What's in a Grade?
School Report Cards and House Prices

David N. Figlio
Walter J. Matherly Professor of Economics
Department of Economics
University of Florida
Gainesville, FL 32611-7140
and National Bureau of Economic Research

Maurice E. Lucas
Director of Research, Evaluation and Zoning
School Board of Alachua County
620 East University Avenue
Gainesville, FL 32601

January 2001

What's in a Grade? School Report Cards and House Prices

1. Introduction

In January 2001 President George W. Bush introduced education reform legislation to

Congress as his first priority as the new president. The cornerstone of the Bush education plan

involves a system of school accountability, in which each school is graded according to its

performance on standardized tests. While it is not clear at the time of this paper's writing what

the specific nature of the final law will be, it appears quite likely that a federal system of assessing

and classifying schools will soon become a reality.

The popular discussion surrounding the Bush education reform proposals has focused on

the distribution of federal aid to K-12 schools (which accounts for only a small fraction of total

school revenues and expenditures in the United States, and even in the affected schools) and the

potential for school vouchers, which are a proposed consequence of persistent school "failure"

under the reform plans.   However, the national conversation has to date been silent on a potential

consequence affecting the assets of a much larger set of Americans–their house values. There is

reason to believe that public information on school quality measures influences house prices. For

instance, compelling work by Black (1999) suggests that student test scores are reflected strongly

in housing values.[1]   However, to date we know of no research that explores the effects of

---

[1]Black (1999) estimates school quality effects on house prices by comparing house prices
on either side of, but adjacent to, elementary school zoning boundaries. Other recent papers using
similar identification strategies, though not at such a fine level of geography, include Bogart and
Cromwell (1997), Downes and Zabel (1997), and Hayes and Taylor (1996). Several other recent
papers address similar questions as well. Bogart and Cromwell (2000) adopt a similar strategy to
measure the degree to which housing prices respond to a disruption of neighborhood schools
following rezoning in Shaker Heights, Ohio. Bradbury, Case and Mayer (forthcoming) also
consider the degree to which school quality measures and tax rates are capitalized into house
prices in Massachusetts. Brunner, Sonstelie and Thayer (2000) show that homeowners paying a

governments' providing additional information about school quality beyond test score data. We ask whether the housing market *further responds* to the discrete grade distinctions between schools assigned by the state, even when publicly observed school attributes like test scores are controlled for. Put differently, we seek to determine whether houses in the enrollment zones of two schools known to have highly comparable attributes (and whose market values previously reflected these attributes) are valued systematically differently if their related schools are assigned different report card grades by the government. That is, *we ask whether there is an independent effect of school grading on the housing market.*

To achieve this goal, we exploit the fact that President Bush is not the first to propose a system of government-imposed school grades. President Bush's accountability plan is very similar in many respects to the accountability system implemented under his brother Jeb's governorship in May, 1999 in Florida.[2] Florida's system, like all classification systems, is necessarily based on judgment calls, and equally plausible grading systems could be constructed using the same data but that reverse some of the grades assigned to schools. We employ a rich house price data set

---

high house price premium for school quality are less likely to support a school voucher initiative in California.

[2]Florida, too, is not unique in having a system of school accountability that assigns discrete school grades to every school in the system. Indeed, forty states use student test scores to rate school performance, and twenty states explicitly classify schools into various performance categories--in essence, providing discrete report card "grades" for schools. For example, Texas grades schools on a four-point scale from "exemplary" to "unacceptable," Oregon grades schools on a five-point scale from "exceptional" to "unacceptable," Louisiana rates schools on a six-point scale from "school of academic excellence" to "academically unacceptable school," and California rates schools on a ten-point scale. Other states (e.g., New Jersey) offer fewer grade distinctions, but aim to identify the "best" set of schools in their states, while still others (e.g., Virginia) base school accreditation on student test performance. Florida explicitly assigns schools a letter grade from "A" to "F" as the centerpiece of its educational accountability system.

from the Gainesville, Florida metropolitan area to investigate the effects of the assignment of school grades in Florida.

Like Black, we are concerned about the potential for omitted variables bias (because, for instance, "better" neighborhoods may have "better" schools) and therefore design an identification strategy that works to circumvent this empirical problem. Specifically, we utilize panel data on every real estate transaction involving single-family houses in platted subdivisions (distinct neighborhoods typically developed at about the same time with very similar houses, in terms of style, square footage, and lot size) from January 1995 until early July 2000 so that we can control for unobserved, time-invariant property-specific fixed effects. Furthermore, because we can identify the precise platted subdivision of each parcel, we can control for neighborhood-year interactions, so that anything common to all properties in the same neighborhood at the same time is controlled for in the analysis. Neighborhood definition is very fine: platted subdivisions in Gainesville have on average 143 houses, and 75 percent of subdivisions have fewer than 200 houses. To this house price panel we map elementary-school zoning areas, so that we can identify for each property its school grade in 1999, as well as the test performance of its school in each year of our study. Because school grades were assigned in the middle of Gainesville's peak real estate season, we observe many transactions taking place in the months immediately prior to letter grade assignment as well as in the months immediately following letter grade assignment. In all, 199 subdivisions in 20 elementary school zoning areas experienced house sales in both the five months prior to letter grade assignment and the five months following letter grade assignment.[3]

---

[3]This excludes sales in neighborhoods zoned for Lawton Chiles Elementary School, which opened in 1999 and was not graded that year.

