

Housing Disease and Public School Finances

Matthew Davis[§] and Fernando Ferreira^{§§}

Abstract: Median expenditure per student in U.S. public schools grew 41% in real terms from 1990 to 2009. We propose a new mechanism to explain part of this increase: housing disease, a fiscal externality from local housing markets in which unexpected booms generate extra revenues that schools administrators have incentives to spend, independent of local preferences for provision of public goods. We establish the importance of housing disease by: (i) assembling a novel microdata set containing the universe of housing transactions for a large sample of school districts; and (ii) using the timelines of school district housing booms to disentangle the effects of housing disease from reverse causality and changes in household composition. We estimate housing price elasticities of per-pupil expenditures of 0.16-0.20, which accounts for approximately half of the rise in public school spending. School districts did not boost administrative costs with those additional funds. Instead, they primarily increased spending on instruction and capital projects, suggesting that the cost increase was accompanied by improvements in the quality of school inputs.

[§]The Wharton School, University of Pennsylvania. Email: mattda@wharton.upenn.edu.

^{§§}The Wharton School, University of Pennsylvania, and NBER. Email: fferreir@wharton.upenn.edu.

We thank the Research Sponsors Program of the Zell/Lurie Real Estate Center at Wharton for financial support. We are grateful to Qize Chen, Stella Yeayeun Park, and Xuequan Peng for providing research assistance. We also would like to thank Moshe Buchinsky, Steven Craig, Caroline Hoxby, Bob Inman, Till Von Wachter, and the seminar participants at University of Houston, UCLA, Insper, and the NBER Economics of Education meeting for valuable comments and suggestions.

I. Introduction

Median expenditure per student in U.S. public schools grew from \$9,131 in 1990 to \$12,907 in fiscal year 2008/9, a real increase of 41%. Figure 1A shows that school districts in the top and bottom percentiles with respect to expenditures per pupil also had similar patterns. While the large amount of resources devoted to public education still sparks a debate over whether money matters for improving school quality,¹ in this paper we focus on understanding why the recent growth happened in the first place. We propose a new mechanism, housing disease, based on spillovers from housing markets. Figure 1B shows real average house prices for U.S. school districts. The median district had a real increase in average prices of 70%, moving from \$159K in 1993 to \$274K by the end of 2007. The 90th percentile district grew by almost 91% and even the 10th percentile district increased by 32%. Are these large swings in house prices responsible for the dramatic changes in education expenditures?

House prices usually have a limited, dependent role in public finance theory. Starting with Tiebout (1956), local finances are determined by individual preferences or by “voting with your feet,” i.e., by households sorting into local communities that provide their preferred quality of public services.² House prices are just a function of local taxes and amenities (Oates 1969) and therefore can be used to recover willingness to pay for local public goods and to test whether those goods are provided at efficient levels.³

Housing disease reverses this standard relationship. First, housing cycles generate unusually high growth rates of housing prices. That triggers a growth in school district revenues given that local governments raise some of their funds via property or land taxes and fees. In turn, school district administrators may have incentives to spend the extra revenues without consulting voters due to rent seeking, budget rules, or frictions in re-optimizing tax rates. The end result is an increase in education expenditures without a corresponding shift in local preferences. This type of mechanism is not unprecedented in the economics literature. In fact

¹ Some key studies on this topic include Coleman et al. (1966), Hanushek (1986), Card and Krueger (1992), Krueger (1999), Hanushek and Rivkin (2006), Jackson, Johnson, and Persico (2016), and Lafortune, Rothstein, and Schanzenbach (forthcoming).

² A long literature shows the importance of household preferences and sorting for determining the quality of public education, such as Epple and Sieg (1999), Fernández and Rogerson (2001), Hilber and Mayer (2009), and Epple, Romano and Sieg (2012).

³ See Bayer, Ferreira and McMillan (2007) on how to estimate willingness to pay for school quality using housing prices, and Brueckner (1979), Barrow and Rouse (2004), and Cellini, Ferreira and Rothstein (2010) on how to test for efficiency in the provision of local public goods.

we use the word “disease” to emphasize its similarity to Baumol and Bowen (1966)’s cost disease, a canonical example of a spillover to the cost of public education stemming from conditions in a separate market. The primary difference is that, whereas Baumol and Bowen’s cost disease originates in the labor market, the housing disease’s genesis is the housing market.

The first challenge in estimating the importance of housing disease is that house prices are endogenous to school quality and household composition. We deal with this issue by using the timeline of housing booms in each school district in our sample. The variation from local housing booms has two features that are key to our research design. First, different school districts began to boom across a decade-long period from mid-1990s to the mid-2000s, some of them multiple times, allowing us to remove the impact of national macroeconomic factors. Second, the beginning of a boom is associated neither with changes in proxies for school quality nor with widespread changes in household composition. In Section IV we show how to estimate the timeline of local booms using time series methods developed by Ferreira and Gyourko (2011) and empirically validate the research design by directly testing the two key features above.⁴

The second challenge is that housing data is generally not available for a large sample of school districts. We solve this problem by amassing the most recent version of the CoreLogic universe of housing transactions from 1993 to 2013, and mapping each home to school district boundaries. Our sample covers more than 2,000 school districts with almost 60% of all total enrollment in public schools. The micro data allow us to use a split-sample approach, such as in Card, Mas and Rothstein (2008), to deal with specification search bias that arises when the same time series is used to estimate both the timeline and magnitude of a housing boom (Leamer 1983). The first random sample is used to create a price index for each district and estimate the timing of local booms. The hold out sample is then used to estimate the magnitude of price changes along the cycle.

We find that school district house prices increase by over 4 percent in the first year of the boom when compared to the pre-boom year, net of other housing booms in the same district and of time and district effects. Prices keep increasing and are nearly 20 percent larger by the end of the fourth year. Next, we estimate how school finances react to the timeline of the boom.

⁴ Charles, Hurst, and Notowidigdo (2015) and DeFusco et al (2017) use a similar methodology to estimate the impact of housing booms on investments in human capital and on price increases in nearby metro areas, respectively.

Expenditures per pupil start to creep up with a one to two-year lag, turning statistically significant at year 3 and becoming 3.3 percent larger by the fourth year of a boom.

With those magnitudes in hand we can back out the house price elasticity of public school finances. We find an elasticity between 0.16 and 0.20. This relatively small elasticity makes sense given that a large fraction of school district revenues now come from state and federal transfers, especially for low income districts.⁵ In fact, we find heterogeneity in the estimation of elasticities that matches this intuition: districts that spent below (above) the median early in the data have smaller (bigger) elasticities than average, though the difference is not statistically different from zero. But the last housing boom was so extraordinary that housing disease can account for almost half of the real increase in public education spending since the early 1990s. Our back-of-the-envelope calculations also show that housing disease can account for a higher fraction of the spending increase for the school districts in the highest percentiles of school expenditure per pupil.

This result is a breakdown of the theoretically efficient choices made by Tiebout-type households. But it does not necessarily imply that all additional resources are being wasted. Pupil-teacher ratios, a proxy for educational quality, improve but at a fairly small rate (less than 1% reduction in pupil-teacher ratio), and capital expenditures increased markedly (18% increase four years after the boom). We also find increases in average salaries and benefits (2.9% and 3.7% respectively), which could reflect increases in productivity or rent-seeking, a possibility raised by Brueckner and Neumark (2014) and Diamond (2017). Interestingly though, we do not find new resources being used to disproportionately increase administrative expenses. Overall, the fact that pupil-teacher ratios increased, capital budgets grew, and administrative expenses remained flat suggests that housing disease is accompanied by improvements in the quality of school inputs, and that bureaucrats are not capturing most of the increased expenditures. We also provide impact estimates based on NAEP test scores, though this analysis suffers from additional data shortcomings, noisy measurements, and limited statistical power. Taken at face value, the results suggest that math scores remain flat, but reading scores may increase in the medium to long term.

⁵ Those transfers now correspond to more than 50% of total revenues, but this number is difficult to properly measure given that the data may not distinguish between the jurisdiction that collects taxes versus the jurisdictions that actually has control over taxes – see Hoxby (1996).

There are a couple of caveats with our results. First, we do not observe how local rules, regulations and tax rates change over time, so we are likely underestimating our elasticities. This lack of data plagues almost all research in school finance and it is due to the nature of the decentralized public system of education. Second, the estimated elasticities may not be symmetric for booms and busts. The descriptive evidence suggests that housing busts are less likely to affect school finances because it is difficult to downsize and also due to the help of state and federal transfers. But we cannot properly test the asymmetry because housing busts are usually accompanied by negative income shocks.

Our estimates contribute to the understanding of the dramatic increase in public education spending of 1990s and 2000s. There are many determinants of the level and quality of local public finance, starting with fiscal federalism,⁶ household preferences and Tiebout sorting, local governmental decisions and transfer schemes,⁷ local autonomy and competition⁸, and more recent “mandates” such as pension benefits and special education. We propose and test a new mechanism, housing disease, that is generally not taken into account by standard theory and is difficult to study given data and design limitations.

Moreover, the relevance of housing disease is likely to increase in the near future because extreme price fluctuations are becoming a feature of the system as opposed to a one-time bug. Housing markets are now characterized by many local housing booms and busts (Ferreira and Gyourko (2011), Sinai (2013)), fueled by both behavioral and financial factors (Shiller (2005), Mian and Sufi (2009), Favara and Imbs (2015)), and exacerbated by regulations that limit the supply of new housing (Glaeser and Gyourko (forthcoming)).

The remainder of the paper is organized as follows: Section II reviews how school district finances work and the potential for housing disease; Section III then describes the data sources and sample construction; Section IV then describes our empirical framework and test the validity of our research design; Section V presents our estimates; and Section VI concludes.

⁶ Reviews of the fiscal federalism literature can be found in Oates (1999, 2005).

⁷ For the impact of local politics see Ferreira and Gyourko (2009) and more recently Macartney and Singleton (2017). For the effects of equalization and transfer schemes in education see Murray, Evans and Schwab (1998), Hoxby (2001), Bradbury, Mayer and Case (2001), Card and Payne (2002), and more recently Jackson, Johnson, and Persico (2016) and Lafortune, Rothstein and Schanzenbach (forthcoming).

⁸ See Hoxby (2000), Rothstein (2007), Hoxby (2007), and Clark (2009).

II. Public School Finances and Housing Markets

School districts in the United States are funded by a mix of local, state, and federal revenue. In 2014, States and localities provide 46% and 45% of total public school revenues, respectively, with federal spending contributing the final 9%. State and federal transfers are generally redistributive in nature. At the state level, movements to reduce inequality in district resources gained traction in the 1970s and accelerated after a series of court cases in the 1990s. Hoxby (2001), Jackson, Johnson, and Persico (2016) and Lafortune, Rothstein, and Schanzenbach (2017) provide analyses and more detailed overviews of these reforms.

It is important to note that the distinction between state and local revenues is not always clear, due to the complexities of state revenue-sharing policies. Hoxby (1996) highlights the importance of distinguishing between the entity that collects revenue – an accounting concept – and the entity that decides how to spend it. For example, California has a system in which school districts collect taxes locally even though revenue rules are determined almost entirely by the state.

Nonetheless, property taxes are the dominant source of local revenue, accounting for 81% of the total. Our empirical analysis focuses on districts with independent taxing authority, i.e. those with the power to levy taxes in order to fund local schools. Mechanisms for selecting property tax rates vary by jurisdiction. Annual budgets, with associated tax rates, are proposed and administered by district officials, and, in some cases, must be approved by voters. District officials have varying levels of accountability to their residents; superintendents and schoolboard members may be directly elected, appointed by other political officers, or a mix of both. In certain cases, citizens may directly vote on school spending measures (Cellini, Ferreira, and Rothstein 2010).

Regardless of the variation in accountability measures and tax rules, households are free to “vote with their feet” by moving to another district if local tax and spending policies stray too far from the household’s preferences. This intuition underlies the Tiebout (1956) model and the extensive literature which follows.⁹ Note that in Tiebout’s original model, districts use head taxes rather than property taxes to screen residents, but in practice, districts cannot use head

⁹ Examples include Epple and Sieg (1999), Fernandez and Rogerson (2001), Hilber and Mayer (2001), Epple Romano and Sieg (2012), and Calabrese, Epple, and Romano (2012).

taxes and instead raise most of their revenue from property taxes. Hamilton (1975) notes that local jurisdictions can still achieve efficient sorting and expenditure policies by combining property taxes with zoning. Lot size restrictions establish a minimum house price in each jurisdiction, mimicking the screening mechanism of Tiebout's head tax.

This class of models has generated significant debate over the proper interpretation of the relationship between local house prices, taxes, and public goods. One point of view – often referred to as the “benefit view” – emphasizes the across-district relationship between taxes and public goods characteristic of the Tiebout/Hamilton tradition. Taxes reflect the price of local public goods, and in the process the screen out households with low willingness-to-pay for these amenities. Thus, the costs of higher taxes are efficiently balanced against residents' valuations of local public goods.

Many other papers qualify this interpretation. Barseghyan and Coate (2016) highlight issues that arise when zoning restrictions – which affect only new construction – are selected by incumbent residents. Banzhaf and Mangum (2017) emphasize that capitalization can take the form of both fixed access costs, a la Hamilton (1975), and an increase in the per-unit cost of housing. When taxes affect the marginal cost of housing services, they also create a consumption inefficiency. Hilber (2017) and Banzhaf and Mangum (2017) provide useful overviews of theoretical and empirical work on this question.

