitting a local polynomial regression. The fitted values from this regression are illustrated in the blue curve in Figure 5. Next, I construct distinct concentration curves for the complier districts alone in 2001 and 2017, and create a new data series represent the change from 2001-2017 across these curves. I fit a second local polynomial regression to this change series. In Figure 5, the dashed black series represents the sum of the fitted values for the global 2001 curve and the fitted values for change among the compliers. Thus, this series represents the change in the global curve through the complier channel alone. Finally, the solid black series plots fitted values for a local polynomial regression of the global 2017 concentration curve. The close alignment of the dashed and solid black curves reveals that essentially all of the increase in inequality from 2001 to 2017 came via the increase in capital spending among the complier districts.

To summarize, Proposition 39 differentially affected the level of capital spending in a meaningfully large subgroup of California school districts. Within-district population heterogeneity is an important determinant of the fiscal effect of reducing the school bond approval threshold. Affected districts are large in population and student enrollment and distributed mainly across urban and suburban areas. In aggregate, the increased expenditure made the distribution of school facilities less equal and more correlated with property values.

5 Empirical Design

The stylized model presented in Section 3 established how school capital spending shapes population sorting. Specifically, for some population composition measurement Y in district B, Y_B is a function of the relative change in capital spending in district B compared to its neighbors, $Y_B = f(\Delta K_B - \Delta \overline{K_{-B}})$. However, both sides of this equation represent equilibrium outcomes. While the descriptive evidence presented above shows that Proposition 39's reduced school bond approval threshold differentially affected spending patterns across districts, translating this variation into credible causal estimates requires careful attention to the endogeneity of capital spending decisions. Districts that invest more in facilities differ systematically in ways that could affect demographic change, and demographic change could also drive the decision to make capital investments to expand school capacity. In this section, I develop an empirical strategy to overcome these challenges by isolating the reform-induced variation in capital investment. This strategy allows me to estimate causal effects on population measures as well as on housing market responses and academic outcomes.

To achieve this, I employ an event study framework that leverages variation in outcomes across school districts that were differentially affected by Proposition 39. In this design, the evolution of outcomes in California districts with low exposure to the reform represents the counterfactual for exposed districts in the absence of the reform. While a variety of research designs with alternative comparison groups are conceivable, low-exposure California districts are especially useful: they operate in the same school finance environment and, crucially, are exposed to the same state and regional population shocks as districts directly exposed to Proposition 39. Furthermore, for sorting outcomes, within-California comparisons align with the conceptual assumption that school facilities primarily affect household location decisions within regions rather than across states—an assumption reflected in my model's focus on a metropolitan area with a fixed population.

My empirical strategy rests on the following premise: districts exposed directly to the reform are those that passed at least one bond with between 55 percent and two-thirds support in the years 2001-2010. Proposition 39 converts these bond measures from failing to passing, providing exogenous variation in available capital funds. The comparison group comprises districts unaffected by the reform—either because their bond support is insufficient for passage under either threshold or because it is sufficient for passage under the original threshold. The response type terminology of compliers, never-takers, and always-takers provides a natural way to categorize these groups, though with the caveat that they play a somewhat different role that in the traditional context in which they are invoked (Imbens and Angrist, 1994).

5.1 Basic Event Study Design

I build up my empirical specification starting from a conventional event study specification. Let Y_{jt} be an outcome of interest in district j, year t. Let C_j represent an indicator function equal to 1 if district j is a complier and 0 otherwise. A basic event study specification has the following form:

$$Y_{jt} = \delta_j + \lambda_t + \sum_{s=1990}^{2017} \beta_s \cdot \mathbb{1}(s = t) \cdot C_j + X_{jt}'\gamma + \varepsilon_{it} \tag{1}$$

In this specification, district fixed effects δ_j control for time-invariant level differences, while year fixed effects λ_t control for time-varying state-level shocks. X_{jt} is a vector of controls, potentially including time-varying components.

The coefficients of interest, β_s, represent the differential change among the compliers in year s relative to an omitted base year t_0. The identifying assumption is that outcomes would have followed parallel trends in the complier and non-complier groups in the absence of the reform. Provided that the assumption holds, $\hat{\beta}_s$ estimates the average treatment effect on the treated (ATT), or the causal effect of the reform on the compliers.

To formalize this estimand and parallel trends assumption, consider the most granular first-stage case where $Y_{jt} = BV_{jt}$ represents the cumulative value of bonds approved by district j. The causal effect of the reform on BV_{jt} is the difference between the year t expected values of BV_{jt} for the complier districts under the two approval thresholds:

$$\mathbb{E}[BV_{jt}|C_j = 1, \text{Threshold} = 55\%] - \mathbb{E}[BV_{jt}|C_j = 1, \text{Threshold} = 2/3] \tag{2}$$

For the event study estimator to provide a consistent estimate of this ATT, the identifying assumption requires that the change in potential outcomes over time in the absence of the reform, $\mathbb{E}[BV_{jt} - BV_{j,t_0}|C_j, \text{Threshold} = 2/3]$ is independent of complier status C_j (de Chaisemartin and D'HaultfŒuille, 2017). A key advantage of the event study design is that I can test the plausibility of this assumption by examining pre-trends (outcomes for $s < t_0$).

5.2 Reduced-Form Policy Effects as the Causal Effect of Bond Approvals

Suppose that for the cumulative bond value outcome BV, the event study estimates reveal parallel pre-trends, suggesting that $\hat{\beta}_s$ is a credible estimate of the causal effect of the reform on bond approvals among the compliers. Further, suppose that $\hat{\beta}_s > 0$. Let $\bar{BV}_s$ be the expected value of bond approvals for compliers in the absence of the reform; thus, $\bar{BV}_s + \hat{\beta}_s$ is the expected value with the reform.

Extending the estimation structure to a downstream outcome Y, we can estimate the reduced-form effect of the policy on Y. An object of more general interest, however, is the causal effect of the new bonds themselves on Y. Under the assumption that the reform affects Y only through its impact on bond approvals BV, we can express this estimand as:

$$\mathbb{E}[Y_{js}|C_j = 1, BV_s = \bar{BV}_s + \hat{\beta}_s] - \mathbb{E}[Y_{js}|C_j = 1, BV_s = \bar{BV}_s] \tag{3}$$

This expression represents the average causal effect of the discrete increase in approved bond value ($\hat{\beta}_s$) on the downstream outcome Y for the compliers. Since all affected districts are compliers and all control districts are non-compliers—with no partial compliance—the event study estimator captures this causal effect without requiring any rescaling via a Wald estimator or related approaches. This is because the reform induces a discrete change in bond approvals among compliers, and the comparison group provides a valid counterfactual for what would have happened in the absence of the reform.

This interpretation of the reduced-form policy effect as the causal effect of new bonds depends on the assumption that the reform influences Y solely through its effect on BV, and not through other channels, such as changes in district policies or administrative practices. This assumption is plausible: Proposition 39's major goal was to lower the approval threshold for school bonds. A potential challenge to this assumption is that Proposition 39 may have also indirectly affected Y by increasing charter school prevalence in complier districts. In Section 6, I address this concern by estimating the policy's effect on charter schools directly and showing the effect to be negligible.

5.3 Challenges to Identification

A number of concerns may be raised about the validity of the identifying assumption for the basic event study design. One subset of these concerns may be addressed simply by appropriate adjustment of the regression through judicious specification of controls. Among these is concern about reverse causality: pre-determined population growth in a district might drive its capital spending decisions. To mitigate this concern, I control for logged linear extrapolations of pre-2001 district-level trends in population and total housing units. A second concern in this category is that the Local Control Funding Formula (LCFF) implemented in 2014 could confound estimates (see Section 2 and Hair, 2021). Fortunately, the LCFF is based on an explicit formula that allocates additional funds based on a measure of low socioeconomic status called the Unduplicated Pupil Percentage (UPP). This formula is linear with a kink where the UPP exceeds 55 percent. I include two controls for the UPP: $UPP/0.55$, trimming values to 1 for districts where the UPP exceeds 0.55, and $(UPP - 0.55)/0.45$, trimming values to 0 for districts where the UPP is less than 0.55. A third concern in this category is that different regions of the state may be exposed to different population shocks that are exogenous to school capital investment. To control for these, I replace the basic specification's year fixed effects with metropolitan area-by-year fixed effects, thus controlling for shocks at the metropolitan level. I define metropolitan areas according to the core-based statistical area definitions developed by the U.S. Census Bureau, including micropolitan statistical areas as metropolitan areas. I treat the 13 counties outside any core-based statistical area as a single metropolitan area.

A different type of concern is that varying geographical nearness to always-taker districts generates spatial spillovers on complier districts that biases the estimates. To control for these spillovers, I define an always-taker exposure measure $\overline{AT}_{Co}$ that represents the enrollment share of always-taker school districts in county Co. I augment the event study specification with an interaction between the complier status and always-taker exposure of district j: $\mathbb{1}(s = t) \cdot C_j \times \overline{AT}_{Co(j)}$. Similar in spirit to interactions used by Cellini (2008), this term controls for the mediating effect of competing school districts when I estimate the causal effect on compliers.

In spite of the inclusion of these controls, the non-complier sample may still exhibit significant differences from the complier sample. To further improve the credibility of my design, I estimate a propensity score model and apply inverse propensity weighting. The design of this model, which is a neural network that estimates a district's probability of being a complier based on its pre-reform observable characteristics, is described in Appendix C. Weighting the event study estimates by the inverse propensity score gives my estimator a 'double robustness' property: the estimator is consistent if either the event study model or the propensity score model is correctly specified, without requiring both.[20]

5.4 Final Event Study Specification

Incorporating the specification refinements just described, I estimate a final event study specification with the following form:

$$Y_{jt} = \delta_j + \lambda_{M\times t} + \sum_{s=1990}^{2017} \left[\beta_s \cdot \mathbb{1}(s=t)\cdot C_j + \delta_j + \zeta \cdot \mathbb{1}(s=t)\cdot C_j \cdot \overline{AT}_{Co(j)}\right] + X'_{jt}\gamma + \varepsilon_{it} \tag{4}$$

To recap, Y_{jt} represents outcome Y in district j, year t. δ_j and $\lambda_{M\times t}$ are district and metropolitan area-by-year fixed effects. C_j is an indicator equal to 1 if j is a complier and 0 otherwise. $\overline{AT}_{Co(j)}$ is the enrollment share of always-taker districts in the same county as district j, $Co(j)$, with the interaction term it enters controlling for spatial spillovers from other high-spending districts. The omitted base year is 2000. The control vector X_{jt} includes controls for the logs of extrapolated pre-period trends in population and housing units and controls for the low-socioeconomic status measure that determines exposure to the 2014 Local Control Funding Formula. Estimation is by weighted least squares with inverse propensity score weights equal to $1/\hat{p}$ for compliers and $1/(1-\hat{p})$ for non-compliers.

For $s > 2000, \beta_s$ represents the average treatment effect of the policy on the treated. For $s < 2000, \beta_s$ provides a test of parallel pre-trends. Furthermore, as argued in Subsection 5.1, under the assumption that the reform only affects outcomes through its effect on capital investment, β_s represents the causal effect of the induced change in bond approvals BV.

[20] Another plausible objection is that my empirical design is biased because I have defined treatment and comparison groups on the basis of outcomes. Recent research on "design-based inference" Rambachan and Roth (2024) find that when defining groups based on outcomes in the presence of selection into treatment, clustering standard errors at the unit level, as I do, provides sufficiently conservative confidence intervals as to completely nullify selection bias. Nevertheless, an alternative approach to estimation would be to define groups based on the propensity score model's predictions of exposure, which are based entirely on pre-period characteristics. When estimated this way, my results are qualitatively similar, but noisier and attenuated due to the presence of classification error. Propensity score weighting cannot be combined with this approach, so the estimates cannot be made doubly robust.

6 Results

6.1 First Stage: Fiscal Impacts

Figure 6 plots event studies estimating the effect of Proposition 39 on school capital. Year-over-year changes are small, but cumulatively, districts that leverage the approval threshold reduction have increased their stock of school capital by $2,535 per student as of 2017. For scale, this change is equivalent to a windfall of 3 years of extra capital spending at average levels. The timing of the effect is consistent with my focus on districts that leverage the change within 10 years.