Hence, we are identifying school grade effects from changes in house prices within a tightly-defined neighborhood in the months surrounding letter grade imposition, while controlling for all time invariant attributes of each property.

Gainesville is an ideal location to study this issue. Florida has county-level school districts, and the entire Gainesville metropolitan area lies within a single school district. This is an advantage because it ensures that every observation in our data is operating under the same policy regime at the same time. Moreover, while the Gainesville metropolitan area has over 200,000 residents today, the number of residents has quadrupled in the last 40 years, such that the vast majority of houses in the metropolitan area lie in well-defined platted subdivisions, helping us to more carefully identify neighborhood effects. Third, despite its rapid growth, Gainesville's zoning boundaries have remained very stable over our study period, with only one new elementary school added to the system and only small disturbances to zoning lines (which are, nonetheless, captured by our identification strategy.) Fourth, Gainesville properties sell at high frequency, allowing us to use repeat-sales data (and thus control for property-specific fixed effects) in our analysis. Indeed, almost 45 percent of the single-family housing stock in platted subdivisions in existence in January 1995 sold at least twice during our sample period, and nearly ten percent sold three or more times during the sample period. Finally, Gainesville had no "F"-rated schools in the 1999 assignment, a fact that is important given that Florida's school voucher program requires a school grade of "F" as a condition of eligibility. Because no student in Gainesville is voucher-eligible (or would face voucher eligibility for at least two more years) under the state's system, we are free of the potentially confounding influence of the state's school voucher program on inferences about the effects of school letter grades *per se.* However, Gainesville is still a large enough metropolitan

4

area to have thousands of real estate transactions annually involving properties that sold in the several preceding years, so that there are sufficient repeat sales to generate reliable results.

Like previous work relating test scores to house values, we find that test scores are capitalized into housing prices. In addition, we observe that the market responds to the assignment of school letter grades. Indeed, we find that the distinction between a grade of "A" and a grade of "B" is valued in the housing market at over $9,000, holding constant other measures of school quality, neighborhood and property attributes, and similar-sized effects (even larger, in percentage terms) are observed surrounding the "B"-"C" distinction. The effects are stronger for larger houses and houses in higher-valued neighborhoods (though only for the "A"-"B" distinction), and are robust to various sensitivity checks.

The rest of this paper is organized as follows. The next section describes the school grading system in Florida, paying particular attention to the distinctions at the top end of the grading system. The third section describes the data and empirical method used in more detail, and the fourth section provides regression results. The fifth section concludes the paper.

2. The Florida School Accountability System

In November 1998, Florida voters elected Jeb Bush governor, and once he took office in January 1999 Governor Bush worked with the state legislature to implement his A+ education plan that coming summer. At its centerpiece is a system of accountability based largely on student test scores, in which each school would receive a letter grade ranging from "A" to "F." Schools on both ends of the spectrum are affected: students attending schools rated "F" in two years out of a four year window are eligible for school vouchers, or "opportunity scholarships," that can be

5

used to send a child to a private school or alternatively, that make a child eligible to transfer to a "C" or higher-rated public school in the same district or an adjacent district. Schools receiving grades of "A" (or in subsequent years, increasing their letter grades from year to year) are eligible for financial rewards totaling about $100 per pupil that can be spent for purposes such as hiring teacher aides or providing teacher bonuses, etc.[4]

Elementary school letter grades are currently based in large measure on aggregate student performance on three examinations: fourth grade reading performance on the Florida Comprehensive Assessment Test (FCAT), fifth grade FCAT mathematics performance,[5] and fourth grade performance on the Florida Writes! examination. The distinction between grades "C", "D", and "F" rests exclusively on this test performance. In order to attain a grade of "C",[6] at least 60 percent of test-takers must achieve at level two or above (on a five-point scale) on the FCAT reading test, at least 60 percent of test-takers must achieve at level two or above (also on a five-point scale) on the FCAT mathematics test, and at least 50 percent of test-takers must achieve at level three or above (on a six-point scale) on the Florida Writes! examination. If a school misses one or two of these thresholds it receives a grade of "D", while if it misses all three

---

[4]Despite being substantive in value, this $100 per pupil reward to the school is still less than two percent of per pupil school spending, so it is doubtful that it in itself should lead to large housing market effects.

[5]Prior to the 1999-2000 academic year, students took the FCAT reading assessment in grades four, eight and ten, the FCAT mathematics assessment in grades five, eight, and ten, and the Florida Writes! examination in grades four, eight, and ten. Beginning in 1999-2000, students now take reading and mathematics tests in grades three through ten.

[6]The basis for school letter grades have changed subtly between 1999 and 2000. Because no observations in our data set are affected by the 2000 letter grade assignment, all discussion of letter grades in this paper refer to the 1999 criteria and assignment.

of these thresholds it receives a grade of "F".