While these models vary in their description of the policy levers available to local governments, they almost uniformly treat house prices based on market-clearing conditions in the housing market. Empirically, however, one of the most salient features of housing markets are strong boom-and-bust cycles, which are difficult to generate in models in which prices depend solely on fundamentals. Glaeser and Nathanson (2015) review models that allow prices to depart from fundamentals, for reasons such as uncertainty about long-run supply, limited rationality, search-and-matching frictions, and lapses in credit standards. Housing disease starts with these departures from competitive equilibrium prices. More precisely, we use the term to refer to the influence of exogenous price increases – i.e. those unrelated to local fundamentals like amenities and productivity – on school district revenues and expenditures.

Some mathematical notation may help clarify how our mechanism departs from standard models. Suppose school district leaders choose both the level of total education expenditure E

and the local property tax rate τ to maximize a value function that increases in expenditure E and decreases in the tax burdens imposed on the local citizenry. Let \mathbf{T} denote the vector of household tax burdens, defined by $T_i = \tau P h_i$, where P is the price of housing and h_i is household i 's housing consumption. Hence, letting H denote the stock of housing, the district solves the following program:

$$\max_{E, \tau} V(E, \mathbf{T}) \text{ subject to } E \leq \tau P H$$

In standard models, the tax rate τ can be frictionlessly adjusted each period.¹⁰ Optimal taxes and expenditures are then determined by an equimarginality condition: taxes are increased until the marginal cost of raising revenue equals the marginal benefit of additional expenditure.¹¹ Suppose now that the district experiences an unexpected housing boom – an increase in P in our framework. If tax rates can be costlessly adjusted, the district can restore the initial allocation by a proportional reduction in the property tax rate. Expenditure and each resident's tax burden is unchanged.¹²

For the reasons discussed above, however, changing tax rates can be a costly process. School district administrators may also hesitate to change tax rules because they may not be able to distinguish housing disease from other mechanisms that produce increases in prices and revenues, such as gentrification or local productivity shocks. Suppose that the district has set E and τ in expectation of a certain price level P . If, after choices are codified, the district learns that prices are actually higher, then revenues will exceed expectations and must be spent (in part because many states and districts have rules that prevent schools from keeping large amounts of rainy day funds). Since the policy variables were chosen to equate marginal costs with marginal benefits, the additional spending induces costs that exceed its benefits. This inefficiency is the

¹⁰ To guarantee a unique solution, we also assume that $V()$ is twice continuously differentiable, strictly concave, and obeys standard Inada conditions ($V_e(0, \cdot) = \infty$, $V_e(\infty, \cdot) = 0$, $V_{T_i}(\cdot, 0) = 0$, and $V_{T_i}(\cdot, \infty) = -\infty$).

¹¹ This is obviously an indirect formulation of the district's decision, rather than a full micro-foundation of the political-economic equilibrium. Instead of taking a stand on the district preferences, resident preferences, and the political process that leads to an equilibrium, we use a general value function that captures the key intuitions.

¹² Note that we are implicitly assuming away several effects that may be important in practice. First, we assume that the population of the town is fixed. While this is perhaps a justifiable assumption in the short and medium term – especially if we think local decision makers place more weight on current residents than potential new residents – it ignores the sorting mechanisms underlying Tiebout models. Second, we assume that expenditures and taxes do not influence prices or quantities, a mechanism emphasized by Hoxby (2001). The setup here can readily accompany such pass-through effects, but they distract from our main point. Finally, by placing tax burdens directly into the value function (rather than, say, citizens' after-tax income) we can ignore the direct effect of the price increase on citizen purchasing power.

cost of housing disease. Of course, if district spending was inefficiently low prior to the housing shock (because of frictions such as state level regulations described in Cellini, Ferreira, and Rothstein (2010)), then housing disease may actually improve efficiency.

This simple set-up assumes away several possible uses for the windfall that deserve discussion. First, our specification of the objective function depends only on total expenditures. In practice, there are many potential sources of educational spending, and district officials could allocate the windfall to sources that benefit them personally, such as their own salaries. Diamond (2017) and Brueckner and Newmark (2014) provide evidence that local officials sometimes use their positions to extract rents in this manner. This effect is more likely if voters pay less attention to tax revenue increases that result from unexpected windfalls as opposed to politically salient increases in rates. We explicitly test for the presence of this type of rent-seeking in our empirical work.

Alternatively, district leaders could save the increased revenues and return them to voters in subsequent years via lower taxes. In some cases, however, districts may have explicit incentives to avoid this behavior, as unspent funds may crowd out future transfers. To account for a full range of possible dynamic effects, our empirical specifications allow prices and expenditures to evolve flexibly over a period of five years following a housing boom. Before turning to our empirical specification though, the next section reviews the school and housing data.

III. Data

School District Data

Our primary data source for school district finances is the School District Finance Survey (often referred to as the F-33 survey), which the National Center for Education Statistics (NCES) has administered annually since 1995. The datasets report detailed revenue and expenditure categories for all school districts in the United States.¹³ School district boundaries are not constant over time, as districts merge and split with some regularity. We contacted all state

¹³ The survey also includes charter school operators, which we do not include in any part of our analysis.

education agencies to request details of the history of district boundary changes. Ultimately we received this information from 36 states, allowing us to create constant-boundary district definitions for most of our sample. We restrict our final analysis sample to districts that have independent taxing authority, “unified” districts that include both elementary and high school students, and districts that never merged or split during that time period. However, we also show that our results are robust to relaxing these restrictions.¹⁴

We supplement the revenue and expenditure data with demographic and staffing information from the District and School Universe Surveys, part of the NCES’ Common Core of Data. These datasets provide a several useful descriptors for our analysis. First, they report the racial background of enrolled students and the number of students eligible for free or reduced-price lunches. These measures allow us to check whether changes in local housing prices might reflect changes in the composition of local students or residents. The files also provide detailed staffing information, which we use to construct measures of average salaries and employment levels for various categories of workers.

Finally, we obtained microdata from the National Assessment of Educational Progress (NAEP) to assess whether changes in spending translated into short-term changes in student achievement. We make use of the State NAEP sample, which contains scores from a national, consistently-normed test administered biannually to a randomly selected subset of students in participating states.¹⁵ We average student scores to the district-year-test level to construct a summary measure of student performance. More precisely, we use NAEP’s reported “plausible values” in lieu of raw test scores, which are not included in the microdata. See Lafortune et al. (forthcoming) and Jacob and Rothstein (2016) for useful discussions of the possible biases that may arise when using model-derived measures of student ability in external analyses.¹⁶ Another limitation of the NAEP is limited coverage in early parts of the sample. Between 1996 and 2002, each biennial testing cycle offered only math or reading – never both. Furthermore, participation

¹⁴ We use per-pupil expenditure and revenue measures throughout our analysis. One shortcoming of the NCES data is that it records “snapshot” enrollment as of October 1st of each schoolyear, which may not reflect district size as accurately as other measures, such as average daily attendance. We are unaware of an annual, national dataset that records districts’ average daily attendance or a similar measure, however.

¹⁵ We are grateful to Julien Lafortune for providing code to link the NAEP microdata to NCES district identifiers.

¹⁶ Fortunately, our results are virtually identical when using NAEP plausible values, ability measures estimated from item-response models, or residualized versions of these measures that control for individual student demographics, suggesting that such biases are not likely to be an important factor in our results.

was optional, and between 41 and 45 states participated in each year during this period. Participation has been mandatory since 2002, however, and the change in sample composition likely explains the sudden change in math scores apparent in Appendix Figure 1A.

Housing Transactions Data

Our house price data come from CoreLogic, a private data vendor that aggregates public deeds records from county recorder's offices in markets across the country. Houses are pre-assigned to their Census block group, which we then match to school district boundaries using Census block relationship files.

We focus attention on districts with sufficient data to at least calculate a continuous quarterly price series between 2000:Q1 and 2007:Q1 (we use data from outside of this time period when it is available).¹⁷ The resulting dataset includes 2,785 school districts and over 28 million transactions. To eliminate bias from specification search (Leamer 1983), we randomly split the sample in half and compute constant-quality hedonic price indices for each sample independently.¹⁸ One sample is used to identify and test for the existence of structural breaks, and the other is used to estimate how prices change in the periods surrounding the break.

Figure 2A plots the 90th, 50th, and 10th percentiles of the resulting district-level price indices. The boom period of the recent cycle is apparent at each part in the distribution. Nevertheless, the magnitude of the bust varied tremendously. Even though we normalize each index to 100 in 2010:Q1, there is considerable variation at the peak just five years earlier.

Figure 2B plots annual growth rates of the same series. To remove the effects of seasonality in the housing market, we calculate growth rates as year-over-year changes in the quarterly series, i.e. $(P_t - P_{t-4})/P_{t-4}$. While the national housing bust starting in early 2005 is immediately apparent, there is no visual evidence of a sudden break during the previous boom

¹⁷Specifically, we only include districts that report at least 16 observations in all quarters during this period, though we also include periods outside of this window

¹⁸We estimate hedonic models because their data requirements are much less stringent than repeat-sales methods, particularly when working with small geographies. In practice, hedonic and repeat-sales estimates are very similar when both are computationally feasible. We construct our hedonic indices by regressing log prices on the square footage of the home (and its square), the number of bedrooms, the number of bathrooms, the age of the home, and an indicator for condominiums. Ferreira and Gyourko (2011) and DeFusco et al (2017) show that this model closely approximates the Case-Shiller index when estimated at the MSA level.

period. This fact is essential to our identification strategy. While most markets experienced a sudden onset of rapid growth, there is considerable cross-sectional variation in the timing of the booms.

Sample Restrictions and Representativeness

Table 1 reports some basic summary statistics and demonstrates how the sample composition changes as we add restrictions. The first column reports summary statistics for the entire sample of school districts in the F-33 dataset. Moving to the right, we add restrictions one by one until arriving at our main regression sample in column (5). The final column summarizes data for districts in the regression sample that we are able to match to test score data.

The most stringent sample restriction is the availability of historical housing transactions data. While the CoreLogic sample covers more than 90% of U.S. counties in 2016, we require sufficient transaction volume to estimate quarterly price indices starting no later than the year 2000. Hence, the merge to the housing sample immediately reduces our sample by 80%. Unsurprisingly, the districts that survive the merge to the housing data tend to be larger than the national average; enrollment in the breakpoint sample (10,221 students per district) is nearly three times that of the average district (3,459), corresponding to almost 60% of the total enrollment in public schools. These districts also have larger minority populations, higher student teacher ratios, and greater portions of the population eligible for free or reduced-price lunch, an indicator of family income. Somewhat reassuringly, revenue per pupil is similar in the housing sample (\$11,047/student) as in the overall sample (\$11,158/student).

Columns (3) through (5) show the effects of restricting the sample to unified districts only (as opposed to districts specific to elementary schools or high schools); districts with independent taxing authority; and districts with constant borders and no missing financial data over our sample period. Enrollment, average revenue, student teacher ratios, and average demographics are largely unaffected by these restrictions. Our favorite sample is based on Column 5, and it represents 42% of all public school students.

IV. Empirical Framework and Validity of Research Design

Identifying Structural Breaks and Estimating Magnitudes

Glaeser et al. (2014) provide the motivation and micro-foundations for the existence of structural breaks in housing prices. In their model, house prices grow at a constant rate in the steady-state. However, the introduction of a shock to the local economy – e.g. a demand shifter or a change in expectations – leads to a discrete jump in the growth rate as the local housing market converges to a new equilibrium. This insight has led to a recent empirical literature exploiting these sharp changes to understand how changes in house values affect other economic variables (Ferreira and Gyourko (2011), DeFusco et al. (2017), Charles, Hurst, and Notowidigdo (2015)). Because we closely follow the breakpoint identification and inference methods described in Ferreira and Gyourko (2011) and DeFusco et al. (2017), we sketch an outline of these procedures here and relegate many of the details to the Appendix.

First, consider the problem of testing for the existence of a single structural break. Denoting the house price growth rate in district i at time t as $d_{i,t}$, the null hypothesis of no structural break is:

$$(1) H_0: d_{i,t} = d_{i,0}, t = 1, \dots, T$$

The alternative hypothesis is that the growth rate changes in the middle of the sample, at a time period t^* , i.e.:

$$(2) H_1: d_{i,t} = \begin{cases} d_{1,i}(t^*), & t = 1, \dots, t^* \\ d_{2,i}(t^*), & t = t^* + 1, \dots, T \end{cases}$$

The first step of our analysis is to identify the value of t^* that minimizes the residual variation in growth rates. We implement this by searching over all values of t' in each districts' price growth series,¹⁹ estimating a regression model with separate intercepts for the pre- and post- t' periods, and selecting the candidate time period that produces the smallest sum of squared residuals.