Several questions may be raised regarding broader fiscal implications. First, have districts completed their new capital expenditures, as suggested by the post-2012 plateau apparent in Figure 6? Figure C2, panel (a) shows that the total value of approved bonds is almost twice the total value of new capital outlays shown in panel (b), suggesting that these districts will likely continue to invest in their facilities.[21] Second, how sustainable is the debt issuance underwriting these investments? Figure C2, panel (c) shows that the increase in total property tax revenues exceeds the increase in long-term debt, suggesting that the mandate is fully funded. Finally, do districts exploit the newly available capital funds to shift other capital funding sources toward instruction or other current expenditures? Figure C2, panel (d) shows no effect on instructional or non-instructional current spending, suggesting not.

6.2 Effects on Population Composition

Next I test the two principal predictions of the theory proposed in Section 3. Recall that the theory predicts that in districts where a change in the voting rule induces a relative increase in public good spending—as just shown to have occurred among the complier districts—two effects are predicted to occur. First, the concentration of low-income households with relatively high taste for public spending will increase. Second, the within-district income heterogeneity of households with high taste for public spending will increase.

6.2.1 Low-Income Households

Since taste for public spending is not directly observable, I use the family status of households as a proxy for relatively high taste for education.[22] In Figure 7, panel (a) plots the event study estimates of the effect of first-stage capital spending on the presence of families with children. The figure plots two effects: families below and above the poverty line. Consistent with the theoretical prediction, in 2017, the number of families below the poverty line has increased by over 10 log points

[21] Local governments commonly take authorized bonds to the market a portion at a time.

[22] The Census defines a family as a group of two or more people related by birth, marriage, or adoption and residing together.

in complier districts compared to non-compliers relative to 2000. The effect on families above the poverty line offers a placebo test since higher-income households are not predicted to sort toward increased investment. The small and insignificant estimates for these households is consistent with this theory. Furthermore, the small and insignificant coefficients in 1990 are reassuring the pre-period trends between the complier and comparison groups are relatively similar.

Panel (b) plots estimated coefficients for effects on all family households compared to all households without children. While not all families with children will have a high taste for education, they likely have a higher taste for education spending than households without children. Estimated effects on families with children are only weakly positive, but it appears unlikely that the total effect on families with children is negative. There is little effect on households without children. These results may suggest that the cost increases associated with increased school spending are too small to have a strong deterrent effect on higher-income households with low taste for education spending.

6.3 Income Heterogeneity

Next, I estimate the reform's effect on the heterogeneity of incomes among family households and on non-family households. Figure 8 illustrates the event study results. Districts that leverage the reform undergo a statistically significant 11.7 log point increase in heterogeneity of incomes among family households, but only a non-significant 2.0 log point increase among non-family households. The small coefficients for 1990 for both cases are reassuring that the estimated effects are not driven by pre-trends.

I interpret the difference between these coefficients as a validity check for the theoretical mechanism. In the theory, the sorting of both families and non-family households can be driven by changes in capital spending and the associated taxation. If families place higher value on capital spending than non-families, then the family households are subject to countervailing forces, attracted by higher spending but not by the price effects of taxation or potential capitalization into house prices. In contrast, the price effects will dominate for non-families. Similar to the results by household type, effects on the high-taste group seem consistent with predictions, while effects on the low-taste group are negligible rather than negative. This again suggests that the cost deterrent effects of increased spending are minimal.

6.3.1 Additional Dimensions of Income Heterogeneity

To further understand the increased income heterogeneity among family households in light of the theoretical framework, I turn to two alternative measures of income heterogeneity, generalized entropy of degree 0 and degree 2 (see subsubsection 4.1.4). These measure heterogeneity across the entire income distribution while emphasizing inequality at the bottom and top, respectively.

The results are presented in Figure 9. I find no evidence of long-run change in heterogeneity at the bottom of the distribution (panel (a)). Panel (b) shows income heterogeneity among families to have increased by 14.0 log points as of 2017, and by 9.4 log points among non-families.

One limitation that is important to note is that my entropy measures are constructed from time-invariant nominal income bins reported by the Census. Secular upward trends in income mechanically increase concentration in the top reported bracket of $150,000 and above, which should have a particular impact on the measure in panel (b) that emphasizes the top of the income distribution. Since the effect on the distribution-neutral measure for non-family households is essentially zero, one interpretation of the effect on the top-end measure is as a lower bound on this mechanical effect. In this light, the higher estimate for family households reaffirms the inference of a positive effect on heterogeneity at the upper end of the distribution.

6.4 House Prices and Property Taxes

The sorting phenomena addressed in the previous section are among the broader effects of school capital investment on demand for the district's educational services. In the education literature, it is common to evaluate this demand by estimating effects on local house prices. These estimates double as a test of efficiency: if education is inadequately provided, willingness to pay for increased spending will exceed the tax cost, and house prices will rise, or vice versa (Brueckner, 1979; Cellini et al., 2010).

I follow this precedent by estimating an event study where the dependent variable is a local house price index (Contat and Larson, 2021). To shed light on the tradeoff between public good benefit and tax price that drives the house price capitalization effect, I also estimate an event study where the dependent variable is an imputation of the local effective property tax rate (see Section 4 for its computation). The estimates from these two regressions are plotted in Figure 10. With respect to house prices, I find a null effect. An initial increase in prices, evident from 2001-2005, is possibly a continuation of a pre-trend, and subsequently falls to zero. With respect to the tax rate, I find that it climbs by 1 mill (0.1%) between 2001 and 2008, and subsequently holds steady. The timing of these effects is consistent with the tax rate offsetting home buyers' valuations of new school facilities in expectation.

6.5 Academic Effects

An important question is whether facilities improvements translate into gains in academic achievement. This is a potential reason that households might value capital spending, but not the only one. For example, capital spending might provide greater athletic opportunities or a healthier environment, which parents might value independently of academic achievement. I test for effects on three academic outcomes: the percentage of students testing proficient in reading and in math, and the dropout rate among the 16-19-year-old population. Figure 11 plots the estimated effects for each of these outcomes as of 2017, all of which are indistinguishable from zero.

Why does increasing capital spending not appear to improve average outcomes? It is worth noting that a highly publicized motivation for Proposition 39 was lowering class sizes. While there is no consensus among researchers that lower class sizes improve academic outcomes, this belief is broadly held among the public. Figure 11 shows exploiting Proposition 39 did not in fact reduce the typical student-teacher ratio in complier districts. To the extent that this is a contributor to academic performance, it may help explain the lack of effect.

7 Discussion

These findings sketch out a picture of how increased fiscal discretion shapes local public good provision and household sorting. The results strongly support the theoretical prediction that high-taste, low-income households are particularly attracted to districts that leverage increased fiscal capacity for facilities investment. The negligible effects on other household types, particularly high-income households with low taste for education, suggest that the deterrent effects of increased local spending are relatively weak. This patterns likely stems from two key mechanisms: First, the limited capitalization of facilities investment into house prices means that new homeowners do not pay a premium for the investment, and existing homeowners do not enjoy a wealth effect that might finance a move. Second, the modest increase in property tax rates (approximately 1 mill or 0.1 percent) implies minimal direct cost increases for households.

The concentration of effects among households with children (the high-α group in the model) suggests that taste for education facilities is more influential in driving sorting decisions than income constraints. This finding advances our understanding of the mechanisms underlying household sorting responses to local public goods. While classic theoretical models like Tiebout (1956) emphasize income-based sorting, and much empirical work focuses on measuring house price capitalization, this paper's results highlight the importance of preference heterogeneity in shaping the distributional consequences of local fiscal policy.

The increase in within-district income heterogeneity among families with children represents an important departure from traditional predictions about local public good provision driving increased homogeneity through sorting. This finding suggests that when facilities investment is valued primarily by households with children, increased spending can actually promote economic integration within communities. The effect appears to be driven by both increased presence of low-income families and greater dispersion at the top of the income distribution, though measurement limitations due to top-coding in Census income data make precise characterization of high-end effects challenging.

These sorting patterns help explain the lack of effects on academic outcomes despite substantial facilities investment. An influx of low-income students may offset any direct benefits of improved facilities on average test scores. This highlights the importance of considering demographic responses when evaluating the impact of school spending programs. It also suggest that traditional measures of school quality may not fully capture the welfare benefits of facilities investment for economically diverse communities.

8 Conclusion

This paper provides new evidence on who benefits when state policy enhances local fiscal capacity for school facilities investment. Using California's 2001 school bond reform as a natural experiment, I show that reducing the voter approval threshold from two-thirds to 55 percent led to a substantial increase in capital investment among affected districts, equivalent to three additional years of typical capital spending. This first-stage effect is concentrated in districts with more heterogeneous populations, consistent with existing research (Grosz and Milton, 2023) and with the theoretical prediction that resident heterogeneity amplifies the impact of lowering the approval threshold.

The primary beneficiaries of this increased investment appear to be low-income families with children who gain access to improved school facilities. Rather than driving economic segregation, the reform promoted income mixing within districts, particularly among households with children. This finding contributes to our understanding of the relationship between local public good provision and economic integration, suggesting that targeted investments valued by specific demographic groups can help create more diverse communities.

These results have direct implications for current policy debates. Several states maintain supermajority requirements for local bond measures, and reforms similar to California's Proposition 39 are regularly proposed. This paper's findings suggest such reforms can effectively increase local investment in public facilities while promoting economic integration rather than exacerbating segregation. However, state matching funds play a crucial role in mediating these effects—a key consideration as jurisdictions beyond school districts explore similar threshold reforms to address infrastructure needs. Understanding how approval thresholds and state transfer policies can be optimized to address cross-jurisdictional externalities while promoting equitable access to public goods remains an important challenge for public finance.

References

Balsdon, E. and Brunner, E. J. (2005). Can competition tame the leviathan? Evidence from California's proposition 39. *National Tax Journal*, 58(4):627–642.

Balsdon, E., Brunner, E. J., and Rueben, K. (2003). Private demands for public capital: evidence from school bond referenda. *Journal of Urban Economics*, 54(3):610–638.

Baum-Snow, N. and Lutz, B. F. (2011). School desegregation, school choice, and changes in residential location patterns by race. *American Economic Review*, 101(7):3019–3046.

Biasi, B., Lafortune, J. M., and Schönholzer, D. (2021). School Capital Expenditure Rules and Distribution. *AEA Papers and Proceedings*, 111:450–454.

Biasi, B., Lafortune, J. M., and Schönholzer, D. (2024). What works and for whom? effectiveness and efficiency of school capital investments across the u.s. Working Paper 32040, NBER.

Billings, S. B., Deming, D. J., and Rockoff, J. (2014). School segregation, educational attainment, and crime: Evidence form the end of busing in charlotte-mecklenburg. *The Quarterly Journal of Economics*, 129(1):435–475.

Borusyak, K., Jaravel, X., and Spiess, J. (2024). Revisiting event-study designs: Robust and efficient estimation. *The Review of Economic Studies*, page rdae007.

Brueckner, J. K. (1979). Property values, local public expenditure and economic efficiency. *Journal of Public Economics*, 11(2):223–245.

Brunner, E. J., Schwegman, D., and Vincent, J. M. (2023). How much does public school facility funding depend on property wealth? *Education Finance and Policy*, 18(1):25–51.

Buchanan, J. M. (1965). An economic theory of clubs. *Economica*, 32(125):1–14.

Caetano, G. and Macartney, H. (2021). What determines school segregation? the crucial role of neighborhood factors. *Journal of Public Economics*, 193:104335.

Caetano, G. and Maheshri, V. (2017). School segregation and the identification of tipping behavior. *Journal of Public Economics*, 148:115–135.

Calabrese, S., Epple, D., Romer, T., and Sieg, H. (2006). Local public good provision: Voting, peer effects, and mobility. *Journal of Public Economics*, 90(6-7):959–981.

Carrell, S. E., Hoekstra, M., and Kuka, E. (2018). The long-run effects of disruptive peers. *American Economic Review*, 108(11):3377–3415.

Cellini, S. R. (2008). Causal inference and omitted variable bias in financial aid research: Assessing solutions. *The Review of Higher Education*, 31(3):329–354.

