Online Appendix

Housing Disease and Public School Finances

Matthew Davis and Fernando Ferreira

Breakpoint Identification and Testing

This section borrows heavily from DeFusco et al. (2018) 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 $t = int(\pi_{i,1}T)$ to $t = int(\pi_{i,2}T)$, where $int(\cdot)$ 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.[1] 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

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

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

[1] 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.

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

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 $0 < \varphi_{i,1} < \cdots < \varphi_{i,m} < 1$ 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 $\pi_{i,0}$[2] where we define $\varphi_{i,0} = 0, \varphi_{i,m+1} = 1$. Further, let $\eta_{i,j} = \frac{\pi_{i,0}}{\varphi_{i,j}-\varphi_{i,j-1}}, j = 1, \ldots, m + 1$. 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.

[2] 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: Effects of MSA Level Booms and Busts on BEA Income

	BEA Income Growth Rate
Relative Year = 1	-0.001
	(0.004)
Relative Year = 2	-0.002
	(0.003)
Relative Year = 3	-0.004
	(0.003)
Relative Year = 4	-0.000
	(0.003)
Relative Year = 5	-0.006**
	(0.003)
R-squared	0.451
Number of observations	1,620
Dependent variable mean	0.047
Time FEs	X
MSA FEs	X

Notes: Reproduced from Ferreira and Gyourko (2021).

Appendix Table 2: 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***	0.077***	-0.007	0.011*
	(0.005)	(0.007)	(0.008)	(0.006)
Relative Year = 2	0.097***	0.142***	0.005	0.015*
	(0.007)	(0.008)	(0.010)	(0.008)
Relative Year = 3	0.133***	0.181***	0.019*	0.023***
	(0.009)	(0.010)	(0.011)	(0.009)
Relative Year = 4	0.154***	0.191***	0.028**	0.036***
	(0.012)	(0.012)	(0.012)	(0.010)
Relative Year = 5	0.161***	0.175***	0.027**	0.031***
	(0.014)	(0.013)	(0.013)	(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 specification.

Appendix Table 3: Municipal Price and Expenditure Effects of Booms and Busts

	Log Price			Log Expenditure		
	Positive	Non-Sig.	Negative	Positive	Non-Sig.	Negative
	(1)	(2)	(3)	(4)	(5)	(6)
Relative Year = 1	0.036***	0.005	-0.015***	-0.013	-0.008	-0.007
	(0.006)	(0.005)	(0.006)	(0.009)	(0.013)	(0.013)
Relative Year = 2	0.107***	0.005	-0.036***	-0.011	-0.007	-0.026
	(0.007)	(0.007)	(0.007)	(0.010)	(0.014)	(0.017)
Relative Year = 3	0.162***	0.000	-0.064***	0.020	-0.000	-0.041**
	(0.008)	(0.009)	(0.009)	(0.013)	(0.018)	(0.017)
Relative Year = 4	0.193***	-0.008	-0.102***	0.037***	-0.019	-0.036*
	(0.008)	(0.012)	(0.012)	(0.014)	(0.018)	(0.019)
Relative Year = 5	0.186***	-0.027*	-0.138***	0.049***	-0.004	-0.062**
	(0.010)	(0.015)	(0.013)	(0.014)	(0.023)	(0.024)
R-squared	0.879	0.879	0.879	0.900	0.900	0.900
Number of observations	48,390	48,390	48,390	10,991	10,991	10,991
Time FEs	X	X	X	X	X	X
Area FEs	X	X	X	X	X	X
Cities	913	913	913	913	913	913

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

Appendix Table 4: Effects on the Residential Tax Base, Total District Revenues, and Total District Expenditures

	Tax Base	Total Revenue	Total Exp.
	(1)	(2)	(3)
Relative Year = 1	1.577e+09***	5.087e+06***	4.366e+06***
	(3.314e+08)	(1.478e+06)	(1.585e+06)
Relative Year = 2	2.408e+09***	1.051e+07***	1.214e+07***
	(6.011e+08)	(2.507e+06)	(2.916e+06)
Relative Year = 3	2.526e+09***	1.665e+07***	1.907e+07***
	(7.891e+08)	(3.600e+06)	(3.898e+06)
Relative Year = 4	1.872e+09***	1.694e+07***	2.280e+07***
	(7.011e+08)	(3.631e+06)	(4.614e+06)
Relative Year = 5	1.816e+09**	1.634e+07***	2.169e+07***
	(7.917e+08)	(4.510e+06)	(5.507e+06)
Relative Year = 6	1.876e+09**	1.777e+07***	2.055e+07***
	(8.923e+08)	(6.444e+06)	(7.180e+06)
R-squared	0.893	0.976	0.965
Observations	24,336	25,740	25,740
Time FEs	X	X	X
Area FEs	X	X	X

Notes: See Table 2 notes for details of the sample and specification. We measure the value of housing stocks by multiplying average transaction prices in the CoreLogic data by the number of housing units in each district. The latter are obtained from the 2000 and 2010 Censuses, and we interpolate linearly in other years.