Of course, this procedure will select a candidate breakpoint regardless of whether a break exists, and some care needs to be taken when constructing tests for the existence of a structural

¹⁹ The endpoints of our series are data-dependent. For each district, the first period is the earliest quarter featuring at least 16 transactions, with a hard minimum of 1993:Q1 to focus attention on the most recent cycle. The final period is the pre-2009 peak of the price level, though our results are robust to capping the series in 2007 for all districts. We do not allow breakpoints to lie in the first two or final two periods of the series.

break. If t^* were known *a priori*, we could test H_1 against H_0 using standard methods. Because we select the break that maximizes the likelihood ratio, however, critical values for testing must be derived from the distribution of the supremum of the likelihood ratio statistic (under the null hypothesis of no break). Andrews (1993) and Bai (1997) derive exact formulas for this distribution, and Estrella (2003) describes numerical methods to calculate p-values efficiently.

Ultimately, we allow for up to three structural breaks in the price growth series for each district. Bai (1999) and Bai and Perron (1998) derive tests for the existence of $b+1$ structural breaks against the null hypothesis of b breaks. Therefore, we test for a second break whenever we detect a first break at the 5% significance level, and a third break whenever we identify a significant second break. This recursive testing procedure is valid because, as shown by Bai (1999) and Bai and Perron (1998), the one-break test remains valid when multiple breaks exist. We identify candidate breakpoints in multiple-break models by looping over all possible pairs (or triples) of breaks in a districts' price growth series.

It is also important to note that the regressions used to identify breakpoint locations do not provide unbiased estimates of the significance and magnitude of the change in price growth rates at the breakpoint. This is due to the specification search issue identified by Leamer (1983), in which the data-dependent manner by which we identify breakpoints contributes to a bias in estimating the magnitude of the break. We address this issue via the split-sample approach suggested by Card, Mas, and Rothstein (2008). That is, we randomly split the dataset in half, and use one sample to estimate the breakpoints and the other to estimate the price response.

We run variants of the panel equation (3) below in order to estimate the magnitude of changes in price (and also for a number of other school district outcomes) along the housing boom. Denote $Y_{i,t}$ the log of the house price index in district i and year-quarter t , $t_{i,b}^*$ the quarter of the b^{th} breakpoint in a district, and B_i the number of breakpoints estimated for district i :

$$(3) Y_{i,t} = \alpha_i + \kappa_t + \sum_{b=1}^{B_i} \sum_{\substack{\rho=-6 \\ \rho \neq 0}}^6 \theta^\rho \mathbf{1}[t - t_{i,b}^* = \rho] + \varepsilon_{i,t}$$

where α_i and κ_t are district and time fixed effects, respectively.

This parameterization allows for flexible dynamics in the break's effects. Each θ^ρ measures the change in the outcome variable ρ years after the break, relative to the year immediately prior to the break (note that we omit the dummy variable for relative year zero.)

Negative values of ρ target the “effects” of future breaks, allowing us to test for the existence of pre-trends that might confound our research design. The controls included in panel equation (3) guarantee that the housing boom effects will be estimated net of calendar effects, school district fixed effects, and also net of other booms and busts that happened in the same district.

In the same specification we estimate separate effects for positive breaks, non-significant breaks, and negative breaks, as we are primarily interested in understanding the effects of sudden booms – i.e. positive structural breaks. Even though all empirical specifications will estimate the effect of housing busts, the validity of such estimates are less credible since many markets begin to decline at essentially the same time, complicating efforts to separate the effects of bust-induced price variation from the national macroeconomic downturn.

Breakpoint Results and Validity of Research Design

For illustrative purposes, each panel of Figure 3 plots price growth rates for four districts, with estimated breakpoints marked in red. The top left panel shows an example of a school district with only one positive and statistically significant breakpoint, which we call a boom. The top right panel has a district with two statistically significant breaks. The bottom left panel has a district with three booms in one district, and finally, the bottom right panel shows the example of a district with one break that is not statistically different from zero. Those examples make the obvious point that the number of breaks we detect depends both on severity of the change in trend as well as the level of idiosyncratic variance in the series.

The three panels of Figure 4 show the full distributions of breakpoint timing for positive breaks, negative breaks, and non-significant breaks. Crucially for our identification strategy, the positive breaks are well distributed between 1998 and 2005. Cross-sectional variation in the timing of housing booms allows us to separate shocks to the local housing market from national trends and changes to the macroeconomy. Negative breaks, on the other hand, are concentrated largely during the onsets of economic downturns in 2001 and 2006. Overall, the 1,725 district time series in our favorite regression sample produce 1,107 booms, 541 busts, and 405 non-significant breaks.

Figure 5A then shows that school district housing booms are not preceded by changes in total expenditures per pupil, pupil-teacher ratios, and mathematics and reading test scores. That is not a surprise given that quality of school amenities are not part of the list of causes of the housing boom. Figure 5B then turns to the demographic composition of school districts. First, there is no evidence of changes in racial composition around booms. Second, while it appears that use of free lunch is lower in the post-boom period, the magnitude of the change is quite small compared to the size of the price effect. To confirm that shifts in demographics are not driving our results, in the next section we report results from models that control for %white, %black, %Hispanic, and % free lunch as a robustness check. Their inclusion does not impact the estimation of the house price elasticity of expenditures per pupil. We also discuss possible mechanisms through which booms could alter unobserved demographic composition at the end of Section 5.

V. Results

House Prices and School Expenditures

The first three columns of Table 2 report how house prices evolved after the start of a school district housing boom, bust, or non-significant breakpoint. Prices jump 4.8% in the first year of a boom, and keep growing in the following years, reaching 20.1% above the baseline in relative year 5. Busts have a symmetric result with cumulative price reductions of 12.0% by relative year 5. Districts that did not boom or bust had negligible price increases.

Estimates for expenditures per pupil are shown in Columns 4, 5 and 6. Expenditures start to creep up in the second year of a housing boom, become statistically significant in year 3, and reach a peak of 3.3% in relative year 4. Busts again have a mirrored pattern of reductions in expenditures. None of the estimates are significant for school districts with non-significant breaks.

Figure 6 plots the impact of local housing booms on prices and expenditures together. Both show no trends prior to the beginning of the boom. But while prices immediately respond to the beginning of a boom, expenditures respond with a lag – matching the institutional features

of school district finances discussed in Section II. Finally, the magnitude of the price effect is an order of magnitude higher than the expenditure effect.

Table 3 explores a number of robustness tests. Column 1 shows our preferred estimates again to facilitate comparisons. Column 2 includes the full sample of school districts in our data, prior to restricting the sample to independent unified school districts that never experienced a split or a merge and that possess a complete panel of finance data.²⁰ The path of the coefficients is similar, but the point estimates are about 20% smaller - which is not surprising given the non-consistent sample. Column 3 then excludes non-independent school districts from the full sample, and the resulting point estimates for expenditures per pupil become slightly larger. Column 4 trims outliers in our preferred sample by excluding districts with expenditure growth rates in the top or bottom 1% of the sample. These estimates are only slightly smaller for house prices and similar for expenditures per pupil.

Column 5 only uses the one-breakpoint model. Estimates are equally larger for both prices and expenditures. The intuition for this result is that such model does not control for a second or third break, and therefore the magnitude of boom is loaded into the one break. Finally, Column 6 uses our original specification with the addition of school demographics. Estimates are practically unchanged, which corroborates the validity of the research design.

House Price Elasticity of Expenditures Per Pupil

In this section we back out the house price elasticity of expenditures per pupil. One complication is that it is difficult to pin down the precise lag structure for these elasticities given the heterogeneity in school finance structures in the United States. We therefore present results from two types of Wald estimator. One divides the point estimates of expenditures per pupil in time t by the price effect in time $t-1$ (the lagged price elasticity) and one that divides the expenditure coefficient by the price coefficient from the same period (the concurrent price elasticity). Standard errors are calculated via the delta method.

²⁰ We have estimated all results in this paper using the full sample of districts that we match to our housing dataset, and our findings are unaffected. The expenditure and revenue coefficients decrease slightly, as one would expect when many districts without independent taxing authority are added to the sample.

The first row of Table 4 shows the estimated elasticities for each relative year. The estimates are remarkably stable, ranging from 0.16 to 0.20. The last column shows the estimate for a specification that bunches relative years three through five, producing a weighted average elasticity of 0.18. The next row uses concurrent estimates as opposed to the lagged structure. These concurrent elasticities are slightly smaller, with a weighted average of 0.16. The table does not report the elasticities for the busts, but our estimates show a number that is larger than the ones for the boom but imprecisely estimated (the pooled elasticity estimate is 0.26 (0.14)). One possible reason for the larger elasticity is that, as we mentioned before, the busts in our sample are bunched in the onset of recessions, and therefore those results might be confounded by other factors, such as drops in employment and wages.²¹

Next we investigate if there is heterogeneity in these elasticities. First we create indicators for districts that were below and above the median expenditure per pupil in 1996, and then fully interact them with the relative year dummies. We run these models for prices and expenditures and calculate elasticities that are reported in the last two rows of Table 4. Although we have a relatively large sample of districts, it is not sufficient to produce heterogeneity estimates that are statistically different from each other. However, the pattern of the point estimates is suggestive: school districts with above median initial expenditures per pupil have a larger elasticity than the below median districts.

These results match a couple of important features of the American school finance system: school districts receive a large fraction of their revenues from state and federal transfers, and those transfers are disproportionately more relevant to low expenditure districts. In this setting, average elasticities should be relatively small, and high expenditure districts should have higher elasticities.

With the elasticities in hand we can back out by how much housing disease impacted the rise in public education spending in the United States during the 1990s and 2000s. The main assumption needed for this exercise is that the estimated elasticities can be applied to all price changes, not just the price changes from the variation used in our research design. While this

²¹ Our estimated elasticities are somewhat smaller than existing estimates of the property-tax elasticity for cities and states. Lutz (2008) estimates a value of 0.4 using national and state level time series analysis, while Vlaicu and Whalley (2011) find a 0.74 elasticity for California cities using an instrumental variable constructed from housing supply constraints.

might seem like a strong assumption, the sample period is characterized by little changes in real wages and incomes. While there is still an ongoing debate about the causes of the last housing boom (i.e., changes in credit supply, changes in house price expectations, or a combination of both) the current consensus is that a small part of the cycle was due to real changes in fundamentals. Finally, we also assume no general equilibrium consequences arising from the initial changes in prices, which is consistent with the lack of changes in demographics observed in Figure 5 (we will discuss changes in unobserved demographic features in more detail in the next subsection).

The underlying data from Figure 2 shows that school districts had an average house price increase of 95.17% from 1995 to 2007 (right before the Great Recession). Multiplying that number by the 0.18 elasticity gives a change in expenditure per pupil of 17.13%. That corresponds to about half of the observed change in average expenditures per pupil from 1996 to 2008, implying that housing disease was the most important determinant of school finances during that period. The main driver of this effect is the unprecedented increase in house prices, a cycle never before seen in the United States (Shiller, 2005).

We also calculate heterogeneity by using the price changes from the bottom and top of the distribution (P90 and P10), and applying the below and above median expenditure heterogeneity in elasticities reported in Table 4. Housing disease can only account for 20% of the expansion of expenditures in below median expenditure districts, but can explain 70% of that increase in above median districts. Again, this reflects the fact that low expenditure districts are much more dependent on state and federal transfers and the fact that housing booms were much larger at the top of the distribution.

School District Revenues

One caveat with our empirical approach is that adjustments in tax rates and other local rules and regulations are not observed in the data. If districts reduce tax rates after the start of a housing boom, then we underestimate the elasticity - but can still interpret the results as a combination of the direct price effect plus the indirect political effect of potential adjustments in tax rates. The school district revenue data do not help solving this problem because of three issues: a) it only reports total revenues as opposed to a breakdown of tax base and tax rates; b)

even the breakdown by local versus state or federal transfer is muddled because it is difficult to disentangle the role of the school district as tax collector versus who in fact has control of the tax resources (Hoxby 1996); c) the revenue data is noisier than the expenditure data because of reporting standards. For example, revenues for capital projects that invest (spend) resources for 5 or 7 years are fully recorded in the first year of the project. A similar phenomenon occurs with private donations.

Further complications arise from state policies that either restrict districts' taxing ability or redistribute revenues. Such policies are quite common; see Hoxby (2001) and Jackson, Johnson, and Persico (2016), who carefully track court cases and state legislation to evaluate the impacts of state policy changes. We are primarily interested in how such policies might mediate housing disease, not the overall impact of these policies. Accordingly, we need only focus on aspects of state formulas that respond to *changes* in the local property tax base. Note that many common formula features, such as foundation formulas or equalization policies, are not directly affected by house price growth, so their impacts are therefore absorbed by district fixed effects.

Therefore, we focus attention on state policies that restrict the growth of local property taxes by placing explicit limits on property tax growth, either by capping growth in assessments or capping revenue growth directly. We draw our classifications from Hightower, Mitani, and Swanson (2010), who surveyed all 50 states and categorized funding formulas along various dimensions.²² In light of these issues, Table 5 reports magnitude estimates for total revenues and for revenue subcategories (local, state, and federal) in states with and without property tax growth caps. Total revenue per pupil follows a similar path observed for expenditures per pupil, albeit with slightly smaller point estimates. As one would expect, local revenues respond to housing booms in uncapped states only. When property tax increases are restricted, housing booms produce small and statistically insignificant effects. State revenues show the opposite pattern: zero effect in uncapped states and positive effects in capped states. The increase in state revenue in areas with local growth caps is likely due to the fact that over the years booms spread out following a geographic pattern (moving from both coasts to the middle of the country), as documented by DeFusco et al. (2017).