For a school to attain a grade of "A" or "B" additional criteria must be met. First of all, the test performance thresholds are higher: In order to attain a grade of "A" or "B", at least 50 percent of test-takers must achieve at level three or above on the FCAT reading test, at least 50 percent of test-takers must achieve at level three or above on the FCAT mathematics test, and at least 67 percent of test-takers must achieve at level three or above on the Florida Writes! examination. In addition, in order to achieve a grade of "A" or "B" the minimum ("C"-level) criteria must be met by test-takers in each of six subgroups of students: Economically Disadvantaged, Black, White, Hispanic, Asian, and Native American. Finally, a minimum of 90 percent of standard curriculum students (including language impaired, speech impaired, gifted, hospital homebound, and limited English proficient students who have been in an ESOL program for more than two years) must have taken the examinations in order for a school to earn a grade higher than "C".

To attain a grade of "A" rather than "B" four additional criteria must be met: (1) the percentage of students absent more than 20 days and the percent suspended must be below the state average; (2) at least 95 percent of standard curriculum students must be tested; (3) reading scores must have "substantially improved" (i.e., the fraction scoring level three or above on the FCAT reading test must have improved by two or more percentage points from the previous year, unless the fraction attaining this level is 75 percent or above already); and (4) math and writing scores must not have "substantially declined" (i.e., the fraction scoring level three or above on the FCAT mathematics test or Florida Writes! must not have fallen by five or more percentage points from the previous year.)

Of the 22 Alachua County regular elementary schools graded in 1999[7] two received grades of "A", five received grades of "B", six received grades of "C", and nine received grades of "D".[8] The higher-rated schools tend to be larger, however, and have disproportionately larger numbers of single-family homes in their catchment areas. Among houses sold between 1995 and 2000, 27.4 percent are zoned for an "A" school, 36.3 percent are zoned for a "B" school, 17.8 percent are zoned for a "C" school, and 18.4 percent are zoned for a "D" school. This paper focuses principally on the six "A" and non-rural "B" elementary school zones, as the distinctions between "A" and "B" graded schools and their communities tend to be finest, thus allowing a cleaner test. In addition, we emphasize the "B"-"C" distinction, where the differences between schools and their communities appear to be greater than the "A"-"B" difference, but where school grades are based on similarly arbitrary criteria. However, the paper still considers the entire metropolitan area, and each regression includes an indicator for the "C"-"D" grade distinction; while these point estimates are not reported in the tables, they are occasionally described in the text.

Table 1 presents some descriptive data on the six elementary schools graded "A" or "B" in the 1999 assignment of school grades. We observe that the "A" and "B" schools in Gainesville, and particularly the "A" schools and "B" schools Finley and Littlewood, appear quite similar in terms of average fourth grade test scores in 1999, a pattern that is present the year before as well.

---

[7]Two charter schools also received letter grades, but are excluded from this analysis because they do not draw from a zoned catchment area..

[8]Two rural elementary schools, High Springs Elementary School and Waldo Community School, have no eligible platted subdivisions in their catchment areas. High Springs earned a grade of "B" and Waldo earned a grade of "C" in the 1999 school grades.

In addition, the schools appear similar along most of the dimensions used to distinguish "A" from "B" schools. In all six cases, at least 95 percent of eligible students took the examination, the excessive absenteeism rate was less than the state average, and there was no "substantial decline in mathematics." Finley missed an "A" by one criterion--the percentage passing the writing examination fell by six percentage points from 1998 to 1999, while five percentage points was the maximum allowable decline to qualify for "A" status. Littlewood and Hidden Oak each missed an "A" by two of the six criteria--both suspended more than the state average for elementary schools, and neither experienced "substantial increases in reading." Both, however, would have passed the latter criterion if the threshold for not requiring reading improvement from cohort to cohort were 70 percent, rather than 75 percent. (Likewise, Talbot, one of the "A" schools would have failed to receive an "A" if the threshold for not requiring reading improvement were higher, as Talbot's reading score did not increase from one cohort to the next.) Wiles is the "B" school least similar to the "A" schools, missing an "A" by three criteria--the two missed by Littlewood and Hidden Oak, as well as having a substantial reduction across cohorts in the fraction passing the writing examination, from 90 percent to 81 percent. Therefore, while the "A" schools appear slightly better than the "B" schools along some observable dimensions, the distinctions are not great.

Moreover, one could easily design a school grading system in which some of the "B" schools appear superior to some of the "A" schools. For instance, consider test score gains on the ITBS from third grade in 1998 to fourth grade in 1999, that is, following the *same cohort*.[9]

---

[9]The state did not have the ability to follow the same cohort of students from one grade to the next in 1999, so could not have used these data to construct school grades.

This is an important alternative grading system, because it represents a criterion similar to the long-run school grading system to be implemented in Florida, as well as being similar to a characteristic of President Bush's education proposals (requiring annual testing and assessment of schools based on year-to-year changes.) While all "B" schools improved in reading, and Finley and Littlewood improved in mathematics, Norton did not improve in reading and both "A" schools fell in mathematics. Looking at outcomes other than test performance, if one were to rate schools on the basis of whether their ratio of excessive absences to student mobility rate is less than the state average, all of the "B" schools would satisfy this criterion, but Talbot, one of the "A" schools, would not. In sum, the distinction between a grade of "A" and a grade of "B" is necessarily imperfect and subject to judgment calls, and regardless of how valid and careful the grade assignment process is, one can construct equally arbitrary (but equally valid) grading criteria that reverse some of the grading distinctions. It is not clear that the grades provide "good" information. Nevertheless, they are readily available and households may use them as sufficient statistics for school quality. The question of interest is to what degree the market responds to this new information.