Appendix Table 5: Effects on Capital and Current Expenditures by High/Low Expenditure Subgroups

	High Expenditure Subgroup		Low Expenditure Subgroup	
	Log Capital	Log Current Expenditures	Log Capital Capital	Log Current Expenditures
Relative Year = 1	-0.061	-0.002	0.04	0.005*
	(0.064)	(0.003)	(0.042)	(0.003)
Relative Year = 2	0.087	0.003	0.029	0.014***
	(0.084)	(0.004)	(0.06)	(0.003)
Relative Year = 3	0.116	0.010**	0.097	0.015***
	(0.093)	(0.004)	(0.067)	(0.004)
Relative Year = 4	0.178*	0.018***	0.157**	0.017***
	(0.096)	(0.005)	(0.076)	(0.004)
Relative Year = 5	0.143	0.019***	0.106	0.016***
	(0.097)	(0.006)	(0.075)	(0.004)
Relative Year = 6	0.016	0.042***	0.148*	0.023***
	(0.106)	(0.008)	(0.084)	(0.006)
R-Squared	0.276	0.953	0.322	0.878
N	12862	12869	12867	12870
Time FEs	X	X	X	X
Area FEs	X	X	X	X

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

Appendix Table 6: Price and Expenditure Impacts of Housing Booms Controlling for Time by County Fixed Effects

	Log Price	Log Expenditure
Relative Year = 1	0.055***	0.007
	(0.004)	(0.005)
Relative Year = 2	0.131***	0.017***
	(0.006)	(0.006)
Relative Year = 3	0.191***	0.030***
	(0.008)	(0.007)
Relative Year = 4	0.225***	0.042***
	(0.01)	(0.008)
Relative Year = 5	0.237***	0.039***
	(0.012)	(0.009)
Relative Year = 6	0.292***	0.046***
	(0.016)	(0.012)
R-Squared	0.926	0.843
N	88534	25740
Time by County FEs	X	X
Area FEs	X	X

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

Appendix Table 7: Local Revenue and Education Expenditure Elasticities of School District House Prices, by High and Low Local Revenue District

	High Revenue Districts	Low Revenue Districts	High Local Rev. Districts	Low Local Rev. Districts
All Positive Breaks:	Local Revenue Elasticities		Total Expenditure Elasticities	
Lagged Year 2:	0.16	-0.03	0.23	0.13
	(0.17)	(0.10)	(0.21)	(0.11)
Lagged Year 3:	0.14	0.04	0.30***	0.12*
	(0.10)	(0.06)	(0.11)	(0.06)
Lagged Year 4:	0.14*	0.02	0.32***	0.15***
	(0.08)	(0.05)	(0.10)	(0.05)
Lagged Year 5:	0.11	-0.02	0.27***	0.12**
	(0.08)	(0.05)	(0.08)	(0.05)

Notes: See Table 4 notes for details of the specification.

Appendix Table 8: Effects on Wages and Benefits by Subcategories

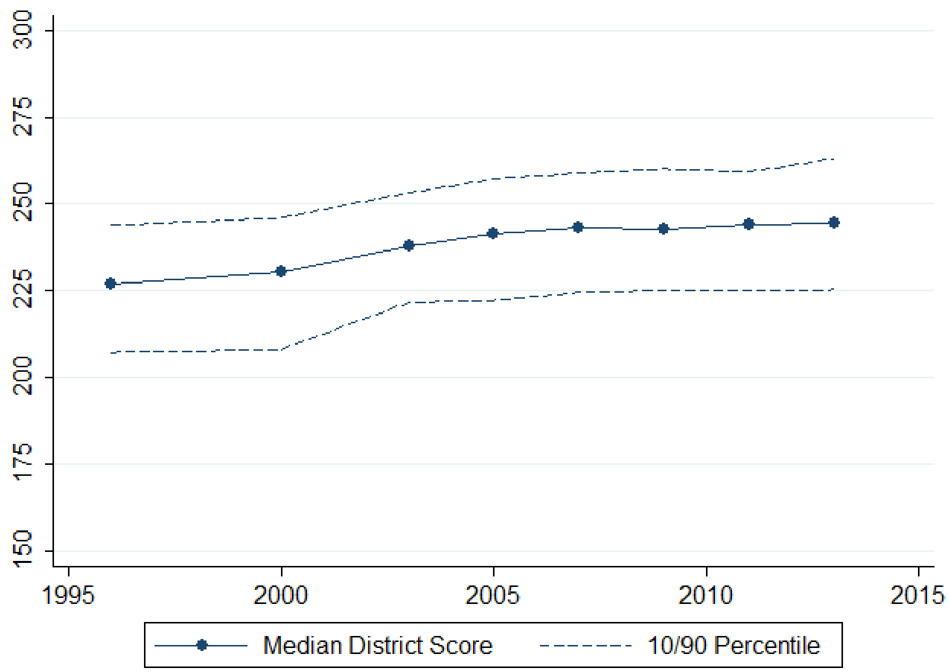
	Log Avg. Salary Instruction	Log Avg. Salary Administrator	Log Avg. Salary Other	Log Avg. Benefit Instruction	Log Avg. Benefit Administrator	Log Avg. Benefit Other
	(1)	(2)	(3)	(4)	(5)	(6)
Relative Year = 1	0.003	-0.058***	-0.051***	0.005	-0.058***	-0.055***
	(0.003)	(0.012)	(0.010)	(0.005)	(0.013)	(0.011)
Relative Year = 2	0.006*	-0.067***	-0.033***	0.018***	-0.062***	-0.037**
	(0.004)	(0.014)	(0.013)	(0.006)	(0.014)	(0.014)
Relative Year = 3	0.011***	-0.066***	0.020	0.021***	-0.059***	0.019
	(0.004)	(0.016)	(0.014)	(0.006)	(0.016)	(0.016)
Relative Year = 4	0.018***	-0.056***	0.058***	0.033***	-0.043**	0.060***
	(0.005)	(0.017)	(0.016)	(0.008)	(0.018)	(0.018)
Relative Year = 5	0.020***	-0.034*	0.084***	0.022***	-0.020	0.082***
	(0.005)	(0.020)	(0.018)	(0.008)	(0.022)	(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