²² We are omitting one formula characteristic that is likely relevant: district spending caps. While not directly tied to growth, when the constraint binds – as they frequently do, in practice – they eliminate the relationship between house prices and revenue.

Type of School Expenditures and Quality

How are the additional resources arising from housing disease spent by school districts? Reported school expenditures are split into three main categories: current (corresponding to 84.7% of the total expenditures during the sample period), capital (10.5%) and others (4.8%).²³ Columns 1 and 2, of Table 6 report estimates for capital and current expenditures. Current expenditures show a similar pattern of lagged increases observed for total expenditures, albeit with smaller point estimates. Capital expenditures have a much larger effect, increasing 17.5% above the baseline in the fourth year of the housing boom.

Next we split current expenditure into its two key categories, instruction and services, and present the estimates in Columns 3 and 4.²⁴ Their estimated effects are quite similar to each other, and also similar to the estimates for total current expenses. Finally, in Columns 5, 6, and 7 we break down the service component into instruction, pupil, and administration.²⁵ While instruction and pupil services show large effects (above 3% in some years), administrative cost point estimates are very small.

Given that instruction expenditures correspond to the largest expenditures in a school district, in Table 7 we test if those increases are due to increases in the quantity of teachers – which reduces pupil-teacher ratios – or to raises in wages and benefits paid to the teachers. We find a mix of both. There are small decreases in average salaries and benefits immediately following the boom, with large increases in years four and five. Pupil-teacher ratios decline moderately after a boom. Ultimately housing disease increases wages by raising teacher compensation, which constitutes the vast majority of district labor expenses. We also report separate effects for instructional salaries, administrative salaries, and other salaries in Appendix

²³ Other types of expenditure include interest on debt and payments to other governments or school systems.

²⁴ Instruction accounts for 60.9% of current expenditures most of which is teacher salaries and benefits (though instructional aides are also included in this category). Services are 33.8%. Examples of service employees include support, administrative, operations, transportation, and business staff.

²⁵ Instructional services are expenses related to instruction that do not involve interaction between students and teachers in the classroom; examples include staff training, curriculum development, and technological services. Pupil support includes administrative, guidance, health, and logistical expenditures, such as counseling, speech therapy, and record maintenance. Administrative services include operations associated with the district office or the office of the school principal.

Table 2, though our construction of these variables requires significant caveats.²⁶ We find that teacher salaries improve steadily after a boom, reaching a 2% increase by relative year 5, while average administrator salaries are 5.6 to 6.7% lower in the four years following a boom. While this negative estimate could be due to noisy data, it corroborates the result that administrative costs are not increasing with housing disease.

Table 8 then presents point estimates for overall test scores for math and reading in columns 1 and 2, and separately for 4th and 8th grade test scores in columns 3 through 6. Because the NAEP test is only administered every two years, we pool relative-year coefficients into groups of two. The estimates are noisy, in part because test scores are never available for 15% of our main regression sample and inconsistently available for other districts. We see no significant effects on math scores in any specification. There is some evidence that reading scores increased in the very long run. We estimate that reading scores increased by 0.090 standard deviations 5-6 years after a boom and 0.099 standard deviations 7-8 years post-boom.

We are hesitant to over-interpret the reading results for several reasons. First, the effects enter with a substantial lag. The estimates are driven by observations long after the boom, placing significant strain on our identification strategy. The extended lag also creates an unbalanced panel; we only observe five post-boom years for districts with positive breaks relatively early in the sample, which are observably different from the late-breaking areas. Furthermore, it is noteworthy that we do not observe a similar increase in math scores. We are unaware of any reason to expect reading scores to respond more strongly to increased expenditures than math scores. In fact, generally speaking, math scores are more responsive to educational intervention than reading scores for school-age children (Fryer 2017). Finally, as explained in the data section, NAEP performance data is based on plausible value predictions of individual test scores, as opposed to the raw tests scores per se.

General Equilibrium Consequences

²⁶ We calculate average salaries by dividing total spending on salaries (obtained from the F-33 Finance file) by the number of employees (obtained from the Common Core of Data survey file). Unfortunately, these two datasets do not group employees into consistent categories, so we aggregate up to the broad groupings described here. Mapping the categories to a common definition nonetheless requires some guesswork. To reduce the influence of misclassification errors, we drop districts with fewer than ten employees in a given category, since errors in employee counts are most harmful in small samples.

Our results assumed that sorting based on unobservables were not driving the estimates, and we corroborated this assumption by looking at how observed school demographics changed around the timeline of local housing booms. However, one could posit that housing booms may induce a higher share of high income families or families with more school-age children to move to better school districts (or districts with higher expenditures per pupil) just because of budget constraints. The mechanism is simple: households with higher unobserved willingness to pay for those school districts will win the bidding war for the limited supply of homes in those districts. The MSA level results shown by Ferreira and Gyourko (2011) seem to potentially corroborate such a story: even though overall income barely changes at MSA level after the beginning of a housing boom, the self-reported income of marginal homebuyers seem to have increased in the first two years (and then declined again to baseline levels).

Such effects, though, may only have a trivial impact on the overall composition of households in a school district because only a small fraction of homeowners move every year. Moreover, new households likely have very little influence in the local political decisions of school boards given that they are mostly newcomers. To the extent that new households do affect tax and spending decisions, one could consider these changes part of the housing disease effect. Notably, if housing disease consists of both direct price effects and any secondary effects on policy, our point estimates are still consistent, though they have a more reduced-form interpretation.

Our discussion also implicitly assumes that housing wealth effects do not operate – that is, the increase in house prices does not cause households to demand higher education expenditures. In theory wealth effects should not occur in this setting because housing consumption remains constant: homeowners would have to sell their current house to tap the new wealth, but the cost of buying a new similar home would completely offset the gains from the previous sale. Behavioral factors could generate some type of illusory wealth effect, but results in the finance literature (e.g. DeFusco (forthcoming)) suggest that such wealth effects are small in practice, and therefore unlikely to drive our results.

Finally, another general equilibrium consequence of housing disease is dependent on the degree of inefficiency of the new expenditure levels. High levels of inefficient spending should

lead to lower future house prices and a reduction in expenditures. Those second order effects may happen with even longer lags though, making its estimation not suitable in our setting.

VI. Conclusion

Both housing prices and educational spending rose dramatically in the 1990s and 2000s. Traditional public finance theory views public school districts as a set of local jurisdictions that provide different degrees of school quality, and access to those benefits is capitalized into house prices. This paper shows that the reverse causal channel should not be ignored: house price increases lead to additional spending per pupil by increasing the local tax base, and local administrators have incentives to spend those extra funds. We refer to this phenomenon as housing disease, as the increase in expenditures comes from a housing market spillover rather than a political decision weighing the benefits of school spending against the costs of increased tax burdens.

The magnitude of the estimated effect is substantial: we estimate house price elasticities of per-pupil expenditures of 0.16-0.20, implying that rising house prices can explain roughly half of the increase in per-pupil expenditures leading up to the Great Recession. Although housing disease is a source of inefficiency in local finances, we find that the spending increases are concentrated on student instruction and not administrator salaries, suggesting that improvements in school quality may have accompanied the increase in school expenditures.

Even after the widespread growth of state and federal revenue sharing rules in the past decades, our results show that district finances are still influenced by local housing conditions. Since there is little reason to believe that housing cycles are disappearing, housing disease will remain a relevant feature of the American landscape. It may even grow in importance, as long as local communities have the power to constrain new housing development through zoning rules. Those regulations not only magnify the housing affordability problem in the United States, but also increase the cost of local services via housing disease. In fact, an interesting area of future work relates to how individuals within a district are bearing the incidence of housing disease, and how jurisdictions interested in reducing localities' exposure to price shocks should alter their taxing framework.

References

- Andrews, Donald W. K. 1993. "Tests for Parameter Instability and Structural Change with Unknown Change Point." *Econometrica* 61 (4): 821-56.
- Bai, Jushan. 1997. "Estimating Multiple Breaks One at a Time." *Econometric Theory* 13 (3): 315-52.
- Bai, Jushan and Pierre Perron. 1998. "Estimating and Testing Linear Models with Multiple Structural Changes." *Econometrica* 66 (1): 47-78.
- Bai, Jushan. 1999. "Likelihood Ratio Tests for Multiple Structural Changes." *Journal of Econometrics* 91 (2): 299-323.
- Banzhaf, H. Spencer and Kyle Mangum. 2017. "Capitalization as a Two-Part Tariff: The Role of Zoning." *Mimeo*
- Barrow, Lisa, and Cecilia Elena Rouse. 2004. "Using Market Valuation to Assess Public School Spending." *Journal of Public Economics* 88 (9): 1747-69.
- Baumol, William J., and William G. Bowen. 1966. *Performing Arts, the Economic Dilemma; a Study of Problems Common to Theater, Opera, Music and Dance*. New York: Twentieth Century Fund.
- Barseghyan, Levon, and Stephen Coate. 2016. "Property Taxation, Zoning, and Efficiency in a Dynamic Tiebout Model." *American Economic Journal: Economic Policy* 8(3): 1-38.
- Bayer, Patrick, Fernando Ferreira, and Robert McMillan. 2007. "A Unified Framework for Measuring Preferences for Schools and Neighborhoods." *Journal of Political Economy* 115 (4): 588-638.
- Bradbury, Katharine L., Christopher J. Mayer, and Karl E. Case. 2001. "Property Tax Limits, Local Fiscal Behavior, and Property Values: Evidence from Massachusetts under Proposition 2 1 2." *Journal of Public Economics* 80 (2): 287-311.
- Brueckner, Jan K. 1979. "A Model of Non-Central Production in a Monocentric City." *Journal of Urban Economics* 6 (4): 444-63.
- Brueckner, Jan K., and David Neumark. 2014. "Beaches, Sunshine, and Public Sector Pay: Theory and Evidence on Amenities and Rent Extraction by Government Workers." *American Economic Journal: Economic Policy* 6 (2): 198-230.
- Calabrese, Stephen M., Dennis N. Epple, and Richard E. Romano. 2012. "Inefficiencies from Metropolitan Political and Fiscal Decentralization: Failures of Tiebout Competition." *Review of Economic Studies* 79(3): 1081-111.

Card, David and Alan B. Krueger. 1992. "Does School Quality Matter? Returns to Education and the Characteristics of Public Schools in the United States." *Journal of Political Economy* 100 (1): 1-40.

Card, David, Alexander Mas and Jesse Rothstein. 2008. "Tipping and the Dynamics of Segregation." *The Quarterly Journal of Economics* 123 (1): 177-218.

Card, David, and A. Abigail Payne. 2002. "School Finance Reform, the Distribution of School Spending, and the Distribution of Student Test Scores." *Journal of Public Economics* 83 (1): 49-82.

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

Charles, Kerwin Kofi, Erik Hurst, and Matthew J. Notowidigdo. 2015. "Housing Booms and Busts, Labor Market Opportunities, and College Attendance." National Bureau of Economic Research.

Clark, Damon. 2009. "The Performance and Competitive Effects of School Autonomy." *Journal of Political Economy* 117 (4): 745-783.

Coleman, James S., Ernest Q. Campbell, Carol J. Hobson, James McParland, Alexander M. Modd, Frederic D. Weinfeld, and Robert L. York. 1966. *Equality of Educational Opportunity* Washington, D.C.: U.S. Government Printing Office.

DeFusco, Anthony. Forthcoming. "Homeowner Borrowing and Housing Collateral: New Evidence from Expiring Price Controls." *Journal of Finance*, forthcoming.

DeFusco, Anthony, Wenjie Ding, Fernando Ferreira, and Joseph Gyourko. 2017. "The Role of Contagion in the Last American Housing Cycle." *Mimeo* University of Pennsylvania.

Diamond, R. 2017. "Housing Supply Elasticity and Rent Extraction by State and Local Governments." *American Economic Journal-Economic Policy* 9 (1): 74-111.

Epple, Dennis, and Holger Sieg. 1999. "Estimating Equilibrium Models of Local Jurisdictions." *Journal of Political Economy* 107 (4): 645-81.

Epple, Dennis, Richard Romano, and Holger Sieg. 2012. "The Intergenerational Conflict over the Provision of Public Education." *Journal of Public Economics* 96 (3-4): 255-68.

Estrella, Arturo. 2003. "Critical Values and p Values of Bessel Process Distributions: Computation and Application to Structural Break Tests." *Econometric Theory* 19 (6): 1128-1143.

Favara, Giovanni, and Jean Imbs. 2015. "Credit Supply and the Price of Housing." *American Economic Review* 105(3): 958-92.

Fernández, Raquel, and Richard Rogerson. 2001. "Sorting and Long-Run Inequality." *The Quarterly Journal of Economics* 116 (4): 1305-41.

Ferreira, Fernando, and Joseph Gyourko. 2011. "Anatomy of the Beginning of the Housing Boom: U.S. Neighborhoods and Metropolitan Areas, 1993-2009." National Bureau of Economic Research.

Ferreira, Fernando, and Joseph Gyourko. 2009. "Do Political Parties Matter? Evidence from US Cities." *Quarterly Journal of Economics* 124 (1): 399-422.