Cellini, S. R., Ferreira, F., and Rothstein, J. (2010). The Value of School Facility Investments: Evidence from a Dynamic Regression Discontinuity Design. *The Quarterly Journal of Economics*, 125(1):215–261.

Cengiz, D., Dube, A., Lindner, A., and Zipperer, B. (2019). The effect of minimum wages on low-wage jobs. *The Quarterly Journal of Economics*, 134(3):1405–1454.

Chakrabarti, R. and Roy, J. (2015). Housing markets and residential segregation: Impacts of the michigan school finance reform on inter- and intra-district sorting. *Journal of Public Economics*, 122:110–132.

Chetty, R., Jackson, M. O., Kuchler, T., Stroebel, J., Hendren, N., Fluegge, R. B., Gong, S., Gonzalez, F., Grondin, A., Jacob, M., Johnston, D., Koenen, M., Laguna-Muggenburg, E., Mudekereza, F., Rutter, T., Thor, N., Townsend, W., Zhang, R., Bailey, M., Barberá, P., Bhola, M., and Wernerfelt, N. (2022a). Social capital i: Measurement and associations with economic mobility. *Nature*, 608:108–121.

Chetty, R., Jackson, M. O., Kuchler, T., Stroebel, J., Hendren, N., Fluegge, R. B., Gong, S., Gonzalez, F., Grondin, A., Jacob, M., Johnston, D., Koenen, M., Laguna-Muggenburg, E., Mudekereza, F., Rutter, T., Thor, N., Townsend, W., ZHang, R., Bailey, M., Barberá, P., Bhola, M., and Wernerfelt, N. (2022b). Social capital ii: Determinants of economic connectedness. *Nature*, 608:122–134.

Contat, J. and Larson, W. D. (2021). A flexible method of house price index construction using repeat-sales aggregates. Technical Report 21-01, Federal Housing Finance Agency.

Cook, J. (2018). The effect of charter competition on unionized district revenues and resource allocation. *Journal of Public Economics*, 158:48–62.

Cowell, F. A. (2011). *Measuring Inequality*. Oxford University Press, Oxford, UK.

Damm, A. P. and Dustmann, C. (2014). Does growing up in a high crime neighborhood affect youth criminal behavior? *American Economic Review*, 104(6):1806–1832.

de Chaisemartin, C. and D'HaultfŒuille, X. (2017). Fuzzy differences-in-differences. *The Review of Economic Studies*, 85(2):999–1028.

Deshpande, M. and Li, Y. (2019). Who is screened out? application costs and the targeting of disability programs. *American Economic Journal: Economic Policy*, 11(4):213–48.

Epple, D. and Platt, G. J. (1998). Equilibrium and local redistribution in an urban economy when households differ in both preferences and incomes. *Journal of Urban Economics*, 43(1):24–51.

Epple, D. and Romano, R. (2008). Educational vouchers and cream skimming. *International Economic Review*, 49(4):1395–1435.

Epple, D. and Sieg, H. (1999). Estimating Equilibrium Models of Local Jurisdictions. *Journal of Political Economy*, 107(4):645–681.

Filardo, M. (2016). State of our schools: America's k-12 facilities 2016. Technical report, 21st Century School Fund, Washington, DC.

Grosz, M. and Milton, R. T. (2023). Relaxing electoral constraints in local eduation funding. Technical Report EdWorkingPaper: 22 -524, Annenberg Institute at Brown University.

Guryan, J. (2004). Desegregation and black dropout rates. *American Economic Review*, 94(4):919–943.

Hair, C. (2021). The local effects of spatially targeted public policies: Evidence from california school finance reform. Job market paper.

Imbens, G. W. and Angrist, J. D. (1994). Identification and estimation of local average treatment effects. *Econometrica*, 62(2):467–475.

Kofoed, M. S. and Fawson, C. (2021). A neighborly welcome? charter school entrance and public school competition on the capital margin. *Public Choice*, 188:75–94.

Krimmel, J. (2022). Reclaiming local control: School finance reforms and housing restrictions. Working paper.

Kuminoff, N. V., Smith, V. K., and Timmins, C. (2013). The new economics of equilibrium sorting and policy evaluation using housing markets. *Journal of Economic Literature*, 51(4):1007–1062.

Lafortune, J. M. and Schönholzer, D. (2022). The impact of school facility investments on students and homeowners: Evidence from los angeles. *American Economic Journal: Applied Economics*, 14(3):254–289.

Legislative Analyst's Office (2000). Proposition 39: Analysis by the legislative analyst. In *Voter Information Guide for 2000, General Election*, pages 38–39. Secretary of State of the State of California, Sacramento, CA.

Manson, S., Schroeder, J., Riper, D. V., Kugler, T., and Ruggles, S. (2023). Ipums national historical geographic information system: Version 18.0 [dataset]. Technical report, IPUMS, Minneapolis, MN.

Martorell, P., Stange, K., and Jr., I. M. (2016). Investing in schools: Capital spending, facility conditions and student achievement. *Journal of Public Economics*, 140:13–29.

Musgrave, R. A. (1959). *The Theory of Public Finance: A Study in Public Economy*. McGraw-Hill, New York.

National Center for Education Statistics (2019). Education demographic and geographic estimates program (edge): School district geographic relationship files.

National Center for Education Statistics (2024). Local eduation agency (school district) universe survey. Available at https://nces.ed.gov/ccd/pubagency.asp.

Neilson, C. A. and Zimmerman, S. D. (2014). The effect of school construction on test scores, school enrollment, and home prices. *Journal of Public Economics*, 120:18–31.

Oates, W. E. (1972). *Fiscal Federalism*. Harcourt Brace Jovanovich, New York.

Park, R. J., Goodman, J., Hurwitz, M., and Smith, J. (2020). Heat and learning. *American Economic Journal: Economic Policy*, 12(2):306–39.

Rambachan, A. and Roth, J. (2024). Design-based uncertainty for quasi-experiments. Working paper.

Reardon, S. F. and Owens, A. (2014). 60 years after *Brown*: Trends and consequences of school segregation. *Annual Review of Sociology*, 40:199–218.

Sacerdote, B. (2011). *Peer Effects in Education: How Might They Work, How Big Are They and How Much Do We Know Thus Far?*, volume 3, chapter 4, pages 249–277. Elsevier, Amsterdam.

Samuelson, P. A. (1954). The pure theory of public expenditure. *The Review of Economics and Statistics*, 36(4):387–389.

Shorrocks, A. F. (1980). The class of additively decomposable inequality measures. *Econometrica*, 48(3):613–625.

Statewide Database (2023). Election data. https://statewidedatabase.org/.

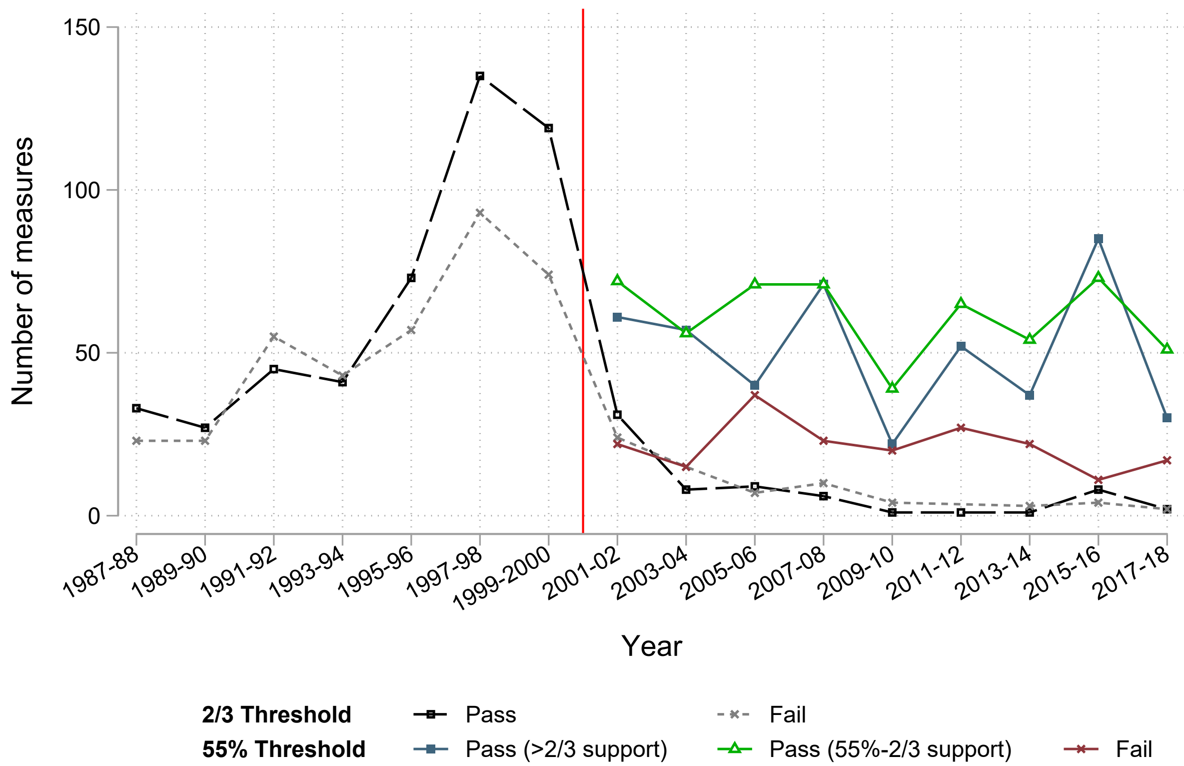
Tiebout, C. M. (1956). A pure theory of local expenditures. *Journal of Political Economy*, 64(5):416–424.

Urban Institute (2024). Education data portal (version 0.22.0). https://educationdata.urban.org/documentation/, made available under the ODC Attribution License. Accessed: 2024-08-26.

Zillow (2023). Zillow housing data. Available at https://www.zillow.com/research/data/.

Figure 1: Local school bond proposals over time

(a) Proposals by approval threshold and outcome, 1987-2018

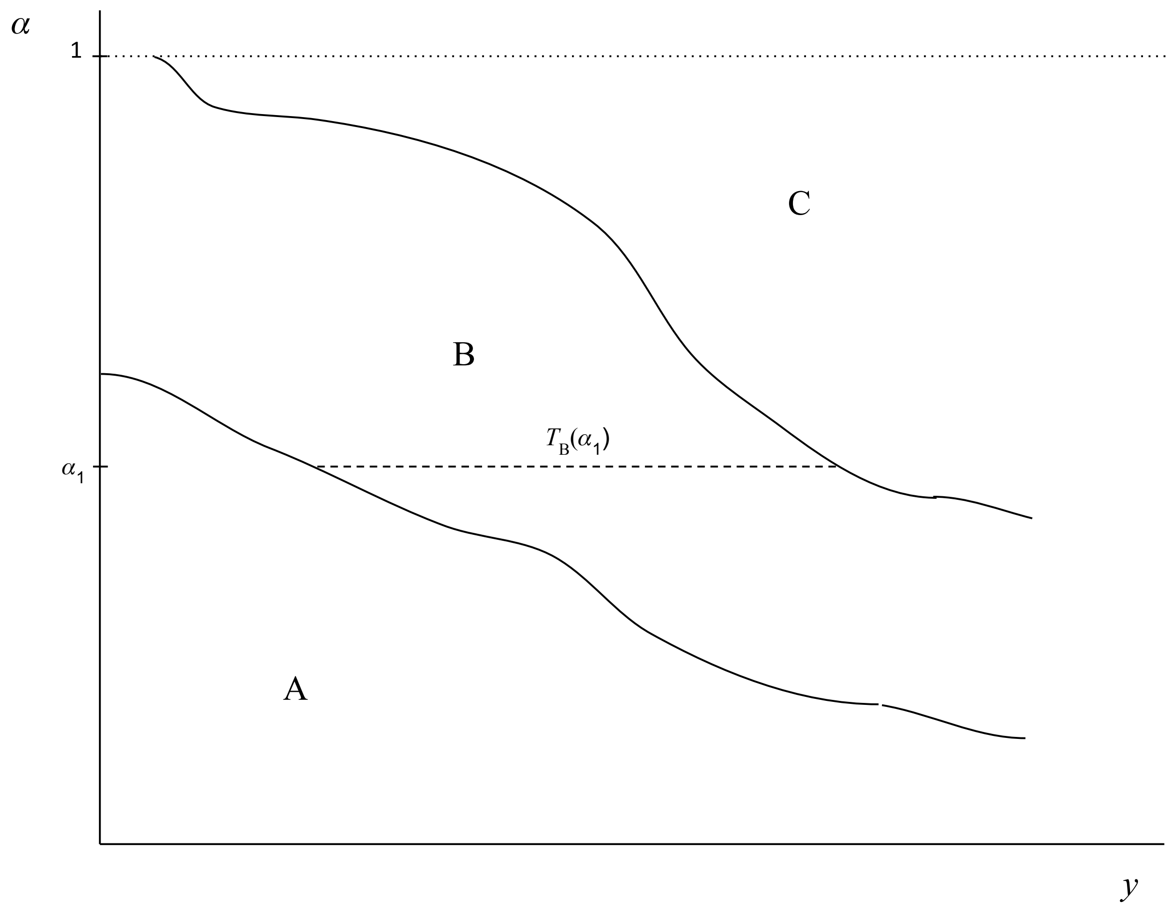


(b) Descriptive statistics for bond proposals, by time period and approval threshold