A preliminary pass at the data suggests that the market responds considerably to this new information. Table 2 presents suggestive evidence to this effect--we identify the neighborhood with an average 1999 sales price closest to $150,000 in each elementary school catchment area (about an 80th percentile neighborhood in "A" or "B" catchment areas) and the neighborhood with an average 1999 sales price closest to $100,000 in each elementary school catchment area (about a 35th percentile neighborhood in "A" or "B" catchment areas.) The only restriction we place on neighborhood selection here is that the neighborhood must have had at least three sales both

before and after the letter grades were released. The table then reports the percentage change in the average house price in that neighborhood in the months of 1999 following the grade reporting versus the 1999 months prior to the grade reporting.

Looking first at the relatively expensive neighborhoods, we observe that house prices in the neighborhoods in "A" rated school districts increased substantially from the first part to the last part of 1999. Robin Lane, the neighborhood in the Norton district, increased in value by 11.2 percent and Blues Creek 4B, the neighborhood in the Talbot district, increased in value by 20.4 percent. These reported values are not raw mean differences, but rather are differences in the residual of a regression of house prices on month dummies and neighborhood dummies, using 1999 data. Therefore, any idiosyncracies that are common to all properties in a neighborhood in 1999 or all properties in Gainesville in a given month are partialed out. In contrast, the $150,000 neighborhoods in three of the four "B" catchment areas decreased in value, and the fourth, Haile Plantation unit 12, phase 2/3, increased in value by only 2.6 percent. This pattern is present, though not as dramatic, when looking at less expensive ($100,000) subdivisions. Both "A" subdivisions increased in value, and, as before, three of the four "B" subdivisions decreased in value, but the fourth, Hamilton Heights, increased in value by 11.2 percent. In sum, these patterns provide suggestive evidence of a relationship between letter grade assignment and property values.

The letter grade system described above is not the first time the state categorized schools. Prior to assignment of letter grades, the Florida Department of Education identified all schools in the state as "level 1" through "level 4" for the 1996-97 and 1997-98 school years. Unlike letter grades, the state's levels did not discern at all among the top schools in the district. Indeed, the

11

vast majority of Alachua County schools were considered level 4 (the highest category) in this grading system: 89.0 percent of real estate transactions in our sample are in the zones for schools graded level 4 in 1998, while 8.6 percent are in level 3 zones and 2.5 percent are in level 2 zones. The same basic pattern is true for 1997, except that Duval Elementary, the only level 2 school in 1998, was graded as a level 1 school in 1997 and Newberry and Shell Elementaries, both graded level 4 is 1998, were graded level 3 in 1997. Every grade "A" or "B" school in 1999 was a level 4 school in 1998, 88 percent of real estate transactions in 1999 "C" school zones were in 1998 level 4 school zones, and 53 percent of real estate transactions in 1999 "D" school zones were in 1998 level 4 school zones. Because this earlier grading system overwhelmingly gave schools its highest level, we control for these earlier grades in our analysis but focus our attention on the 1999 grading system, which provided distinction at the top and middle of the school distribution.

3. Data and Method

A more thorough investigation of this question requires that we reduce the probability that unobserved factors are driving the patterns alluded to above. Our solution is to estimate a highly parameterized model that controls for a series of fixed effects:

$$\text{price}_{isnmy} = \alpha_i + \beta_m + \delta_{ny} + \eta\text{test}_{smy} + \sum_g \theta_g \text{grade}_{smy} + \sum_l \varsigma_l \text{level}_{smy} + \varepsilon_{isnmy} \text{ ,}$$

where price_{isnmy} is the price of house i in neighborhood n in school catchment area s in month m in year y. Our primary variables of interest, grade_{smy}, are a series of dummy variables reflecting the assignment of each particular letter grade to school s in 1999. These variables take on a value of

12

zero prior to July 1999,[10] and a value of one following that time if school s was given the relevant grade. We also control for a series of similar dummy variables ("level" in the above expression) reflecting the "level 1-4" distinction assigned to schools mentioned in the preceding section.[11] Because there is little variation in this categorization system, we focus our attention on the 1999 grading system, and only occasionally report the point estimates of the earlier grading system in the text. We include these last variables in the model for completeness, but the results presented below are unaffected, either in terms of magnitude or statistical significance, by the inclusion or exclusion of these "level" classifications in the model.