Fryer, Roland G. 2017. "Chapter 2: The Production of Human Capital in Developed Countries: Evidence from 196 Randomized Field Experiments." In *Handbook of Field Experiments* Edited by Abhijit Vinayak Banerjee and Esther Duflo, 95-322. Amsterdam: Elsevier B.V.

Glaeser, Edward and Charles Nathanson. 2015. "Housing Bubbles." In *Handbook of Urban Economics*. Edited by Gilles Duranton and Vernon Henderson, 701-751. Amsterdam: Elsevier B.V.

Glaeser, Edward, Joseph Gyourko, Eduardo Morales and Charles Nathanson. 2014. "Housing Dynamics: An Urban Approach." *Journal of Urban Economics* 81:45-56.

Glaeser, Edward, and Joseph Gyourko. Forthcoming. "The Economic Implications of Housing Supply." *Journal of Economic Perspectives*.

Hamilton, Bruce W. 1974. "Zoning and Property Taxation in a System of Local Governments." *Urban Studies* 12 (2): 205-221.

Hansen, Bruce E. 1997. "Approximate Asymptotic P Values for Structural-Change Tests." *Journal of Business and Economic Statistics* 15: 60-67.

Hanushek, Eric A. 1986. "The Economics of Schooling: Production and Efficiency in Public Schools." *Journal of Economic Literature* 24 (3): 1141-1177.

Hanushek, Eric A., and Steven G. Rivkin. 2006. "Chapter 18 Teacher Quality." In *Handbook of the Economics of Education*, edited by Eric A. Hanushek, Stephen Machin and Ludger Woessmann, Vol. 2, 1051-1078. Amsterdam: Elsevier B.V.

Hightower, Amy M., Hajime Mitani, and Christopher B. Swanson. 2010. "State Policies That Pay: A Survey of School Finance Policies and Outcomes." Editorial Projects in Education and Pew Center on the States.

Hilber, Christian A.L. 2017. "The Economic Implications of House Price Capitalization: A Synthesis." *Real Estate Economics* 45 (2): 301-339.

Hilber, Christian A. L., and Christopher Mayer. 2009. "Why Do Households Without Children Support Local Public Schools? Linking House Price Capitalization to School Spending." *Journal of Urban Economics* 65 (1): 74-90.

Hoxby, Caroline M. 1996. "Are Efficiency and Equity in School Finance Substitutes or Complements?" *The Journal of Economic Perspectives* 10 (4): 51-72.

Hoxby, Caroline M. 2000. "Does Competition among Public Schools Benefit Students and Taxpayers?" *American Economic Review* 90 (5): 1209-1238.

Hoxby, Caroline M. 2001. "All School Finance Equalizations Are Not Created Equal." *The Quarterly Journal of Economics* 116 (4): 1189-231.

Hoxby, Caroline M. 2007. "Does Competition among Public Schools Benefit Students and Taxpayers? Reply." *American Economic Review* 97 (5): 2038-2055.

Jackson, C. Kirabo, Rucker C. Johnson, and Claudia Persico. 2016. "The Effects of School Spending on Educational and Economic Outcomes: Evidence from School Finance Reforms." *The Quarterly Journal of Economics* 131 (1): 157-218.

Jacob, Brian, and Jesse Rothstein. 2016. "The Measurement of Student Ability in Modern Assessment Systems." *Journal of Economic Perspectives* 30 (3): 85-108.

Krueger, Alan B. 1999. "Experimental Estimates of Education Production Functions." *The Quarterly Journal of Economics* 114 (2): 497-532.

Lafortune, Julien, Jesse Rothstein, and Diane Whitmore Schanzenbach. Forthcoming. "School Finance Reform and the Distribution of Student Achievement." *American Economic Journal: Applied Economics*.

Leamer, Edward E. 1983. "Let's Take the Con Out of Econometrics." *The American Economic Review* 73 (1): 31-43.

Lutz, Byron. 2008. "The Connection Between House Price Appreciation and Property Tax Revenues." *National Tax Journal* 61(3), 555-572.

Macartney, Hugh and John D. Singleton. 2017. "School Boards and Student Segregation." National Bureau of Economic Research.

Mian, Atif and Amir Sufi. 2009. "The Consequences of Mortgage Credit Expansion: Evidence from the U.S. Mortgage Default Crisis." *The Quarterly Journal of Economics* 124(4): 1449-1496

Murray, Sheila E., William N. Evans, and Robert M. Schwab. 1998. "Education-Finance Reform and the Distribution of Education Resources." *The American Economic Review* 88 (4): 789-812.

Oates, Wallace E. 1969. "The Effects of Property Taxes and Local Public Spending on Property Values: An Empirical Study of Tax Capitalization and the Tiebout Hypothesis." *Journal of Political Economy* 77 (6): 957-71.

Oates, Wallace E. 1999. "An Essay on Fiscal Federalism." *Journal of Economic Literature* 37 (3): 1120-1149.

Oates, Wallace E. 2005. "Toward a Second-Generation Theory of Fiscal Federalism." *International Tax and Public Finance*. 12 (4): 249-374.

Rothstein, Jesse. 2007. "Does Competition Among Public Schools Benefit Students and Taxpayers? Comment." *American Economic Review* 97 (5): 2026-2037.

Shiller, Robert J. 2005. *Irrational Exuberance* (Second Edition). Princeton, NJ: Princeton University Press.

Sinai, Todd. 2013. "House Price Moments in Boom-Bust Cycles." In *Housing and the Financial Crisis*, edited by Edward Glaeser and Todd Sinai. University of Chicago Press (2013).

Tiebout, Charles M. 1956. "A Pure Theory of Local Expenditures." *Journal of Political Economy* 64 (5): 416-24.

Vlaicu, Razvan and Alexander Whalley. 2011. "Do Housing Bubbles Generate Fiscal Bubbles? Evidence from California Cities." *Public Choice* 149: 89-108.

Table 1: Sample Restrictions and Representativeness

	All Districts	Merged with Housing	Unified Districts Only	Indep. Districts Only	Final District Finance Sample	Test Score Sample
Number of Districts	13850	2785	2070	1748	1716	1465
Enrollment	3459	10221	12334	11750	11706	13200
Revenue Per Pupil	11158	11047	10929	10696	10725	10562
Student/Teacher ratio	14.37	16.84	16.61	17.11	17.06	17.07
Percent Black (K-4)	0.07	0.10	0.10	0.10	0.10	0.11
Percent Hispanic (K-4)	0.10	0.17	0.15	0.16	0.16	0.16
Percent Free-Lunch	0.27	0.20	0.20	0.21	0.21	0.21

Notes: All variables are reported for the year 2005. Restrictions are added cumulatively; hence each column is a subset of the column directly to its left. The district finance regression sample includes only districts with constant boundaries and no missing finance data during our sample period.

Table 2: Price and Expenditure Impacts of Housing Booms and Busts

	Log Price			Log Expenditure		
	Positive	Non-Sig.	Negative	Positive	Non-Sig.	Negative
	(1)	(2)	(3)	(4)	(5)	(6)
Relative Year = 1	0.048*** (0.004)	0.003 (0.003)	-0.014*** (0.004)	0.004 (0.005)	-0.003 (0.005)	0.002 (0.007)
Relative Year = 2	0.116*** (0.006)	0.008* (0.004)	-0.032*** (0.005)	0.012* (0.006)	-0.001 (0.007)	-0.004 (0.010)
Relative Year = 3	0.168*** (0.007)	0.010 (0.006)	-0.048*** (0.008)	0.023*** (0.007)	0.004 (0.008)	-0.017 (0.010)
Relative Year = 4	0.196*** (0.008)	0.007 (0.008)	-0.079*** (0.011)	0.033*** (0.008)	0.003 (0.008)	-0.012 (0.010)
Relative Year = 5	0.201*** (0.009)	-0.003 (0.010)	-0.120*** (0.013)	0.030*** (0.008)	0.008 (0.010)	-0.025** (0.012)
R-squared	0.859	0.859	0.859	0.797	0.797	0.797
Number of observations	88,534	88,534	88,534	25,740	25,740	25,740
Time FEs	X	X	X	X	X	X
Area FEs	X	X	X	X	X	X

Notes: Prices are estimated using quarterly data, while expenditures are only available annually. All models also include a dummy for all pre-break years, a dummy for all relative years 6 and above, district fixed effects, and year fixed effects. The sample includes all independent, unified districts with no missing finance data, constant borders, and sufficient housing data to calculate breakpoints (see text for the precise criterion). Standard errors allow for clustering at the district level. ***, **, and * reflect statistical significance at 1%, 5%, and 10% confidence, respectively.

Table 3: Robustness of Price and Expenditure Effects of Housing Booms

	Main Sample	All Districts	All Indep. Districts	Trimmed Sample	Single Break	Demog. Ctrls.
	(1)	(2)	(3)	(4)	(5)	(6)
<i>Panel A. Effects on ln(Price)</i>						
Relative Year = 1	0.048*** (0.004)	0.041*** (0.003)	0.046*** (0.004)	0.046*** (0.005)	0.061*** (0.003)	0.050*** (0.004)
Relative Year = 2	0.116*** (0.006)	0.106*** (0.004)	0.115*** (0.005)	0.114*** (0.006)	0.142*** (0.005)	0.117*** (0.006)
Relative Year = 3	0.168*** (0.007)	0.154*** (0.005)	0.163*** (0.005)	0.165*** (0.007)	0.210*** (0.007)	0.166*** (0.007)
Relative Year = 4	0.196*** (0.008)	0.174*** (0.006)	0.182*** (0.006)	0.192*** (0.008)	0.254*** (0.009)	0.193*** (0.008)
Relative Year = 5	0.201*** (0.009)	0.172*** (0.007)	0.179*** (0.007)	0.196*** (0.009)	0.281*** (0.012)	0.198*** (0.009)
R-squared	0.859	0.874	0.871	0.864	0.857	0.859
Number of observations	88,534	144,605	126,495	74,316	88,534	86,682
<i>Panel B. Effects on ln(Expenditures Per Student)</i>						
Relative Year = 1	0.004 (0.005)	0.004 (0.004)	0.006 (0.004)	0.002 (0.004)	0.007 (0.005)	0.004 (0.005)
Relative Year = 2	0.012* (0.006)	0.009* (0.005)	0.012** (0.005)	0.015** (0.006)	0.014** (0.007)	0.012* (0.006)
Relative Year = 3	0.023*** (0.007)	0.013** (0.005)	0.015*** (0.006)	0.024*** (0.006)	0.030*** (0.008)	0.023*** (0.007)
Relative Year = 4	0.033*** (0.008)	0.022*** (0.006)	0.024*** (0.006)	0.032*** (0.007)	0.042*** (0.010)	0.033*** (0.008)
Relative Year = 5	0.030*** (0.008)	0.017*** (0.006)	0.018*** (0.007)	0.029*** (0.007)	0.046*** (0.010)	0.030*** (0.008)
R-squared	0.797	0.801	0.795	0.849	0.797	0.798
Number of observations	25,740	41,678	36,578	21,405	25,740	25,274

Notes: Column (1) reproduces the results in Table 2; see Table 2 notes for details of the sample and specification. Column (2) includes all districts with sufficient housing data to estimate breakpoints (see text for the precise criterion). Column (3) restricts this sample to districts with independent taxing authority. Column (4) imposes the other restrictions in our main regression sample and also removes districts whose annual revenue growth falls in the top or bottom one percent of observed values in our sample. Column (5) follows the main regression sample, but uses breakpoint results calculated from a model that allows at most one break per district. Column (6) follows the main regression sample and specification but adds controls for the percentage of minority students and the percentage of students eligible for free lunch.

Table 4: Education Expenditure Elasticities of Local House Prices

	Relative Year 2	Relative Year 3	Relative Year 4	Relative Year 5	Pooled Yrs. 3-5
	(1)	(2)	(3)	(4)	(5)
<i>All Positive Breaks</i>					
Lagged Price Elasticity	0.17** (0.09)	0.18*** (0.05)	0.20*** (0.04)	0.16*** (0.04)	0.18*** (0.04)
Concurrent Price Elasticity	0.09** (0.05)	0.14*** (0.04)	0.19*** (0.04)	0.17*** (0.04)	0.16*** (0.04)
<i>Heterogeneity in Lagged Price Elasticities</i>					
High-Expenditure Districts	0.10 (0.21)	0.20* (0.11)	0.21** (0.09)	0.17** (0.08)	0.20** (0.08)
Low-Expenditure Districts	0.20** (0.10)	0.16*** (0.06)	0.20*** (0.05)	0.16*** (0.05)	0.17*** (0.05)

Notes: Elasticities are the ratio of coefficients on log expenditures and log price. Lagged price elasticities divide expenditure coefficients by price coefficients from the previous year. Concurrent price elasticities divide expenditure and price coefficients from the same year. We collapse price data to the annual level to create a common estimation dataset and estimate models via seemingly unrelated regression to compute standard errors. Otherwise, the sample and specification follow the description in Table 2. High (low) expenditure districts are districts with per-student expenditures above (below) the sample median in 1996. Standard errors allow for clustering at the district level. ***, **, and * reflect statistical significance at 1%, 5%, and 10% confidence, respectively.