Notes: See Table 2 notes for details of the sample and specification. 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) after aggregating the different classification schemes in each file up to the broad groupings described here. Districts with fewer than ten employees in a given category are dropped.

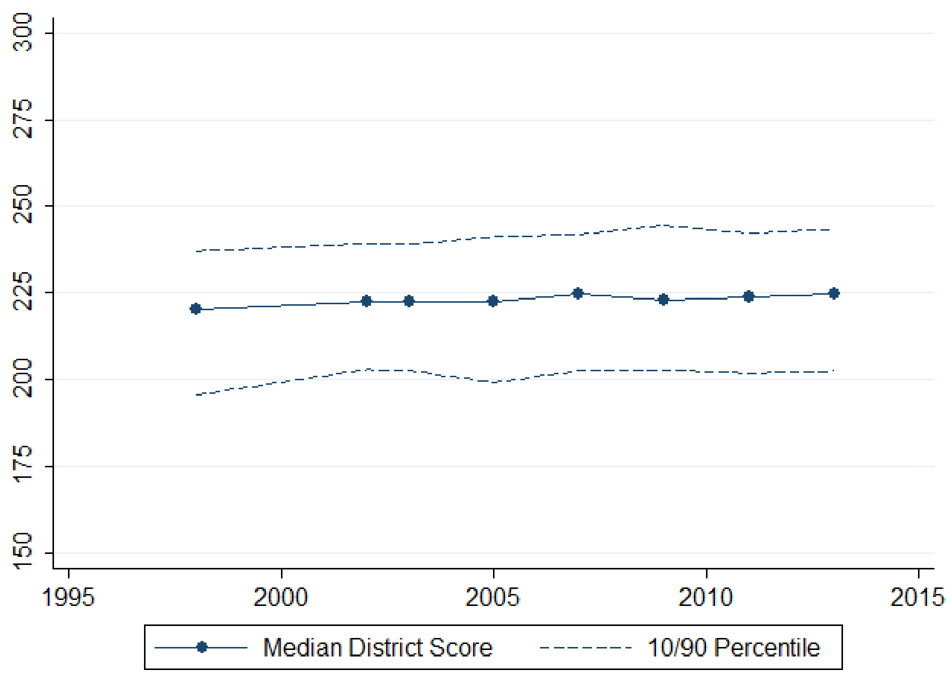
Appendix Table 9: Effects on Total District Revenues and Revenue Sources

	Full Sample		
	Log Local Revenues	Log State Revenues	Log Federal Revenues
	(1)	(2)	(3)
Relative Year = 1	0.005	0.010	0.007
	(0.005)	(0.007)	(0.007)
Relative Year = 2	0.007	0.010	-0.003
	(0.006)	(0.008)	(0.008)
Relative Year = 3	0.016**	0.020**	-0.019**
	(0.007)	(0.009)	(0.010)
Relative Year = 4	0.016**	0.026**	0.005
	(0.008)	(0.011)	(0.010)
Relative Year = 5	0.009	-0.001	0.011
	(0.008)	(0.011)	(0.012)
R-squared	0.949	0.866	0.907
Number of observations	25,739	25,739	25,721
Time FEs	X	X	X
Area FEs	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).