Of secondary interest to us is the coefficient on $test_{smy}$, which represents the most recently reported average test score of school s, as of month m and year y. The model also controls for a series of fixed effects. The fixed effects α_i capture any characteristics about property i that did not change over our time period, 1995 to 2000. This would include any characteristics such as the square footage, number of bedrooms, quality of the trees, general condition of the house, and lot size, among other features. Because it is possible that a house might change substantially, sensitivity tests reported later control for the value of any changes to the house that require a building permit (e.g., addition of a pool or screened enclosure) between house sales and the results remain substantively unchanged; hence, we are confident that factors that are special about

---

[10]We assume a one-month lag before new information on test scores or school grade assignment is reflected in house prices, given the typical duration between agreement and closing on real estate transactions.

[11]As with the school letter grade variables, these dummy variables take on a value of zero during the time in which the "level" classification regime is not in place. We also estimated models that assign the most recent "level" classification to schools even after the letter grade regime began; the letter grade results are nearly identical to those reported herein.

a house are adequately controlled for in this model. The fixed effects δ_{ny} capture any

characteristics about all properties in a given platted subdivision that change together over time.

These fixed effects are defined at the neighborhood-calendar year level. This might include, for

example, general neighborhood deterioration or beautification, the addition of a neighborhood

pool or playground, increased traffic, or a rash of burglaries in a neighborhood. Finally, the

housing market in Gainesville is highly seasonal; therefore, we control for month-of-year dummies

β_m to reflect factors that are common to all sales in, say, January, regardless of the year.[12] In sum,

for any remaining omitted variable to explain our findings, they would need to change

systematically within a neighborhood halfway through a year, while also covarying with changes

in the most recent publicly released information about the school which the neighborhood's

children attend.

To implement this empirical strategy, we employ house price data from the Alachua

County Tax Collector's Office. Our sample consists of every sale of a single-family house in a

platted neighborhood in Alachua County, Florida, between January 1, 1995 and July 10, 2000.[13]

We use only single-family houses for purposes of data comparability, and we use only houses in

platted subdivisions (which, as mentioned above, accounts for virtually the entire urban and

suburban portion of the Gainesville metropolitan area) because otherwise we have no distinct

measurement of "neighborhood." In addition, most subdivisions are reasonably homogeneous in

---

[12]There is a significant seasonal cyclicality in house prices in Gainesville. House prices tend to be highest in the summer months and lowest in the winter months, according to these estimated month effects.

[13]The end of the sample period is arbitrary; it is the last day of sales data available at the time of the data extract provided for us by the county.

terms of age, size and quality of house and lot, making them a desirable measurement of neighborhood for the purposes of this analysis. The county database includes both neighborhood and parcel identifiers, making it possible to match properties to neighborhoods, and to match sales of the same house over time.

Our dependent variable is the dollar value of the real estate transaction. While the dependent variable in the regressions reported in the paper is expressed in levels, we have estimated all models with the dependent variable expressed in logs, and the results are similar to those presented herein. We attempt to exclude token sales (e.g., intra-family sales) and likely typographical errors by excluding from our sample all observed house sales for less than one-quarter of the property's assessed value, or for more than ten times the property's assessed value. These are admittedly arbitrary thresholds, and the results are robust to changing these thresholds. These data restrictions leave us with 12,715 real estate transactions over the sample period. The mean house price of these transactions is $106,994 with a standard deviation of $66,365. The median house price is $94,480. Ten percent of houses sold for more than $185,707, while ten percent of houses sold for less than $38,649. These house prices are deflated to 2000 dollars using the average sale price in a month-year combination as a deflator, but all results are unchanged if we utilize nominal sales prices instead. Unsurprisingly, houses zoned for highly-graded schools sell for more than houses zoned for lower-graded schools, although this relationship is not monotonic: the average house sale in "A" school zones is $115,543, as compared to $128,336 in "B" zones, $96,941 in "C" zones, and $61,874 in "D" zones.

Real estate transactions are mapped to school zones using data from the School Board of Alachua County. Our sample period overlaps one minor rezoning that occurred concurrently with

15

the opening of Chiles Elementary in summer 1999; for transactions prior to July 1999 we use the

pre-rezoning school zone, while for transactions after that point we use the post-rezoning school

zone. This rezoning only affects two percent of our sample,[14] and our results are virtually

unchanged if we exclude the affected properties from the analysis. Elementary school zones

generally do not split platted subdivisions, but about two percent of houses in Gainesville are

located in neighborhoods split into multiple school attendance zones. While these split

neighborhoods have the potential to provide additional variation for our model, in fact our results

are invariant to including or excluding neighborhoods split into multiple attendance zones.

Test score data come from the School Board of Alachua County, and are assigned to real

estate transactions beginning the month following the public release of these data, generally in

May (but varying somewhat from year to year.) Therefore, within any calendar year houses are

assigned two different sets of test scores: one from before the release of updated information and

another from after the release of updated information. We include separately the mean fourth

grade Iowa Test of Basic Skills mathematics and reading scores for each elementary school in the

district. For data prior to June 1996 we use the analogous scores on the California Achievement

Test, which was used prior to the 1996-97 academic year; the rank order of schools is basically

the same across the two tests, which in unsurprising, given that both tests are nationally norm-

referenced.[15] The mean school average mathematics test score over the course of our sample is

68.8 with a standard deviation of 12.9, while the mean school average reading test score over the

---

[14]As mentioned above, properties zoned for Chiles Elementary are excluded from the
analysis, as we have no test scores or letter grades for Chiles during the sample period.