Table 5: Effects on Total Revenues and Revenue Sources

	Full Sample	No Property Tax Growth Cap			Property Tax Growth Cap		
	Log Total Revenues	Log Local Revenues	Log State Revenues	Log Federal Revenues	Log Local Revenues	Log State Revenues	Log Federal Revenues
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Relative Year = 1	0.007 (0.004)	-0.005 (0.005)	0.001 (0.007)	-0.008 (0.010)	0.014 (0.009)	0.018 (0.013)	0.024** (0.010)
Relative Year = 2	0.005 (0.005)	-0.000 (0.007)	-0.001 (0.008)	-0.026*** (0.010)	0.011 (0.009)	0.030** (0.014)	0.025** (0.011)
Relative Year = 3	0.014*** (0.005)	0.020** (0.009)	0.011 (0.010)	-0.036*** (0.012)	0.007 (0.010)	0.041*** (0.014)	0.002 (0.016)
Relative Year = 4	0.025*** (0.005)	0.023** (0.010)	0.010 (0.012)	-0.004 (0.014)	0.003 (0.012)	0.056*** (0.015)	0.014 (0.012)
Relative Year = 5	0.012** (0.005)	0.029*** (0.010)	-0.012 (0.013)	-0.004 (0.017)	-0.019 (0.012)	0.023 (0.015)	0.025 (0.016)
R-squared	0.909	0.950	0.867	0.907	0.950	0.867	0.907
Number of observations	25,740	25,739	25,739	25,721	25,739	25,739	25,721
Time FEs	X	X	X	X	X	X	X
Area FEs	X	X	X	X	X	X	X

Notes: See notes to Table 2 for details of the sample and specification.

Table 6: Effects on Expenditure Subcategories

	Log Expenditure						
	Current	Capital	Current Instruction	Current Services	Service Pupil	Service Instructional	Service Administrative
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Relative Year = 1	0.003 (0.002)	0.004 (0.036)	0.006** (0.002)	-0.001 (0.003)	-0.008 (0.006)	0.003 (0.008)	-0.005 (0.004)
Relative Year = 2	0.010*** (0.003)	0.062 (0.049)	0.012*** (0.003)	0.008** (0.004)	0.008 (0.007)	0.027** (0.011)	-0.001 (0.005)
Relative Year = 3	0.014*** (0.003)	0.116** (0.055)	0.017*** (0.003)	0.012*** (0.004)	0.022*** (0.008)	0.029** (0.013)	0.003 (0.006)
Relative Year = 4	0.018*** (0.003)	0.175*** (0.060)	0.021*** (0.003)	0.016*** (0.005)	0.032*** (0.009)	0.045*** (0.013)	0.004 (0.007)
Relative Year = 5	0.017*** (0.003)	0.126** (0.060)	0.019*** (0.004)	0.016*** (0.005)	0.042*** (0.010)	0.035** (0.013)	0.008 (0.007)
R-squared	0.958	0.293	0.956	0.929	0.893	0.795	0.857
Number of observations	25,739	25,729	25,739	25,739	25,272	25,279	25,279
Time FEs	X	X	X	X	X	X	X
Area FEs	X	X	X	X	X	X	X

Notes: See notes to Table 2 for details of the sample and specification.

Table 7: Effects on Wages, Benefits, and Teacher Employment

	Log Avg. Salary	Log Avg. Benefits	Log Pupil Tchr. Ratio
	(1)	(2)	(3)
Relative Year = 1	-0.013*** (0.004)	-0.014** (0.005)	-0.001 (0.002)
Relative Year = 2	-0.005 (0.005)	-0.002 (0.006)	-0.011 (0.008)
Relative Year = 3	0.009 (0.005)	0.012 (0.008)	-0.009*** (0.003)
Relative Year = 4	0.029*** (0.006)	0.037*** (0.009)	-0.007 (0.005)
Relative Year = 5	0.046*** (0.007)	0.051*** (0.011)	-0.008** (0.004)
R-squared	0.793	0.866	0.831
Number of observations	24,178	24,178	24,864
Time FEs	X	X	X
Area FEs	X	X	X

Notes: See notes to Table 2 for details of the sample and specification. All dependent variables are in logs.

Table 8: Effects on NAEP Test Scores

	Math	Reading	Grade 4 Math	Grade 4 Reading	Grade 8 Math	Grade 8 Reading
	(1)	(2)	(3)	(4)	(5)	(6)
Relative Year in [-3, -2]	-0.0140 (0.0289)	0.0439* (0.0266)	-0.0254 (0.0499)	0.00409 (0.0489)	0.00136 (0.0444)	0.0429 (0.0331)
Relative Year in [1, 2]	0.0104 (0.0228)	0.0310 (0.0223)	0.0412 (0.0383)	0.0322 (0.0418)	-0.00268 (0.0338)	0.0102 (0.0280)
Relative Year in [3, 4]	-0.00309 (0.0231)	0.0425* (0.0253)	0.00658 (0.0390)	0.0484 (0.0453)	-0.00641 (0.0345)	0.0247 (0.0305)
Relative Year in [5, 6]	0.0278 (0.0291)	0.0899*** (0.0294)	0.0239 (0.0465)	0.111** (0.0537)	0.0229 (0.0402)	0.0558 (0.0346)
Relative Year in [7, 8]	0.0289 (0.0359)	0.0993*** (0.0367)	0.0137 (0.0546)	0.133** (0.0666)	0.0147 (0.0535)	0.0471 (0.0432)
Observations	7,430	7,547	3,711	3,751	3,719	3,796
R-squared	0.772	0.762	0.816	0.762	0.829	0.783
Time FEs	X	X	X	X	X	X
Area FEs	X	X	X	X	X	X

Columns (1) and (2) pool 4th and 8th grade test results together and include grade-level dummies. See the notes to Table 2 for other details of the specification and sample restrictions.

Figure 1A: School District Expenditures Per Student

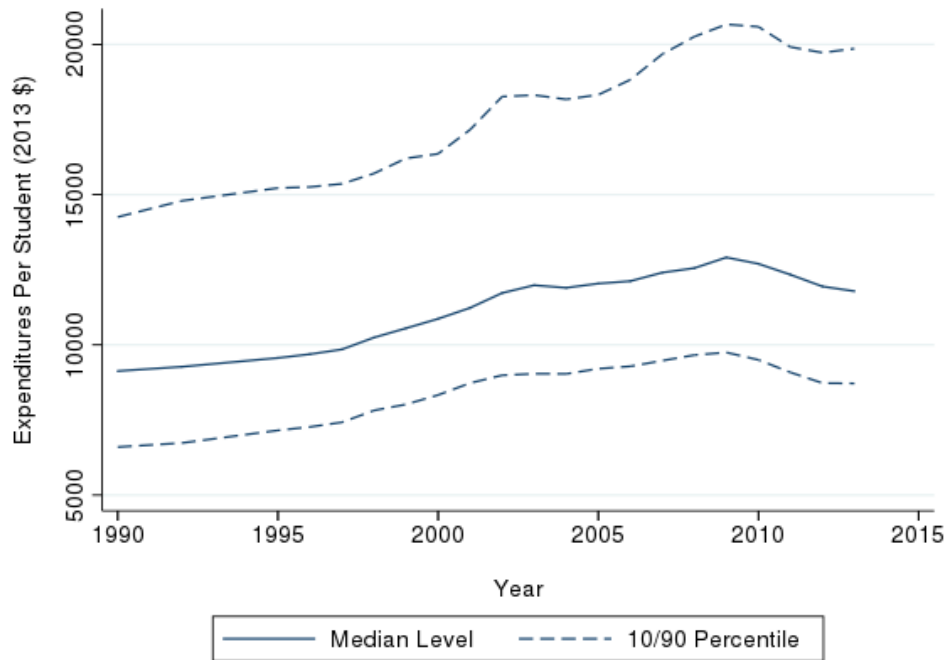
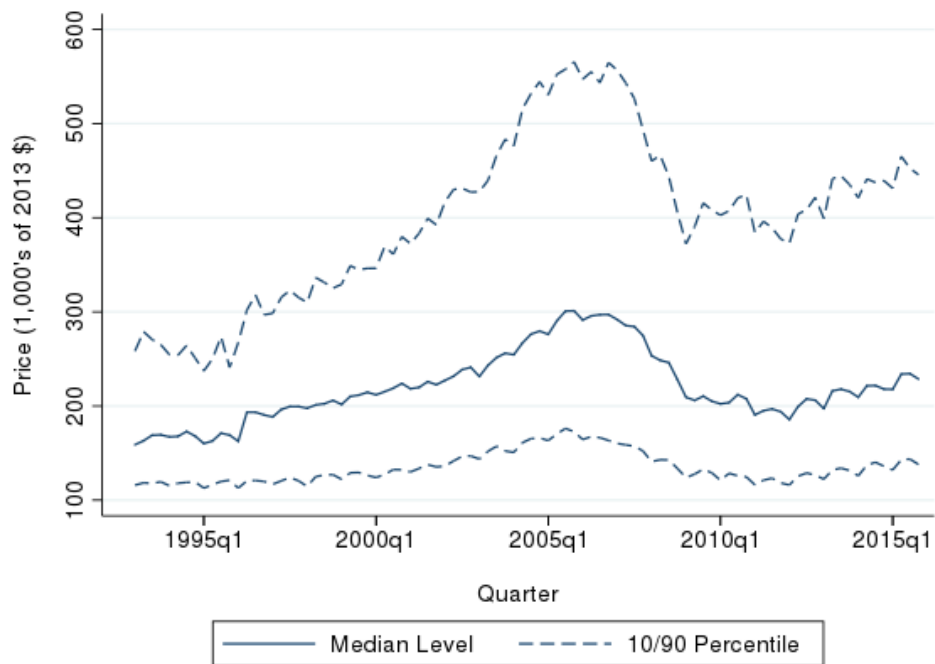


Figure 1B: School District Average House Prices



Notes: Plots show percentiles among school districts in our final regression sample (i.e. all independent, unified districts with no missing finance data, constant borders, and sufficient housing data to calculate breakpoints).

Figure 2A: School District House Price Indices

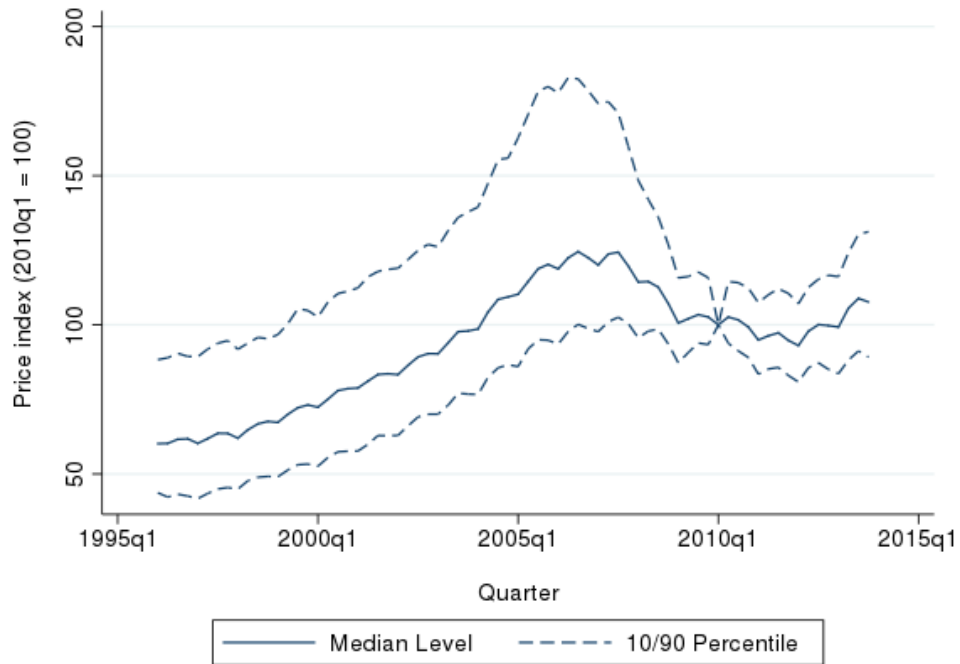
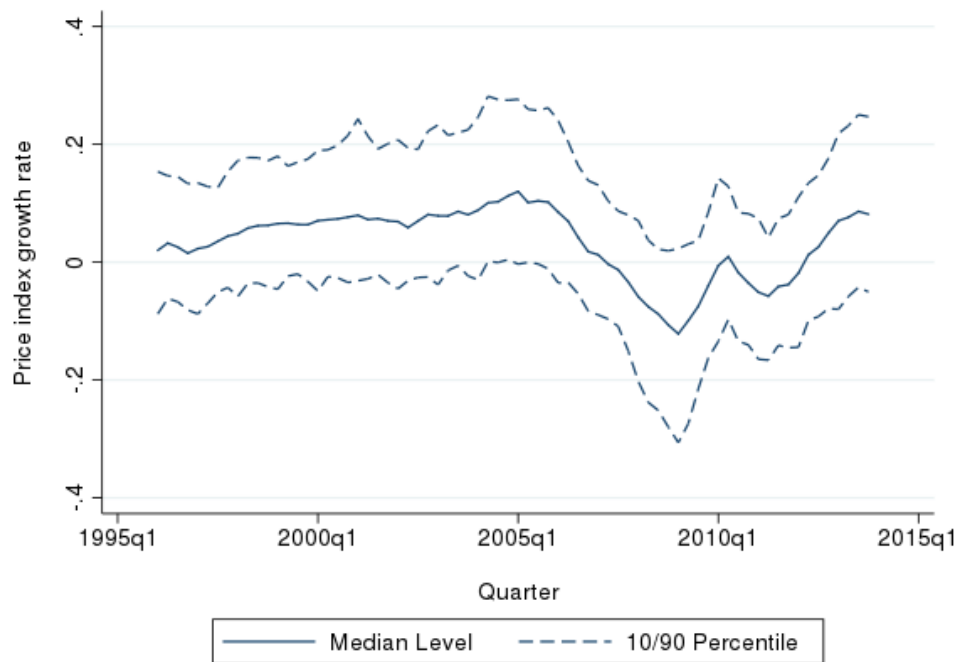


Figure 2B: School District House Price Index Growth Rates



Notes: Plots show percentiles among school districts in our final regression sample (i.e. all independent, unified districts with no missing finance data, constant borders, and sufficient housing data to calculate breakpoints).