	1986–2001 Threshold $= \frac{2}{3}$	2002–2019 Threshold $= \frac{2}{3}$	2002–2019 Threshold $= 55\%$
Passage rate (%)	56.6	43.7	81.9
Voter support (%)	67.1	64.4	63.2
	[10.5]	[13.3]	[8.7]
Bond size (2001 $)	37,630,659	60,830,335	82,301,995
	[102,253,995]	[116,232,323]	[144,149,454]
Number of proposals	875	103	1,238

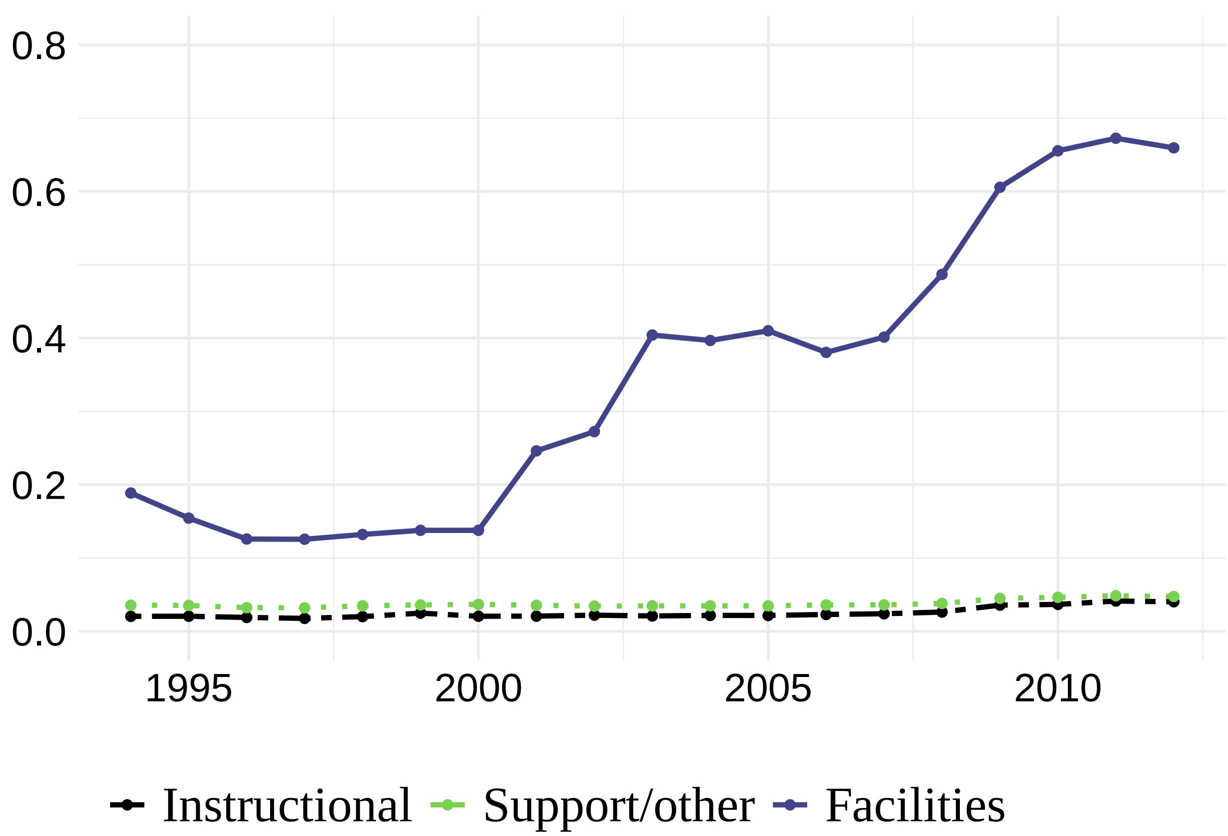
Notes: Panel (a) illustrates the number of passing and failing bond proposals in each 2-year interval from 1987-88 through 2017-18. Two-year intervals are used to smooth cyclical fluctuation between even- and odd-numbered years. Dashed lines pertain to proposals that require $\frac{2}{3}$ voter support for approval, and solid lines pertain to proposals that require 55% voter support. Panel (b) reports the passage rate, mean share of yes votes, and mean bond size for $\frac{2}{3}$-threshold bonds before Proposition 39 and for $\frac{2}{3}$-threshold and 55%-threshold bonds after Proposition 39. Bond size is reported in 2001 dollars. Standard deviations for voter support and bond size are reported in square brackets. Both panels are calculated from the set of all bond elections between 1986 and 2019, without restriction to districts observed in the main sample.

Figure 2: Theoretical illustration: Equilibrium stratification



Notes: An equilibrium of the model described in Section 3. Households are stratified across the three districts A, B, and C by their income y and taste for the public good α. Housing demand in B is given by the integral between the boundaries of B with A and C. $T_B(\alpha_1)$ represents the dispersion of income in district B at a given taste α_1.

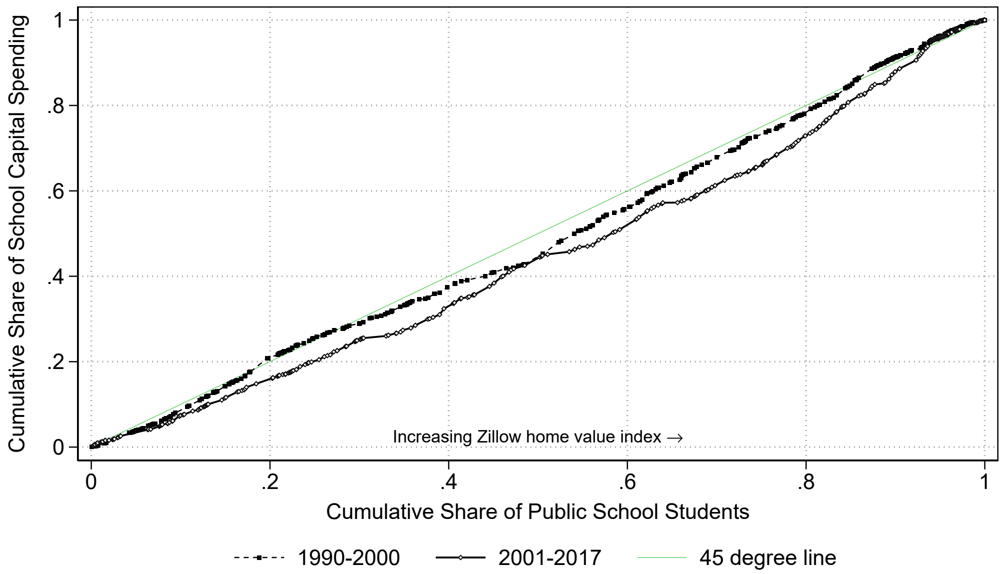
Figure 3: Cross-District Variation in Capital Outlays



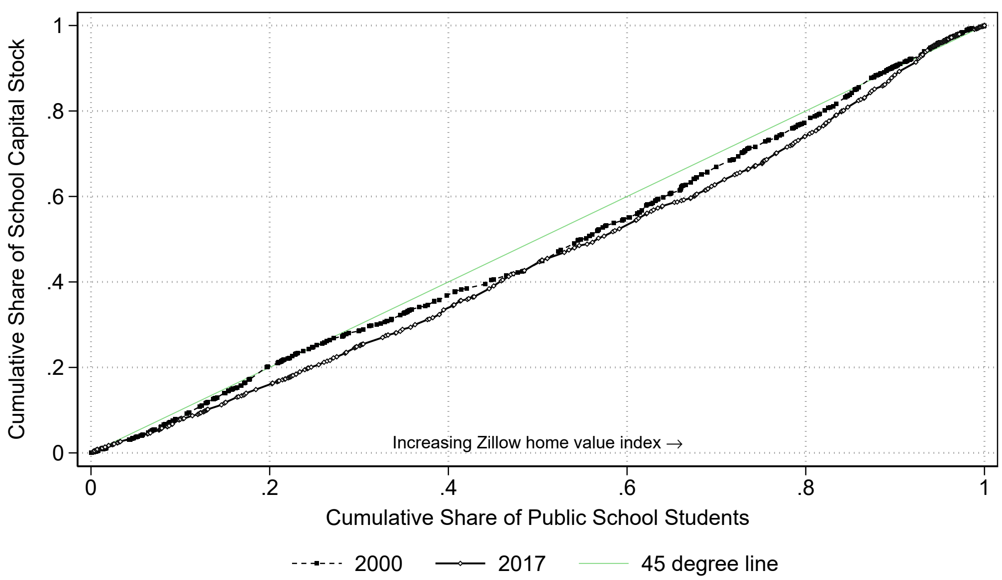
Notes: This figure plots a normalized measure of cross-district expenditure variance for three categories of school expenditure: instruction, support services and other current expenditures, and facilities expenditures. Specifically, the measure plotted is the standard deviation across districts divided by the mean expenditure level in the category in question. This is done to facilitate comparison of variance across categories.