[15]Our estimates of the effects of school grading are stronger if we restrict our analysis to
the years with the same test (Iowa), and are reported later in the paper.

course of our sample is 64.2, also with a standard deviation of 12.9. The mean mathematics test score for properties in "A" graded school zones over the course of our sample is 77.5. The comparable figures are 74.9 for "B" zones, 64.7 for "C" zones, and 47.7 for "D" zones. A similar pattern emerges with respect to reading scores across the letter grade levels.

Because of our fixed effects specification, we do not control individually for property or neighborhood characteristics. However, to fix ideas, the mean heated square footage of a property in our sample is 2167 square feet with a standard deviation of 848 square feet. One percent of our sample is a five-bedroom house, 21 percent are four-bedroom houses, and 67 percent are three-bedroom houses. Nine percent have three or more bathrooms, while 15 percent have one or one and a half bathrooms. The average house was 16 years old at the time of sale, with a standard deviation of 13 years.

4. Results

Table 3 presents the primary results of the paper. Column 1 of the table reports the results of a model that excludes school letter grade effects (or the effects of the prior school level assignment) to get a sense of the degree to which test scores are capitalized in house prices. We observe that even in this highly parameterized model, mathematics test scores appear to be valued by the housing market--the estimated effect is significantly different from zero at conventional significance levels--and, while not statistically significant at conventional levels, reading test scores also appear to be independently valued by the market.[16] The results are reasonably large in

---

[16]Standard errors are adjusted for clustering at the school attendance zone x time level to reflect the fact that the variables of interest only vary at the school attendance zone level at any given time.

magnitude, as well. For instance, the results suggest that the differences in math and reading

scores between the typical "A" school and the typical "B" school are valued at $1,492, and the

differences in math and reading scores between the typical "B" school and the typical "C" school

are valued at $5,435. These results are similar in magnitude to those reported by Black (1999),

and provide additional evidence that the housing market reflects changes in information

concerning test scores at neighborhood schools.

More to the point of this paper are the results presented in the second column of Table 3.

This column reports the results of the specification that includes the effects of grade assignment in

1999, as well as the effects of the previous "level" assignments in 1997 and 1998. We first

observe that the mathematics test score result reported in column 1 is strengthened by the

inclusion of the state-assigned classification variables in the model, while the reading test score

coefficient remains statistically insignificant and actually changes sign. More importantly, for the

purposes of this paper, the evidence suggests that state assignment of a grade of "A" versus a

grade of "B" is valued at $9,179, and statistically distinct from zero at the seven percent level.

Given that the average house price in neighborhoods zoned for "A" and "B" schools is $122,830,

this suggests that the market value of an "A" versus a "B" is about seven percent.[17] Because "C"-

zoned houses are less expensive still, the "B"-"C" distinction results are even larger in percentage

terms. While not reported in the table, the results of the other coefficients suggest that the

classification of schools into levels in 1997 and 1998 had some market effects as well. For

---

[17]The results are very similar if the dependent variable is expressed in logs. The coefficient estimate on assignment of a "B" grade is -0.06, suggesting that a grade of "A" is associated with a six percent increase in house prices relative to a grade of "B". The coefficient estimate on assignment of a "C" grade (relative to a "B") is -0.12.

instance, the value of classification as "level 4" (the highest level) rather than "level 3" is estimated to be \$6,728, and is significant at the one percent level. In sum, the results suggest that the housing market reflects the assignment of school report card grades.

4.1 Heterogeneity in Effect Size

The remaining columns of Table 3 present the results of regressions in which we further interact the school grade variables and prior level assignment variables with a specific house characteristic--specifically, in turn, we consider the differential effect of letter grade assignment on houses based on square footage, average neighborhood price, and relative square footage in a neighborhood. That is, we estimate variants of the model:

$$price_{isnmy} = \alpha_i + \beta_m + \delta_{ny} + \eta test_{smy} + \sum_g \theta_g grade_{smy} + \sum_g \gamma_g grade_{smy} X_i + \zeta X_i + \varepsilon_{isnmy} \ ,$$

where all notation is as before, and as before, the prior level assignment variables are included in the model but omitted from the equation above. The variable X_i represents, depending on model specification, the house's square footage, the average neighborhood price, or the ratio of house square footage to the average square footage in the neighborhood. Of course, the parameter ζ is not independently estimated, but is rather subsumed into the property fixed effect α_i but we include X_i in the equation for completeness.