Figure 3: Examples of Breakpoint Estimates



Figure 4: Timing of Structural Breaks in School District House Prices

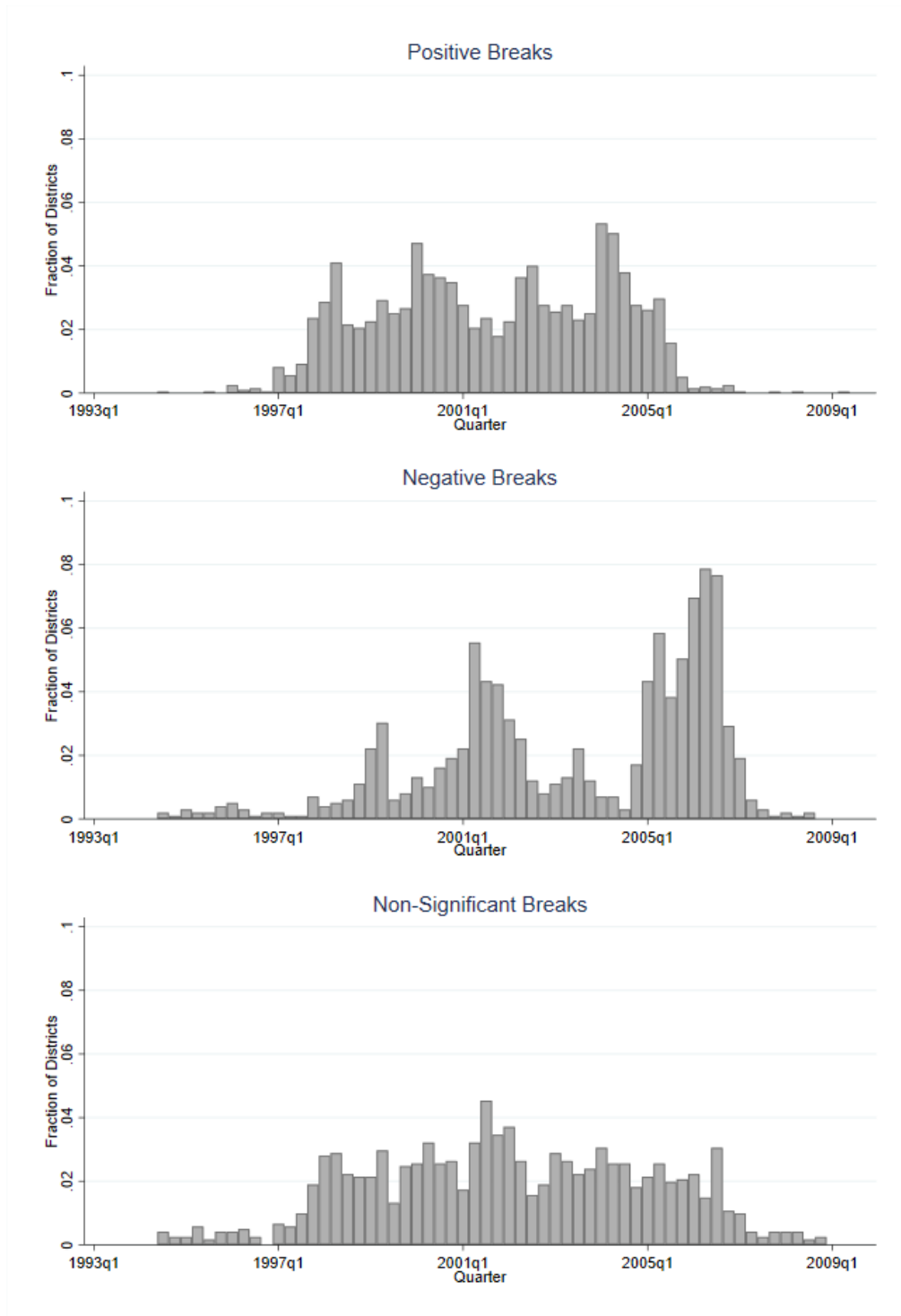
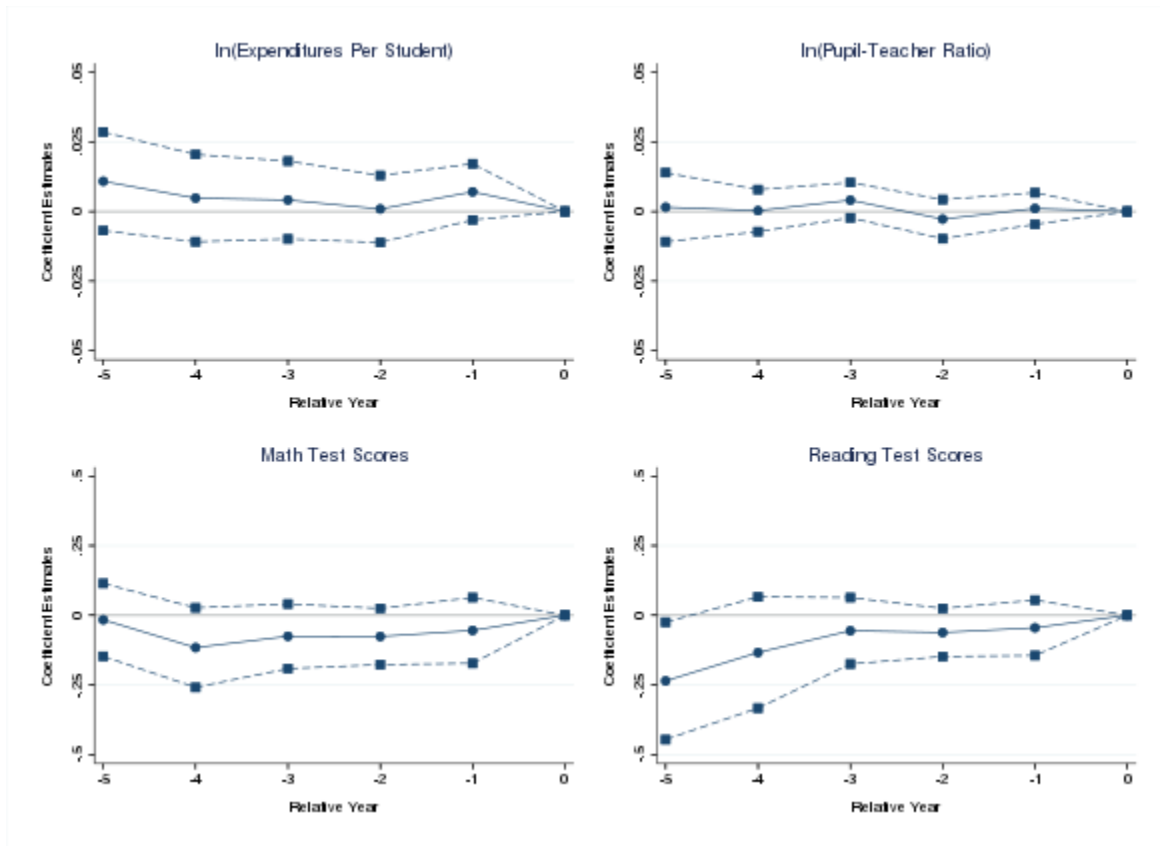
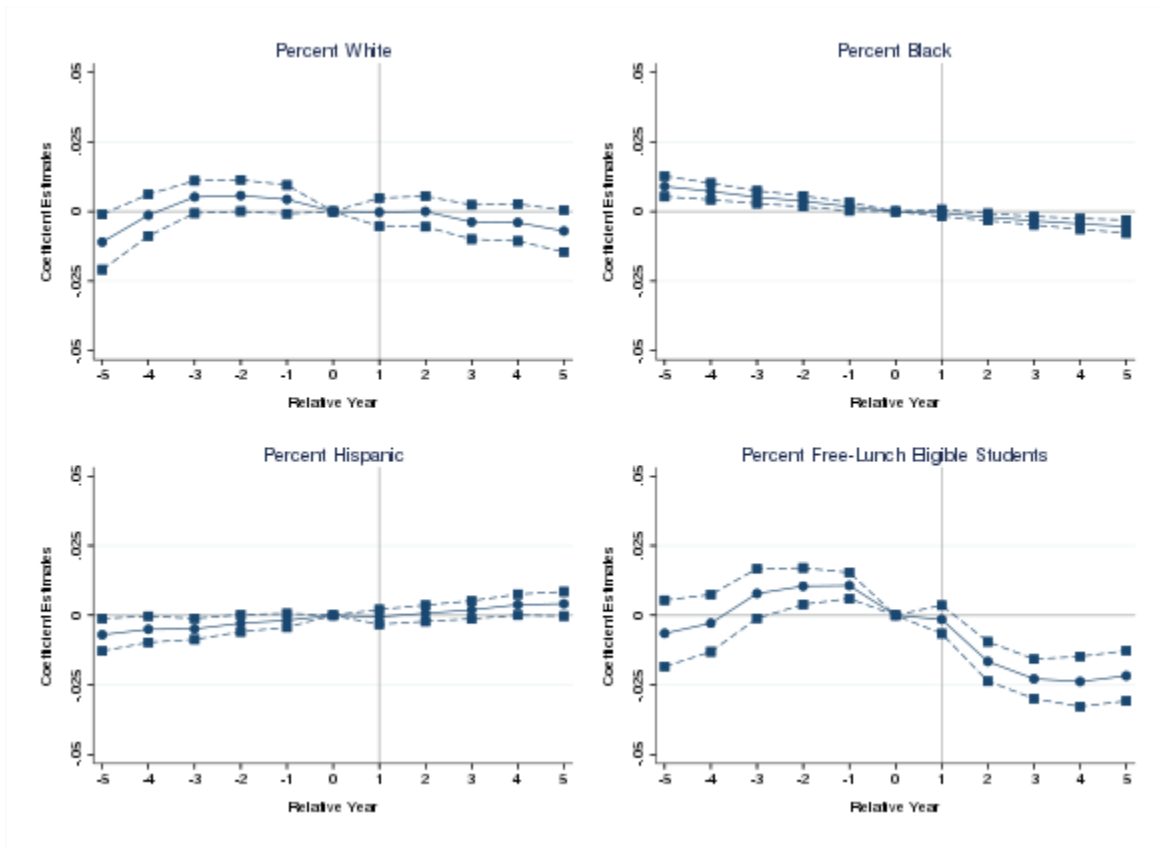


Figure 5A: Pre-Boom Trends in School Quality Proxies



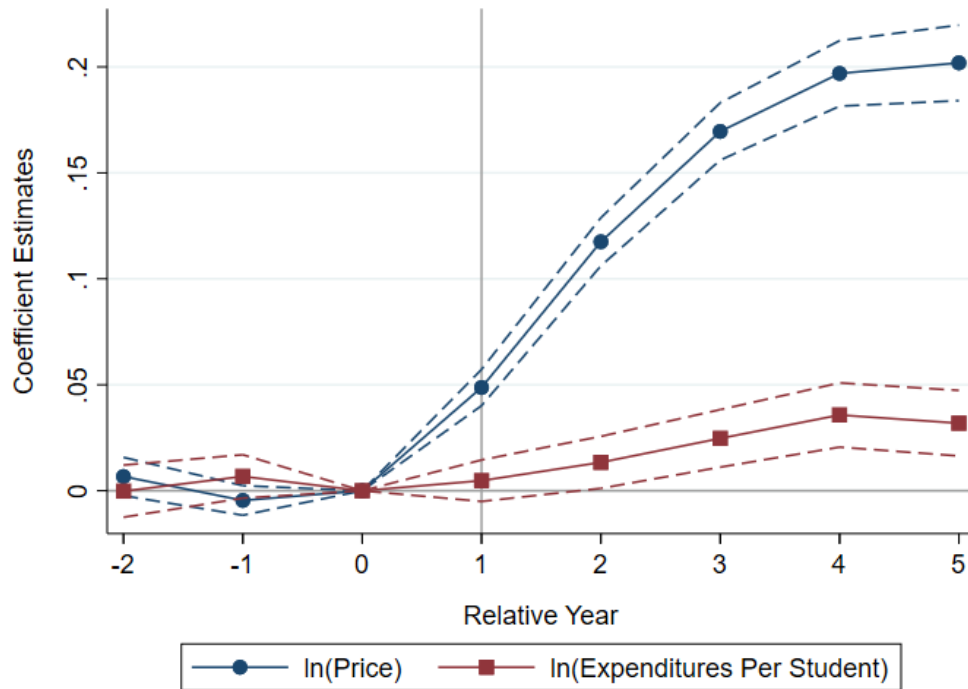
Notes: Plotted coefficients are from a model with five relative year dummies ranging between -5 and -1, a post-break dummy, a dummy for relative years less than -6, district fixed effects, and year fixed effects. The sample includes all independent, unified districts with no missing finance data, constant borders, and sufficient housing data to calculate breakpoints. Dotted lines show 95% confidence intervals, with standard errors clustered at the district level.