Figure 4: Distribution of School Capital by Median Home Value

(a) Capital Outlays

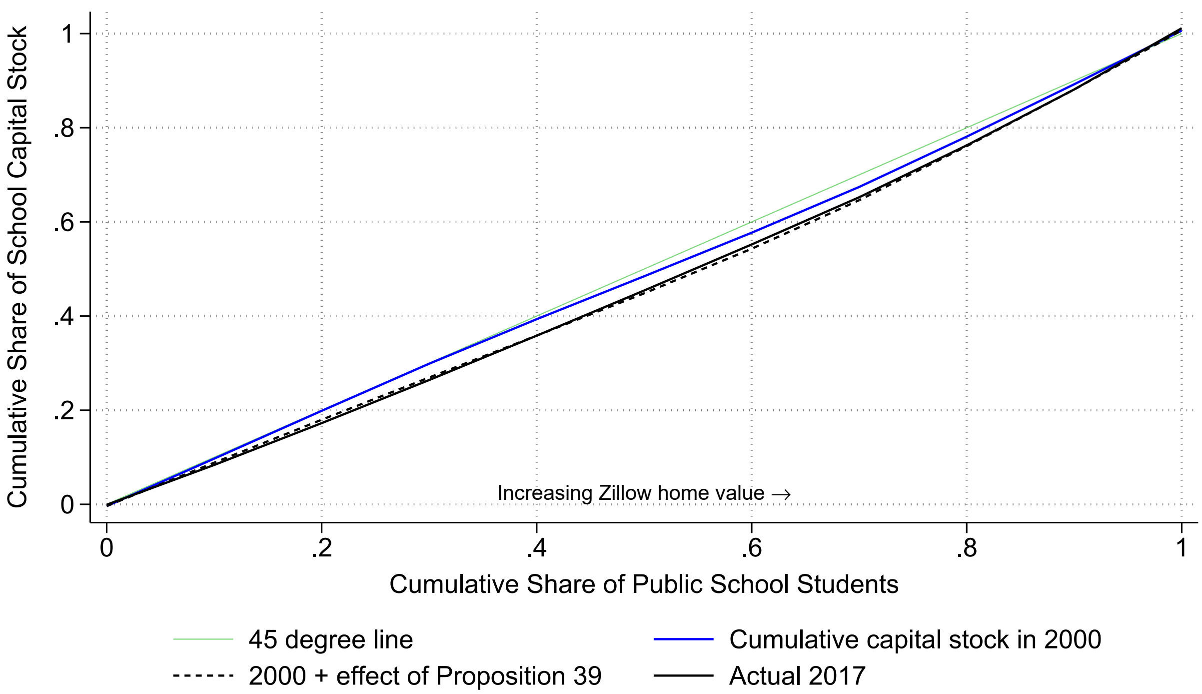


(b) Capital Stock



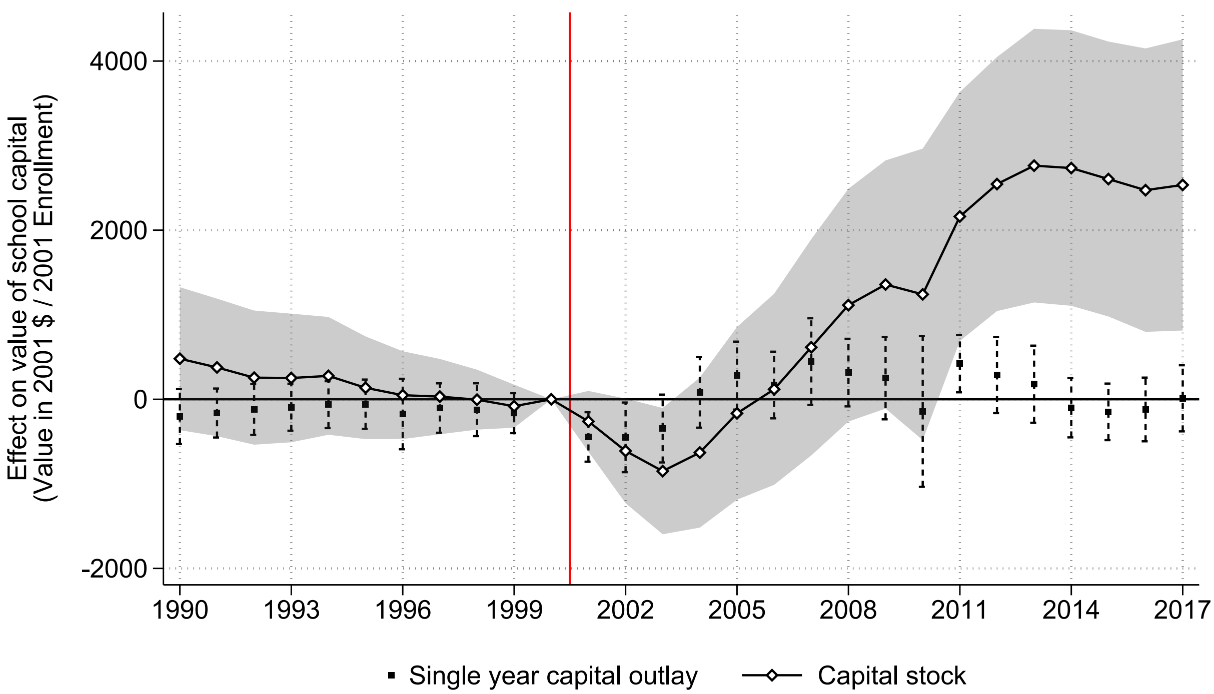
Notes: This figure plots concentration curves for school capital spending and stock before and after Proposition 39. The x-axes represent the cumulative share of students included in all districts left of x, and the y-axes represent the cumulative share of capital spending and stock, respectively. Panel (a) illustrates the distribution of spending from 1990-2000 and from 2001-2017. Panel (b) illustrates the distribution of capital stock in 2000 and 2017, including all previous capital spending since 1972, discounted at 5 percent annually.

Figure 5: Decomposition of Shift in Capital Stock Distribution



Notes: This figure decomposes the global shift in the concentration curve of school capital stock (defined in the notes for Figure 4, panel (b)) into the share attributable to changes in expenditure among districts directly affected by the reduction of the school bond approval threshold. The blue curve plots fitted values from a local polynomial fit for the 2000 global concentration curve. The dashed black curve adds fitted values from a local polynomial fit for the 2001-2017 change among directly affected districts to the 2000 global curve. This represents the shift in the global curve attributable to direct effects of the reform. The black curve represents the global fitted curve in 2017.

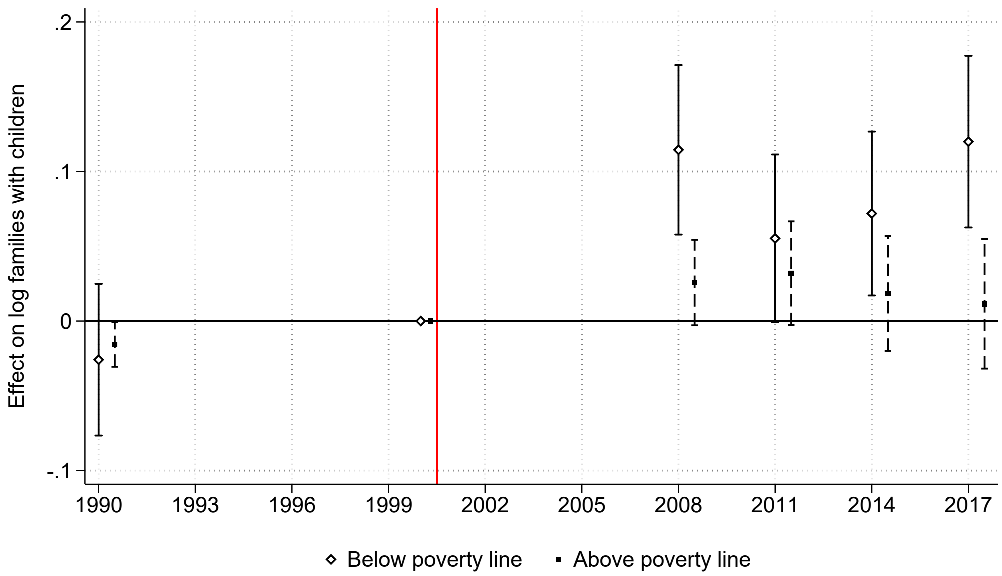
Figure 6: Effect of threshold reduction on outlays and stock of school capital



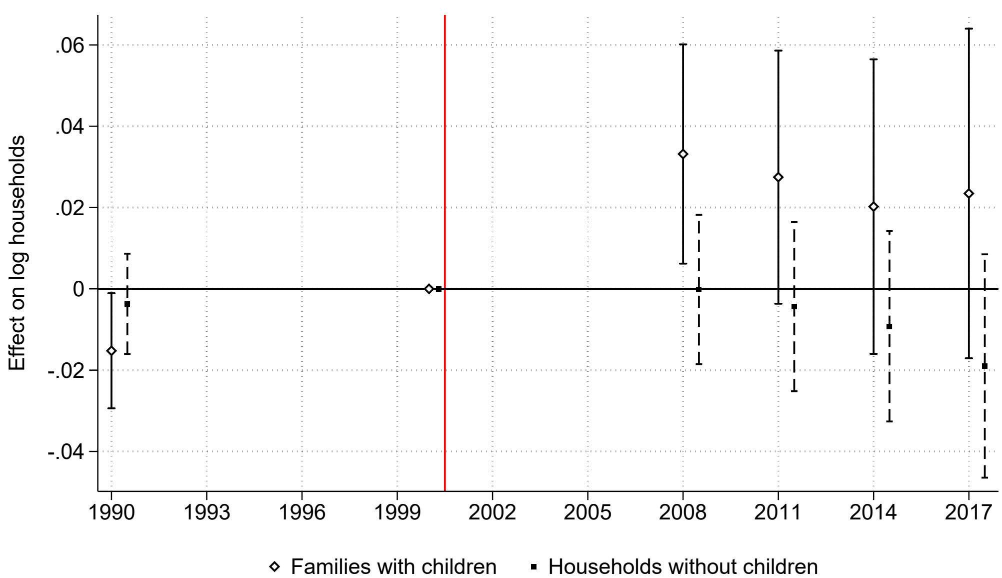
Notes: This figure reports the event-study estimates obtained after fitting Equation 4 on single year capital outlay (solid square points with dashed confidence intervals) or on total capital stock (hollow diamond points with shaded confidence interval). The omitted year is 2000, one year before the approval threshold reduction. Total capital stock is measured as the cumulative sum of capital outlays since 1972, depreciated at a 5% annual rate with a 30-year lifespan. The confidence intervals represent 95% confidence intervals using standard errors clustered at the district level.

Figure 7: Effect of threshold reduction on composition of households

(a) Effect on number of families with children by poverty status

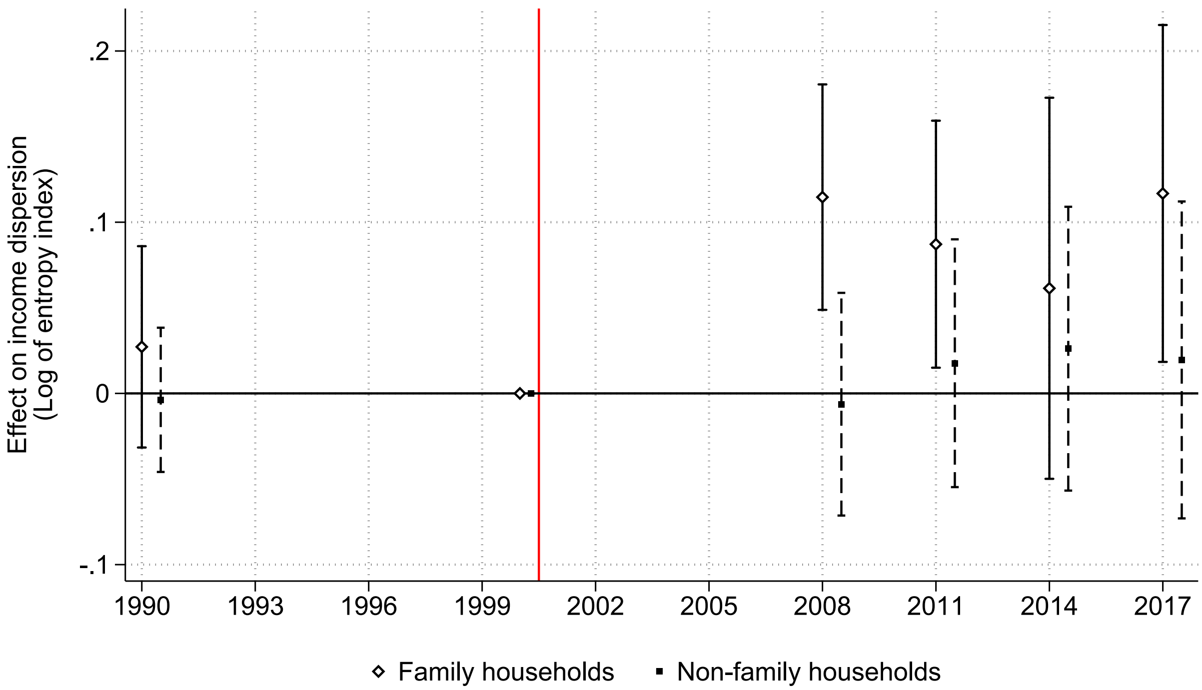


(b) Effect on number of households by presence of children



Notes: This figure reports the event-study estimates obtained after fitting Equation 4 on the log of households of different types, corresponding to a test of Prediction 2A in Section 3. Panel (a) shows estimates for family households with children above and below the poverty line. Panel (b) shows estimates for family households with children and for all households without children. The omitted year is 2000, one year before the approval threshold reduction. The confidence intervals represent 95% confidence intervals using standard errors clustered at the district level.