Column 3 of Table 3 resents the results of the first of these models. We observe that the negative effect of assignment of a "B" (or, alternatively, the positive effect of "A" assignment) increases, albeit only marginally significantly (11 percent) with square footage.[18] While only

---

[18]This negative interaction term is considerably more statistically significant when the dependent variable is expressed in logs. The point estimate of -0.000223 (implying that the effect increases in magnitude by two percent for each 100 square foot increase) is significant at the one

marginally significant, this interaction term is large in magnitude. Interpreting the coefficients together, the results suggest that the value of an "A" versus a "B" for a house the size of a typical two-bedroom house in Gainesville (1,233 square feet) is -$1,022 (p=0.85). For a house the size of a typical three-bedroom house (1,663 square feet) the effect is $8,507 (p=0.07) and for a house the size of a typical four-bedroom house (2,311 square feet) the effect is $22,869 (p=0.06). We find similar results if, instead of interacting square footage with the policy variables, we simply estimate the initial regression equation separately for above-median and below-median houses, classified in terms of square footage. For below-median square footage houses, we estimate that an "A" is worth -$2,725 (p=0.46) relative to a "B" while for above-median square footage homes, we find that an "A" is worth $19,709 (p=0.02) relative to a "B". Given that the average above-median-sized house sells for $142,970, this suggests that for above-median houses, the estimated effect of an "A" is just under 14 percent of a typical house price.   While we find that the mean effect of a grade of "C" versus a "B" is significant, we find no evidence that a grade of "C" differentially affects house prices among houses of different sizes.

In other results not presented in the tables, we interact the grade variables with number of bedrooms. In this model, we find that an "A" is worth $10,225 (p=0.38) more than a "B" for three-bedroom houses relative to two-bedroom houses, and the differential valuation of an "A" versus a "B" is $23,451 (p=0.03) for four-bedroom houses relative to three-bedroom houses. These results, taken together with those mentioned above, jointly suggest that the effects of an "A" versus a "B" are considerably greater in the market for larger houses than they are in the market for smaller houses. This result seems sensible, given that families with children are

percent level.

probably more likely to reside in larger houses.[19]

The fourth column of Table 3 reports the results of a similar exercise, except this time interacting the policy variables with average house price in the neighborhood. We find that the estimated effects of an "A" versus a "B" increase strongly in magnitude and statistical significance with neighborhood value. The results suggest that the effect of an "A" increases by $3,999 for each $10,000 increase in neighborhood value, and is positive for any neighborhood with a value greater than $8,563.[20] If instead we were to estimate the effects of an "A" versus a "B" in separate regressions for low-value (below median) and high-value (above median) neighborhoods, we estimate the value of an "A" to be -$1,513 (p=0.70) for below-median neighborhoods and $36,975 (p=0.00) for above-median neighborhoods. Therefore, the evidence suggests that the effects of school grading (or at least the distinction between an "A" and a "B") are concentrated in the higher-value neighborhoods.[21] As with the square footage results, we find no evidence that similar differential effects occur at the "B"-"C" distinction.

---

[19]Supporting evidence for this claim is found in the data. Houses where children currently enrolled in Alachua County public elementary schools live are on average eight percent larger than houses where no children currently enrolled in Alachua County public elementary schools live. Given that the comparison group includes households with private school attendees and public secondary school attendees as well, the difference between houses with children and houses without children is probably even larger.

[20]The coefficient on the interaction term in the logarithmic specification is -0.0029 (p=0.00), suggesting that the effect of an "A" increases by 2.9 percentage points for each $10,000 increase in neighborhood value.

[21]We note that the coefficient on reading test scores becomes much larger (and negative), and approaches statistical significance, in this specification and the specification reported in column 5. We have no good explanation for this finding, except that reading and math scores tend to be relatively collinear, and the two effects may cancel out on net once we include the interactions between grade and neighborhood (or house) characteristic.

The fifth column of Table 3 reports the results of a specification that interacts the policy variables with a variable reflecting the ratio of the house's square footage to the average square footage in the neighborhood. Here, we seek to discover whether the effects of school letter grade assignment have a differential effect on houses within the same neighborhood. We observe that, indeed, the effect of a grade of "A" versus "B" is statistically and economically significantly greater for relatively large houses within a neighborhood.[22] For instance, we find that the effect of an "A" is $6,930 greater for a house whose square footage is ten percent above the neighborhood average than for a house whose square footage is ten percent below the neighborhood average. As with the other two sets of interactive results, we find no evidence that similar differential effects occur at the "B"-"C" distinction. Taken together, these results, and those presented above, indicate that the effects of school grading are heterogenous, and that the assignment of high-end school grades particularly affects the higher end of the market, and that small houses and low-value neighborhoods do not appear affected. However, these results only appear to occur at the "A"-"B" distinction, and not at lower distinctions of school grading.

4.2 Sensitivity Analysis

To gauge the sensitivity of the results presented in Table 3, we perform a number of sensitivity tests. (Here, for purposes of presentation, we focus only on the "A"-"B" distinction, as the "B"-"C" distinction continues to follow similar patterns.) Our first set of sensitivity tests

---

[22]The results are similar in the logarithmic dependent variable specification. The coefficient estimate on the interaction between "B" grade and the ratio of house size and average size in the neighborhood is -0.250 (p=0.00), implying that the effect of an "A" is five percentage points greater for a house whose square footage is ten percent above the neighborhood average than for a house whose square footage is ten percent below the neighborhood average.

concerns what happens when we change some of the conditions under which we draw our sample. First, recall that, to eliminate token sales and probable typographical errors, we eliminate real estate transactions of less than one-quarter of assessed value, or greater than ten times assessed value. But perhaps these thresholds are not sufficiently restrictive. Therefore, the second row of Table 4 presents the results of a model in which we require that sales must be at least half of assessed value to be included in the sample, and the third row of Table 4 presents the results of a specification in which we require that sales must be less than five times assessed value to be included in the sample. (The first row of the table repeats the full-sample specification from Table 3.) We observe that the results are very similar when we impose tighter restrictions on sample inclusion. Our next set of sensitivity checks concerns whether our sample start date affects our results; therefore, rows four and five of Table 4 present the results of specifications in which we start our sample in January 1996 or January 1997, respectively. We observe that, in fact, our estimated effects of a grade of "A" relative to a "B" are conservative in magnitude as a result of our starting our analysis in 1995: as we bound the sample to a set of transactions that occur closer to the policy change, the results only get stronger.