Figure 5B: Demographic Shifts During Housing Booms



Notes: Plotted coefficients are from a model with five relative year dummies ranging between -5 and -5, a dummy for relative years greater than 6, a dummy for relative years less than -6, district fixed effects, and year fixed effects. The sample includes all independent, unified districts with no missing finance data, constant borders, and sufficient housing data to calculate breakpoints. Dotted lines show 95% confidence intervals, with standard errors clustered at the district level.

Figure 6: The Effects of a Housing Boom on Prices and School District Expenditures



Notes: Plotted coefficients are from a model with 10 relative year dummies ranging between -5 and 5 (omitting zero), a dummy for all relative years -6 and lower, a dummy for all relative years 6 and above, district fixed effects, and year fixed effects. The sample includes all independent, unified districts with no missing finance data, constant borders, and sufficient housing data to calculate breakpoints. Prices are estimated using quarterly data, while expenditures are only available annually. Dotted lines show 95% confidence intervals, with standard errors clustered at the district level.

Appendix

Breakpoint Identification and Testing

This section borrows heavily from DeFusco et al. (2017) and adopts the notation in Estrella (2003).

Our goal is to estimate t^* and assess its statistical significance. Let $\Pi_i = [\pi_{i,1}, \pi_{i,2}]$ be a closed interval in $(0,1)$ and let S_i be the set of all observations from $t = \text{int}(\pi_{i,1}T)$ to $t = \text{int}(\pi_{i,2}T)$, where $\text{int}(\cdot)$ denotes rounding to the nearest integer. The estimated break point is the value t^* from the set S_i that maximizes the likelihood ratio statistic from a test of H_1 against H_0 .²⁷ That is, for every $t \in S_i$ we construct the likelihood ratio statistic corresponding to a test of H_1 against H_0 for that value of t , and we take the t that produces the largest test-statistic as our estimated break point for district i .

Assessing the statistical significance of this breakpoint estimate requires knowing the distribution of the supremum of the likelihood ratio statistic as calculated from among the values in S_i . Let $\xi_i = \sup_{S_i} LR$ denote this supremum. Andrews (1993) shows that this distribution can be written as

$$(A1) \quad P(\xi_i > c) = P(\sup_{\pi_i \in \Pi_i} Q_1(\pi_i) > c) = P\left(\sup_{1 < s < \lambda_i} \frac{\|B_1(s)\|}{s^{1/2}} > c^{1/2}\right)$$

where $\|B_1(s)\|$ is the Bessel process of order 1, $\lambda_i = \pi_{i,2}(1 - \pi_{i,1})/\pi_{i,1}(1 - \pi_{i,2})$, and

$$Q_1(\pi_i) = \frac{(B_1(\pi_i) - \pi_i B_1(1))'(B_1(\pi_i) - \pi_i B_1(1))}{\pi_i(1 - \pi_i)}.$$

Direct calculation of the probability in (2) is non-trivial and prior research has relied on approximations that typically are based on simulation or curve-fitting methods (Andrews 1993, Hansen 1997). However, Estrella (2003) provides a numerical procedure for calculating exact p -values that does not rely on these types of approximations. We use this method to calculate p -values for the estimated break point, π_i , for each district in the sample.

We have not yet said where the interval endpoints $\pi_{i,1}$ and $\pi_{i,2}$ come from. We do not allow breakpoints to fall in the first two or last two quarters in our sample. These values vary by district because the length of the available series depends on both data availability and the timing of the peak of the housing market in each district.

Multiple Breaks

²⁷ We use the terms supremum and maximum interchangeably in this exposition. Technically, all of the results are in terms of the supremum of the likelihood ratio statistic.

In estimating the break points, we allow for the possibility that a given market might experience more than one housing boom during the course of our sample period. Our method is recursive in that we first test for the existence of one break point against the null hypothesis of zero. Given the existence of at least one break point, we can then test the hypothesis of $m + 1$ break points against the null of m using the results from Bai (1999). Bai and Perron (1998) show that the test for one break is consistent in the presence of multiple breaks, which is what allows for this sequential estimation procedure.

More specifically, let $0 < \varphi_{i,1} < \dots < \varphi_{i,m} < 1$ mark the proportions of the sample generated by the m break points estimated under the null hypothesis for district i . For technical reasons, we require that $\varphi_{i,j} - \varphi_{i,j-1} > \pi_{i,0}$ for some small $\pi_{i,0}$ ²⁸ where we define $\varphi_{i,0} = 0, \varphi_{i,m+1} = 1$. Further, let $\eta_{i,j} = \frac{\pi_{i,0}}{\varphi_{i,j} - \varphi_{i,j-1}}, j = 1, \dots, m + 1$. The likelihood ratio test compares the maximum of the likelihood ratio obtained when allowing for $m + 1$ breaks to that from only allowing for m . The distribution of this likelihood ratio statistic is given by

$$(A2) P(LR > c) = 1 - \prod_{i=1}^{m+1} \left(1 - P \left(\sup_{\pi_i \in [\eta_{i,j}, 1 - \eta_{i,j}]} Q_1(\pi_i) > c \right) \right),$$

which we calculate by recursive application of the method provided in Estrella (2003).

We apply this procedure to test for the existence of two break points against the null of one as well as three against the null of only two among those districts for which we find at least two statistically significant break points.

²⁸ In practice, we require breakpoints to be separated by at least three quarters. Hence $\pi_{i,0} = 1/T_i$, where T_i denotes the number of periods in the time series for district i .

Appendix Table 1: Effect Heterogeneity by Baseline District Expenditures

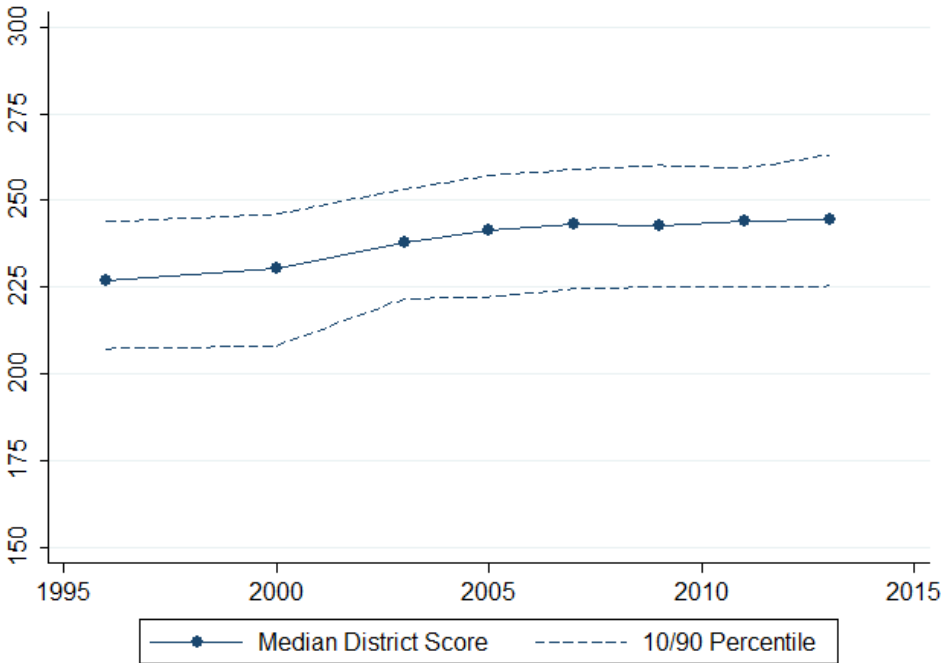
	Log Price		Log Exp. Per Student	
	High-Exp. Districts	Low-Exp. Districts	High-Exp. Districts	Low-Exp. Districts
	(1)	(2)	(3)	(4)
Relative Year = 1	0.044*** (0.005)	0.077*** (0.007)	-0.007 (0.008)	0.011* (0.006)
Relative Year = 2	0.097*** (0.007)	0.142*** (0.008)	0.005 (0.010)	0.015* (0.008)
Relative Year = 3	0.133*** (0.009)	0.181*** (0.010)	0.019* (0.011)	0.023*** (0.009)
Relative Year = 4	0.154*** (0.012)	0.191*** (0.012)	0.028** (0.012)	0.036*** (0.010)
Relative Year = 5	0.161*** (0.014)	0.175*** (0.013)	0.027** (0.013)	0.031*** (0.010)
R-squared	0.878	0.867	0.758	0.594
Number of observations	11,395	11,775	11,395	11,775
Time FEs	X	X	X	X
Area FEs	X	X	X	X

Notes: To create a common analysis dataset for prices and expenditures, we average the quarterly price series to the district-year level. See the Table 2 notes for other details of the sample and specification.

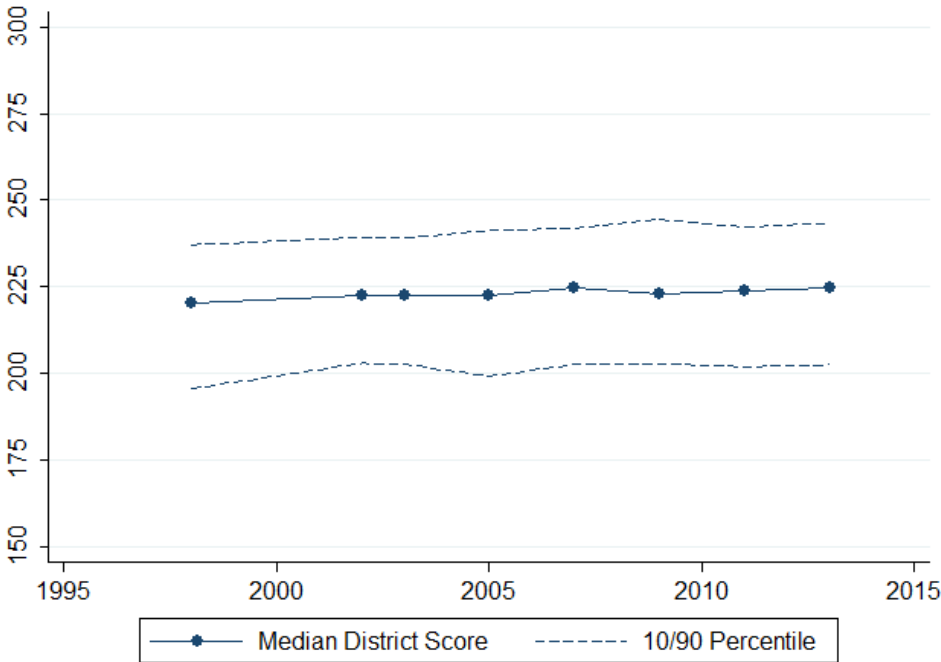
Appendix Table 2: Effects on Wages and Benefits by Subcategories

	Log Avg. Salary Instruction	Log Avg. Salary Administrator	Log Avg. Salary Other	Log Avg. Benefit Instruction	Log Avg. Benefit Administrator	Log Avg. Benefit Other
	(1)	(2)	(3)	(4)	(5)	(6)
Relative Year = 1	0.003 (0.003)	-0.058*** (0.012)	-0.051*** (0.010)	0.005 (0.005)	-0.058*** (0.013)	-0.055*** (0.011)
Relative Year = 2	0.006* (0.004)	-0.067*** (0.014)	-0.033*** (0.013)	0.018*** (0.006)	-0.062*** (0.014)	-0.037** (0.014)
Relative Year = 3	0.011*** (0.004)	-0.066*** (0.016)	0.020 (0.014)	0.021*** (0.006)	-0.059*** (0.016)	0.019 (0.016)
Relative Year = 4	0.018*** (0.005)	-0.056*** (0.017)	0.058*** (0.016)	0.033*** (0.008)	-0.043** (0.018)	0.060*** (0.018)
Relative Year = 5	0.020*** (0.005)	-0.034* (0.020)	0.084*** (0.018)	0.022*** (0.008)	-0.020 (0.022)	0.082*** (0.021)
R-squared	0.807	0.616	0.698	0.877	0.718	0.756
Number of observations	23,082	23,082	23,082	23,082	23,082	23,082
Time FEs	X	X	X	X	X	X
Area FEs	X	X	X	X	X	X

Appendix Figure 1A: School District NAEP Math Scores (Fourth Grade)



Appendix Figure 1B: School District NAEP Reading Scores (Fourth Grade)



Notes: Plots show percentiles among school districts in our final regression sample (i.e. all independent, unified districts with no missing finance data, constant borders, and sufficient housing data to calculate breakpoints).