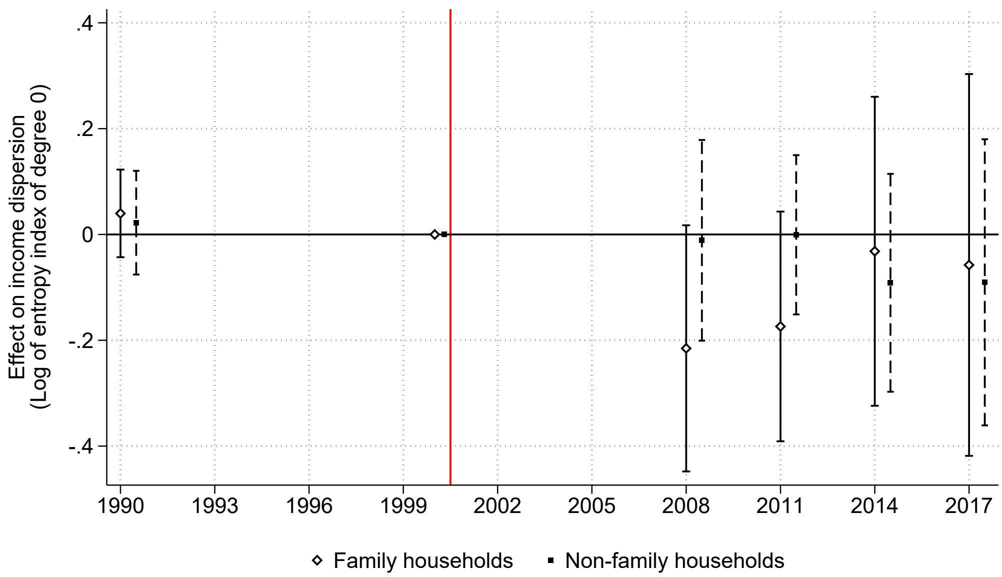
Figure 8: Effect of threshold reduction on income heterogeneity of households, conditioning on family status



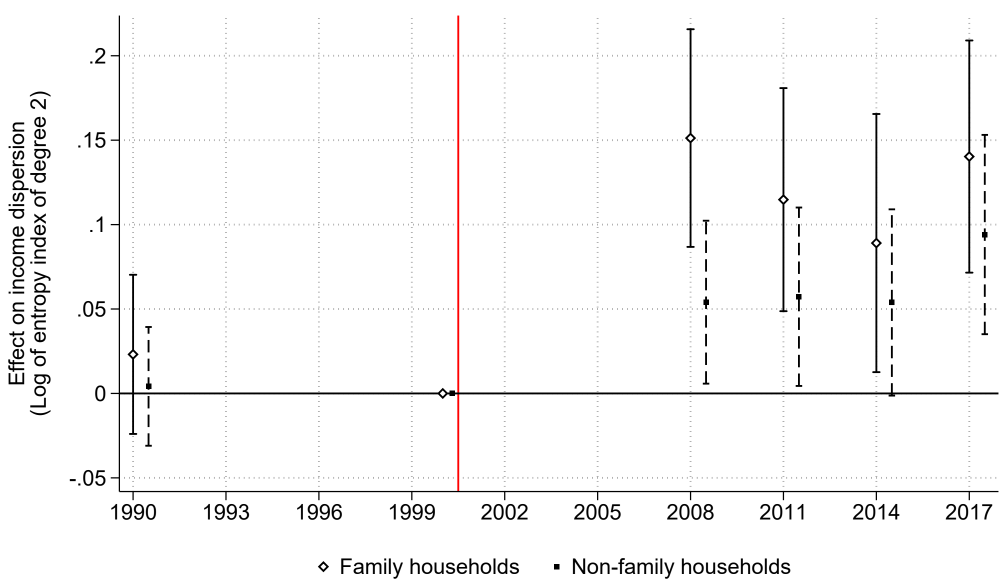
Notes: This figure reports the event-study estimates obtained after fitting Equation 4 on the log of income heterogeneity for family households (hollow diamond points with solid confidence intervals) or for non-family households solid square points with dashed confidence intervals). The former corresponds to a test of Prediction 2B in Section 3. The omitted year is 2000, one year before the approval threshold reduction. The measure of income heterogeneity is Theil's T statistic computed from the 15 income bins reported in Census/ACS data. Appendix B describes the calculation of Theil's T. The confidence intervals represent 95% confidence intervals using standard errors clustered at the district level.

Figure 9: Effect of threshold reduction on alternative measures of income heterogeneity of households, conditioning on family status

(a) Effect on low-end income heterogeneity (mean log deviation)

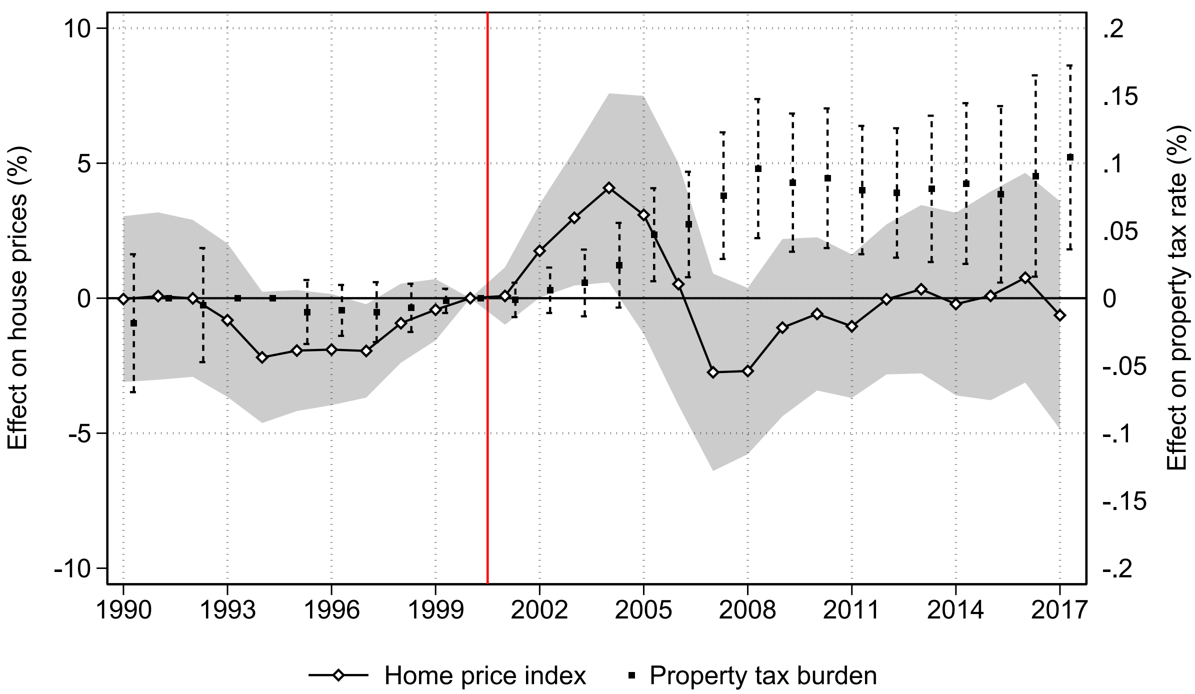


(b) Effect on high-end income heterogeneity (half the squared coefficient of variation)



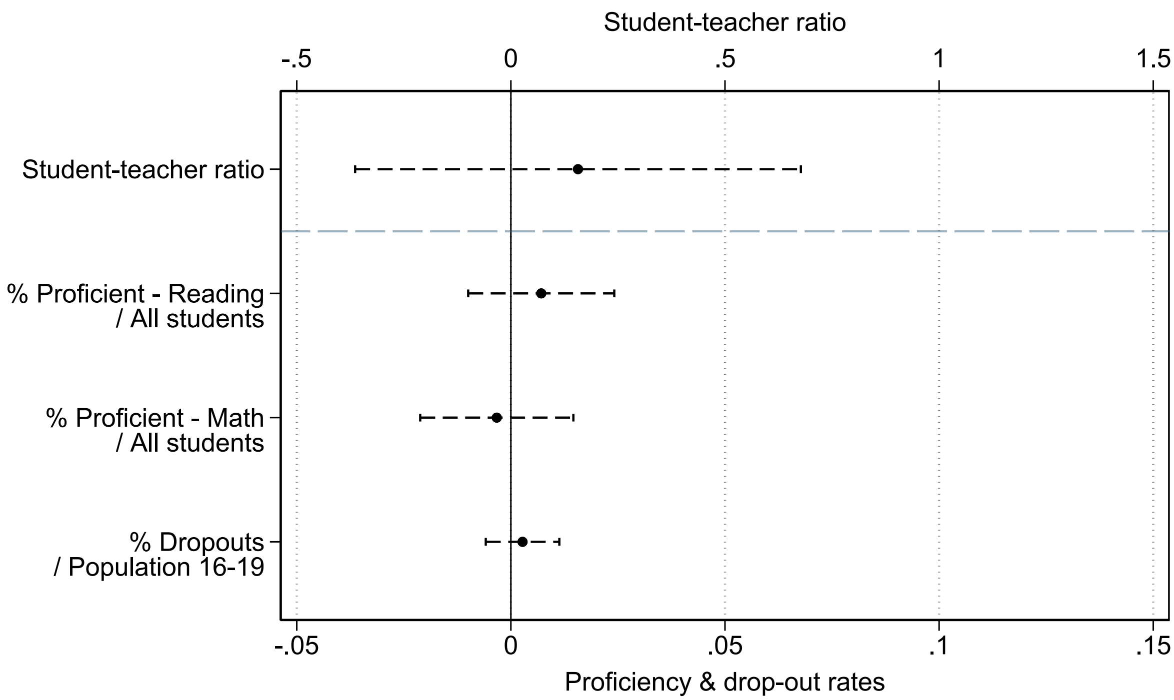
Notes: This figure reports the event-study estimates obtained after fitting Equation 4 on measures of income heterogeneity for family households (hollow diamond points with solid confidence intervals) or for non-family households solid square points with dashed confidence intervals). The measure in panel (a) is the mean log deviation (or generalized entropy of degree 0), which is particularly sensitive to heterogeneity in the low end of the distribution. The measure in panel (b) is Theil's L (or generalized entropy of degree 2), which is particularly sensitive to heterogeneity in the high end of the distribution. Both are computed from the 15 income bins reported in Census/ACS data. subsubsection 4.1.4 describes the calculation. The omitted year is 2000, one year before the approval threshold reduction. The confidence intervals represent 95% confidence intervals using standard errors clustered at the district level.

Figure 10: Effect of threshold reduction on house prices and effective property tax rates



Notes: This figure reports the event-study estimates obtained after fitting Equation 4 on the Contat-Larson house price index (hollow diamond points with shaded confidence intervals; Contat and Larson, 2021) or on an estimate of the effective property tax rate (square points with dashed confidence intervals). The omitted year is 2000, one year before the approval threshold reduction. The effective property tax rate is estimated by dividing total property tax revenues by aggregate home values; see Section 4 for a discussion. The confidence intervals represent 95% confidence intervals using standard errors clustered at the district level.

Figure 11: Effect of threshold reduction on academic outcomes in 2017



Notes: This figure reports the event-study estimate for the year 2017 obtained after fitting Equation 4 on the academic outcome indicated on the y-axis. The dropout rate is the share of the 16- to 19-year-old population that has not graduated high school and is not currently enrolled. The confidence intervals represent 95% confidence intervals using standard errors clustered at the district level.

Table 1: Summary Statistics: California Districts across Time

	1995		2000		2017	
	Mean	SD	Mean	SD	Mean	SD
	(1)	(2)	(3)	(4)	(5)	(6)
Enrolled pupils	5,792	[23,369]	6,322	[25,644]	6,639	[23,570]
Schools per 1,000 pupils	4.99	[9.65]	5.83	[12.45]	6.86	[15.84]
Charters per 1,000 pupils	0.00	[0.00]	0.09	[0.64]	0.30	[0.98]
Cumulative approved bond value	855	[2,884]	2,235	[4,939]	9,898	[14,262]
Long-term debt per pupil	689	[2,041]	2,140	[4,194]	6,517	[8,439]
Total revenue per pupil	6,362	[1,996]	8,073	[4,305]	10,776	[4,892]
Federal share	0.07	[0.05]	0.07	[0.06]	0.07	[0.04]
State share	0.52	[0.16]	0.55	[0.16]	0.54	[0.20]
Capital outlay per pupil (5 years)	4,393	[6,103]	4,393	[6,103]	4,393	[6,103]
State share	0.14	[0.24]	0.14	[0.24]	0.14	[0.24]
Instructional spending per pupil	3,441	[1,371]	4,075	[1,056]	5,780	[3,888]
Student-teacher ratio	23.21	[3.35]	19.90	[2.85]	21.78	[3.57]
Capital stock value per pupil	6660.37	[14582.26]	8950.53	[12602.02]	16047.75	[15734.25]
Families with children			8641.94	[23284.98]	8570.75	[20198.58]
Households without children			29600.86	[84349.03]	16339.57	[44997.87]
Poverty share among ages 5-17	0.19	[0.12]	0.18	[0.11]	0.19	[0.11]
Population			63916.75	[1.8e+05]	75480.02	[1.9e+05]
White			32308.65	[59242.70]	30183.10	[58062.39]
Black			3421.95	[19628.12]	3567.14	[16981.13]
Hispanic			19573.89	[81584.27]	28534.40	[93898.33]
College educated			10680.48	[29444.81]	17044.66	[46554.13]
Owner-occupied housing units			13148.59	[26337.73]	14665.31	[27738.42]
Renter-occupied housing units			8426.75	[34129.02]	10245.00	[38154.26]
Moved in last year			5807.65	[16159.77]	9485.80	[22421.88]
Moved in last year – homeowners			1447.35	[2688.34]	716.90	[1256.84]
Moved in last year – renters			2979.72	[10633.35]	1611.85	[5241.02]
Moved in last year – within metro			3479.48	[11690.26]	6179.70	[16181.24]
Moved in last year – into metro			2398.05	[4954.81]	3245.06	[6587.83]
Median home value, Zillow	1.9e+05	[1.1e+05]	2.5e+05	[2.0e+05]	5.5e+05	[4.6e+05]

Notes: This table reports summary statistics for California school districts six years before the implementation of Proposition 39 (columns (1) and (2)), in the year of implementation (columns (3) and (4)), and at the end of this paper's study period (columns (5) and (6)). Per-student values are calculated by dividing the numerator by the number of students reported to be enrolled in the entire district in the specified year. Student-teacher ratios are winsorized at the 5th and 95th percentiles. State capital outlays are the amount of state transfers to school districts for capital outlay and debt service programs. Federal capital outlays on schools are negligible and not collected in the school district finance survey in any of these years. Monetary values are 2001 dollars.

Table 2: Income spread explains differential take-up of bonds following Proposition 39

	(1)	(2)	(3)
	Approves at least 1 bond		Log approved bond value per student
(a) No history of bond approval before 2001			
2000 IQR of household income, SD	0.10^{**}	0.02	0.38^{***}
	(0.03)	(0.02)	(0.09)
Total proposals, 2001-2010		0.37^{***}	
		(0.02)	
N	570	570	224
adj. R^2	0.028	0.511	0.106
(b) At least one bond approval before 2001			
2000 IQR of household income, SD	0.08^{**}	0.04^{*}	0.21^{***}
	(0.02)	(0.02)	(0.03)
Total proposals, 2001-2010		0.40^{***}	
		(0.03)	
N	359	359	179
adj. R^2	0.026	0.500	0.101

Notes: This table reports results from a regression of an indicator for approving at least 1 bond between 2001 and 2010 (columns 1 and 2) and the log of total approved bond value per student (column 3) on the interquartile range of 2000 household income. Column 2 controls for the total number of bond proposals during the ten-year period. $^{*}p < 0.5,^{**} p < 0.01,^{***} p < 0.001$.

Table 3: Summary statistics: California districts by bond behavior, 2000

	Never-takers (No new bonds)		Compliers (Bond passes <2/3)		Always-takers (All bonds pass >2/3)	
	Mean	SD	Mean	SD	Mean	SD
	(1)	(2)	(3)	(4)	(5)	(6)
Share unified	0.32	[0.47]	0.57	[0.50]	0.49	[0.50]
Location type (locale)						
City	0.07	[0.25]	0.23	[0.42]	0.25	[0.44]
Suburb	0.22	[0.41]	0.40	[0.49]	0.46	[0.50]
Town	0.13	[0.34]	0.20	[0.40]	0.11	[0.31]
Rural	0.57	[0.50]	0.17	[0.38]	0.18	[0.39]
Enrolled students	2,322	[5,220]	11,694	[49,678]	9,703	[15,013]
Schools per 1,000 students	9.93	[17.51]	2.41	[2.36]	2.70	[3.73]
Charters per 1,000 students	0.14	[0.93]	0.05	[0.20]	0.05	[0.27]
Cumulative approved bond value	3,326	[6,957]	1,933	[3,576]	3,070	[5,250]
Long-term debt per student	1,718	[4,271]	2,058	[3,063]	2,965	[5,928]
Total revenue per student	8,728	[6,144]	7,263	[1,166]	7,970	[2,024]
Federal share	0.08	[0.07]	0.06	[0.04]	0.07	[0.04]
State share	0.56	[0.16]	0.56	[0.14]	0.53	[0.19]
Annual capital outlay per student	948	[1,702]	780	[569]	857	[637]
State share	0.12	[0.23]	0.17	[0.19]	0.15	[0.20]
Instructional spending per student	4,320	[1,349]	3,813	[538]	4,049	[947]
Number of districts	402		208		138	

Notes: This table reports summary statistics for California school districts according to their observed bond usage in the years 2001-2010. Districts that pass a bond with less than 2/3 support in 2011 or later are excluded as contaminated controls. Columns (1) and (2) report means and standard deviations for the never-takers, i.e. districts that passed no new bonds between 2001 and 2010. Columns (3) and (4) report means and standard deviations for the compliers, districts that approved bonds with 55%-2/3 of votes in the same ten-year span. Columns (5) and (6) report means and standard deviations for the always-takers, districts that gave all approved bonds greater than 2/3 support in the ten-year span. Covariates are observed in 2000, except capital outlay, which is averaged over the years 1996-2000, and locale, for which the 2017 value is reported. Monetary values are 2001 dollars.

A Theory Appendix

A.1 Voting Equilibrium with an Arbitrary Supermajority Approval Threshold

EP provide a proof of sufficient conditions for the existence of a unique equilibrium tax-spending bundle when the winning outcome requires a simple majority ($d = 0.5$). I provide a simple extension that shows this result extends to an arbitrary supermajority approval threshold $d > 0.5$. In EP's proof for their Proposition 2, assumption (ii) refers to a simple majority threshold of 0.5. This can be generalized by replacing 0.5 with any threshold $d > 0.5$ representing the share required for passage. The remainder of the proof follows as in the original paper, with the locus of pivotal voters shifting according to the new threshold d, proving existence. This change may limit the existence of a *non-zero* voting equilibrium because under a voting rule $d > 0.5$, the existence of a function $\tilde{y}^j(\alpha)$ satisfying assumption (ii) may be less likely to exist. Note that the proof does not follow for $d < 0.5$ because it cannot rule out a larger coalition for a different level of spending.

A.2 Rotation of Indifference Boundaries

This section shows that the rotation of the indifference boundaries that follows from a relative increase in spending in B is an analytical feature of the model. The intuition for this result is described in Section 3. Here I provide a formal proof in two steps. I begin by deriving the shifts in the indifference boundary following a relative increase in K_B under the assumption that housing prices in all communities are fixed. This provides the core intuition for how budget-balanced changes in public spending and tax levels influence stratification. Subsequently, I extend the proof to a case where house prices in all communities are fully endogenous to spending and tax levels.

A.2.1 Fixed-Price Case

The indifference boundary between communities A and B is determined by the following condition on indirect utilities:

$$V(K_A, p_a; \alpha, I) = V(K_B, p_b; \alpha, I), \tag{5}$$

where K_j is public spending in district j, p_j is the fixed unit price of housing, α is household taste for education, and I is household income. This equality implicitly defines this indifference boundary as a function $\alpha^A(I)$. Under the assumptions imposed by EP—strictly quasi-concave utility with public good-housing cost indifference curves with decreasing slopes in both income and taste, as noted in Footnote 9—the marginal rate of substitution between K and the numeraire, denoted MRS_K, is continuous and decreasing in income, while the marginal rate of substitution between housing and the numeraire, MRS_h is continuous and increasing in income. Thus, for a sufficiently low income

I, there is a neighborhood around $(\alpha^A(I), I)$ with $MRS_K > MRS_h$. Following a joint increase in K_B and τ_b, $V(K_A, p_a; \alpha, I) < V(K_B, p_b; \alpha, I)$ for the households in this neighborhood. Visually, this is represented as a down and left shift of the indifference boundary at $(\alpha^A(I), I)$. Due to the continuity and monoticity of the marginal rates of substitution, the magnitude of this neighborhood, and consequently of the shift, shrinks as I increases. Furthermore, there is some sufficiently large I such that this pattern is reversed, and the indifference boundary shifts up and right following a joint increase in K_B and τ_B. In aggregate, the indifference boundary rotates counter-clockwise. By a symmetrical argument, the indifference boundary between B and C rotates clockwise.

A.2.2 Endogenous Housing Prices

Next, I extend the derivation to incorporate endogenous adjustments to housing prices in all districts.

The vector of housing prices $p = \{p_A, p_B, p_C\}$ is a function of the vector of public spending and tax levels $(K, \tau) = \{(K_A, \tau_A), (K_B, \tau_B), (K_C, \tau_C)\}$. Any change in (K, τ) affects p through residential re-sorting, with p_j rising if aggregate housing demand rises in j. Considering households along the $A - B$ indifference curve with sufficiently low income such that $MRS_K > MRS_h$, if housing demand falls in B, the downward and leftward shift in the boundary is amplified by the increased affordability of housing in B. If housing demand increases in B, the shift is reduced but not reversed, as $MRS_K > MRS_h$ still holds. Equivalent arguments hold at all other points on the original indifference boundaries for any set of changes in p. Thus, while endogenous price adjustment affect the magnitude of boudnary rotations, they do not alter their directionality.

B Data and Measurement Appendix

$GE(1)$ represents neutrality among these three measures, and is closely related to the better-known Theil's H index of multi-group entropy. Theil's H measures the inequality of a distribution across K categorical groups in a population, taking the following form:

$$H = \sum_{k=1}^{K} \frac{N_k}{N} \ln\left(\frac{N_k}{N}\right) \tag{6}$$

In the formula above, N_k represents the count of observations in group k and N represents the size of the entire population. H is a generalization of a more specific inequality measure, $GE(1)$ (Theil's T). Given average incomes $\bar{y}_l$ (or another cardinally measured dimension of inequality) for each group k, the measure is:

$$\text{Theil's } T = GE(1) = \sum_{k=1}^{K} \frac{N_k}{N} \times \frac{\bar{y}}{\bar{y}_k} \ln\left(\frac{\bar{y}}{\bar{y}_k}\right) \tag{7}$$

where N_k is the count of observations in group k; N is the total population; $\bar{y}_k$ is the average income in k; and $\bar{y}$ is the average income in the population.

Conceptually, both of these statistics represent the amount of information needed to describe the deviation of the distribution from perfect equality. They are bounded below by 0, representing perfect equality. $GE(1)$ is unbounded above, by while Theil's H is bounded above by $\ln K$.

Where $GE(1)$ is neutral with respect to where inequality is present in the distribution, generalized entropy of degree 0, also referred to as mean log deviation, describes inequality throughout the distribution with an emphasis on dispersion at the bottom:

$$\text{mean log deviation} = GE(0) = \frac{1}{N} \sum_{k=1}^{K} \ln\left(\frac{\bar{y}}{\bar{y}_k}\right) \tag{8}$$

On the other end of the spectrum, generalized entropy of degree 2, or half the squared coefficient of variation, describes inequality throughout the distribution with an emphasis on dispersion at the top:

$$\text{half the squared coefficient of variation} = GE(2) = \frac{1}{2N} \sum_{k=1}^{K} N_k \left(\frac{\bar{y}}{\bar{y}_k} - 1\right)^2 \tag{9}$$

A special characteristic of these indices is their utility in additively aggregating and decomposing inequality. For example, considering the case of $GE(1)$, the total inequality in a school district d with total population N_d, consisting of J Census tracts, can be represented as

$$GE_d(1) = \underbrace{\sum_{j=1}^{J} \frac{N_j}{N_d} \times \frac{\bar{y}_d}{\bar{y}_j} \ln\left(\frac{\bar{y}_d}{\bar{y}_j}\right)}_{\text{between-tract inequality}} + \underbrace{\sum_{j=1}^{J} \frac{N_j}{N_d} \times T_j}_{\text{within-tract inequality}} \tag{10}$$

In Equation 10, the first right-hand-side summation represents inequality in average incomes across districts, and the second summation is a population-weighted average of income inequality within districts. Summing these, $GE_d(1)$ represents overall inequality in the district. Analogously, H_d can be calculated at the district level and decomposed into between-school and within-school inequality.

In practice, I observe different information about income distributions at different levels of observation. At the Census tract and school district level, the Census/ACS data report household counts in 15 time-invariant nominal income bins topping out at $150,000 and above. In these cases, I calculate $GE(0), GE(1)$, and $GE(2)$ using the bracket midpoints as approximations of the average income in brackets 1 through 15, and I use $250,000 for bracket 15. At the school enrollment level, lunch subsidy eligibility (free/reduced-price/ineligible) is the most detailed proxy for income status, and I calculate Theil's H.

C Propensity Score Estimation and Weighting

This appendix describes the approach used to estimate propensity scores for the inverse probability weighting procedure in the main analysis. The goal is to estimate each district's probability of leveraging the reduced approval threshold based on pre-reform characteristics, allowing for complex interactions among district features that might influence selection into take-up. For instance, a district's property wealth might predict different take-up patterns depending on its demographic composition and existing capital stock. Weighting the event study regressions by the inverse of these estimated probabilities mitigates the potential selection bias introduced by systematic differences between complier and non-complier districts.

To capture these potential nonlinearities and interactions without requiring them to be specified ex ante, I employ a neural network to estimate the propensity scores. This approach offers advantages over traditional methods like logistic regression, particularly in its ability to discover interaction effects in the data and naturally accommodate the multi-class nature of the outcome (never-takers, compliers, and always-takers). Empirically, it also outperforms alternative machine learning algorithms, including random forests and gradient boosting.

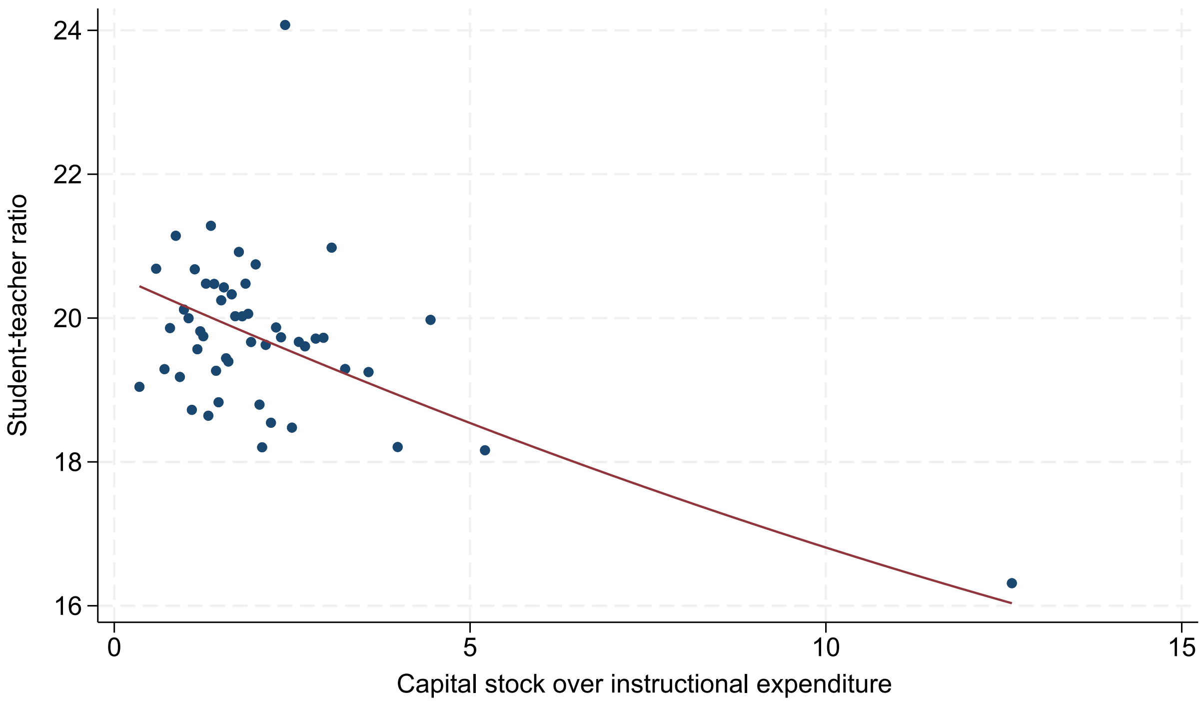
The neural network takes as input 42 pre-reform district characteristics, including capital outlays and their growth rates from the 1990s, operational metrics such as schools per pupil and long-term debt, student demographics and enrollment, housing market conditions including home values and rental rates, and population characteristics like racial composition and poverty rates. These inputs are normalized using batch normalization. The network architecture consists of five hidden layers containing [30, 50, 50, 50, 30] neurons respectively, with ReLU activation in the first layer and Leaky ReLU activation (negative slope $= 0.1$) in subsequent layers. Dropout regularization ($p = 0.5$) between hidden layers prevents overfitting. The output layer contains three nodes with softmax activation, producing probabilities for each class.

The model is trained on California school districts with known outcomes using a random 80-20 train-test split. To address class imbalance, where compliers and always-takers represent minority classses, the cross-entropy loss function employs class weights of [1.0, 4.0, 2.0]. The model is optimized using Adam with learning rate 0.001 and weight decay 0.0001, with the learning rate decaying by half every 1,000 epochs. Model architecture and hyperparameters were selected through manual tuning, focusing on balancing predictive power against overfitting concerns. The final model achieves moderate predictive performance, with out-of-sample accuracy of 0.72 for identifying complier districts and overall AUC-ROC of 0.69. These metrics represent a meaningful improvement over random assignment and indicate the model captures substantive variation in districts' propensities to leverage the reform.

The predicted probabilities are used to construct stabilized inverse propensity weights for event study analysis. Standard inverse propensity weights can exhibit high variance when predicted probabilities are close to 0 or 1, potentially leading to imprecise estimates and sensitivity to model specification. Stabilized weights adress this issue by incorporating the marginal probability of treatment, reducing the influence of extreme predictions while maintaining the weights' balancing properties. For each district, the predicted probability of being a complier is first winsorized at the 5th and 95th percentiles to further mitigate the influence of extreme weights. Let p_c denote the sample proportion of complier districts. The stabilized weights are then constructed as $p_c/\hat{p}$ for complier districts and $(1 - p_c)/(1 - \hat{p})$ for non-complier districts, where $\hat{p}$ represents the winsorized predicted probability of compliance. This weighting scheme, combined with the event study specification, provides a "doubly robust" estimation approach that remains consistent if either the event study or propensity score model is correctly specified.

D Additional Figures and Tables

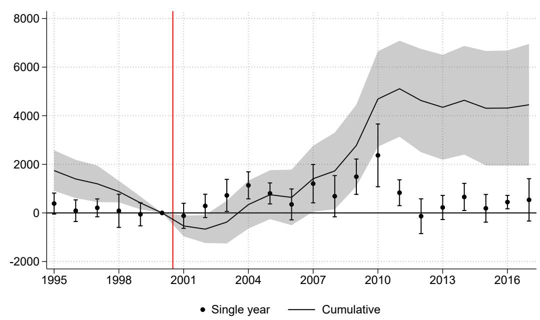
Figure C1: Higher capital stock is correlated with lower student-teacher ratios (2000)



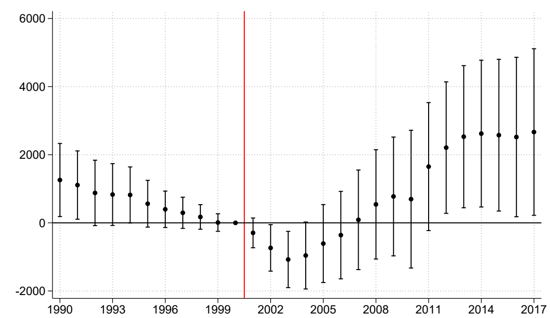
Notes: This figure is a bin scatter plotting the ratio of total capital stock over annual instruction expenditure on the x-axis against student-teacher ratios on the y-axis. The red curve illustrates fitted values from a quadratic regression of student-teacher ratio on capital stock over instructional expenditure. The results are essentially unchanged if rural school districts are dropped.

Figure C2: More fiscal impacts of approval threshold reduction on compliers

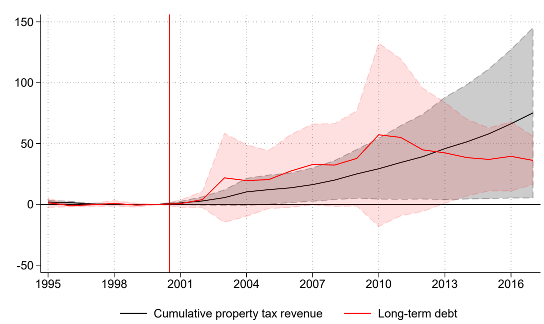
(a) Value of bonds approved



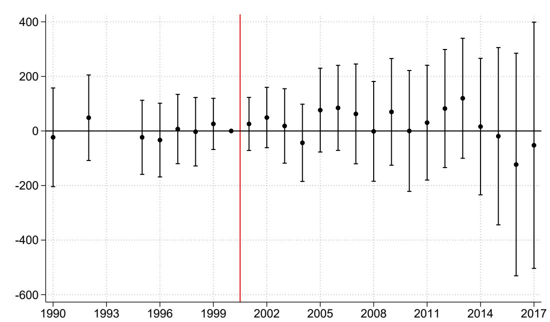
(b) Cumulative capital outlays, ignoring depreciation



(c) Cumulative tax revenue versus long-term debt



(d) Instruction spending per student



Notes: The dependent variable represented by circular points is single year capital outlay, the flow of capital spending. The dependent variable represented as a solid line is the total stock of school capital.

Table C1: Statewide school bond proposals in California, 1986-2020

Election	Measure name	Proposed funding	Outcome	Vote in favor	Vote against
1986 General	Proposition 53	$800 million	Passed	60.73%	39.27%
1988 Primary	Proposition 75	$8 million	Passed	64.95%	35.05%
1988 General	Proposition 79	$800 million	Passed	61.23%	38.77%
1990 Primary	Proposition 123	$800 million	Passed	57.52%	42.48%
1990 General	Proposition 146	$800 million	Passed	52.53%	47.47%
1992 Primary	Proposition 152	$1.9 billion	Passed	52.92%	47.08%
1992 General	Proposition 155	$900 million	Passed	51.80%	48.20%
1994 Primary	Proposition 1B	$1 billion	Failed	49.57%	50.43%
1996 Primary	Proposition 203	$3 billion	Passed	61.95%	38.05%
1998 General	Proposition 1A	$9.2 billion	Passed	62.49%	37.51%
2002 General	Proposition 47	$13.05 billion	Passed	59.08%	40.92%
2004 Primary	Proposition 55	$12.3 billion	Passed	50.85%	49.15%
2006 General	Proposition 1D	$10.416 billion	Passed	56.90%	43.10%
2016 General	Proposition 51	$9 billion	Passed	55.18%	44.82%
2020 General	Proposition 13	$15 billion	Failed	46.99%	53.01%

Notes: This table catalogs all statewide proposals decided by voters between 1986 and 2020 that included bonds for K-12 facilities. Statewide bond proposals require a simple majority in favor for approval. In some cases, the proposed funding includes allocations for post-secondary facilities (Proposition 1A 1998 and all proposals since 2004). Approved bonds are obligations of the state's General Fund and are not directly associated with a tax increase in the way that local bonds are. Source: UC Law SF Scholarship Repository. https://repository.uclawsf.edu/ca_ballots/

.