We next seek to determine whether certain small sets of observations are driving our results. Row six of Table 4 presents the results of a model that excludes from our analysis the two percent of properties located in platted subdivisions split by elementary zone lines, and row seven of the table presents our results when we exclude the two percent of properties rezoned in 1999 when Chiles Elementary opened. The results in both cases are very similar to our full-sample model. Rows eight and nine of the table present the outcomes of specifications in which we exclude, in turn, the bottom five percent of neighborhoods (ranked by average sales price) and

23

the top five percent of neighborhoods. We find that in both cases, the effect of an "A" versus a

"B" remains statistically significant and at least as strong in magnitude as in our full sample, and

that our results are strengthened when we exclude the very most expensive neighborhoods from

our model. This may be due to the likelihood that children in these neighborhoods tend to attend

private schools.

The final four rows of Table 4 present the results of sensitivity checks in which we explore

the potential presence of some omitted variable that might be driving our results. Row ten

displays the estimated effects of a grade of "A" if the grades had been assigned in **May 1998**

rather than in the actual month, May 1999.[23] If we were to find similar types of results in this

specification, it would raise concerns that some factor other than the assignment of grades led to

the results presented above. However, we observe that this is not the case: in this "pseudo-grade"

specification, we find no effect on an "A" relative to a "B". The estimate effect is $335.31

(p=0.93). Row eleven performs a similar analysis, this time assuming the grades were released in

November 1998, when Jeb Bush was elected governor. Again, we find effects relatively small in

magnitude ($2,760.53) and statistical significance (p=0.48), suggesting little or any anticipation by

the market of the actual grades that were released, even after Bush was elected and school

grading became much more likely.

Row twelve presents a different type of sensitivity check. Here, we assume that the

announcement of school grades *per se* did not matter, but rather made the real estate market

aware of the data that were used to generate the school letter grades. For instance, while changes

---

[23]As above, we assume a one-month lag between grade assignment and real estate
transactions, given the length of time it takes to close on a house.

in year-to-year changes in a school's test scores were published in the newspaper for years prior to imposition of the letter grade regime, perhaps the market did not fully take these changes into account until the system placed them more into the spotlight. Alternatively, perhaps the school grades provide information about the *distribution* of performance within a school, which has been argued (Brown and Saks, 1975) to matter substantially to parents and schools. To gauge the degree to which the apparent market response to school grading is truly due to the publicizing of other important school evaluation criteria, the specification reported in row twelve controls for all of the variables used to construct the school letter grades, interacted with a dummy variable representing that school letter grades had been imposed. That is, we "activate" these variables at the same time we "activate" the school letter grades. We observe that, rather than eliminating the independent school grade effect, the estimated effect of an "A" versus a "B" gets even larger in magnitude--an almost implausibly large $39,933.82 (p=0.00). While we hold little stock in the magnitude of this result, it suggests compellingly that the imposition of school letter grades had a significant effect independent of the variables used to define the letter grades. Also, in related sensitivity exercises not reported in the table, we find that including over the entire time series all of the variables that ultimately became part of the accountability system (e.g., attendance, suspension rates, and changes in test scores across cohorts) only strengthens our estimated effects of school grade assignment.

The final row of Table 4 presents the results of a model in which we control for all observed changes to each house between sales. Here, we control for the dollar value of new improvements, as reported by the tax assessor's office. We observe that, once we control for these improvements, our results get slightly stronger in magnitude as well as more significant

25

statistically. In sum, it is difficult to think of any omitted variables that might be driving our results reported above.

5. Conclusion

This paper provides the first evidence of the effects of school grade assignment on the housing market. Our results suggest that the housing market responds significantly to the new information about schools provided by these "school report cards," even when taking into consideration the test scores or other variables used to construct these same grades. These results suggest that innocuous-seeming school classifications may have large distributional implications, and that policy-makers should exercise caution when classifying schools.

One should be careful, however, in interpreting these results. While we find large effects of school grading in the year following grade imposition, we have no way of knowing whether these market effects are permanent or transitory. The early evidence, however, suggests that the results may be transitory: when we estimate a model in which we interact the school grades with the number of months since grade imposition, we observe that the effects of an "A" grade versus a "B" grade are strongest in the months immediately following the grade implementation, and decrease in magnitude in each subsequent month. Indeed, in results not presented in the tables, we find that the effect of an "A" versus a "B" in the month following imposition is estimated to be $21,229 (p=0.01) and falls by $2,397 (p=0.01) per month in each subsequent month. Hence, while it is too early to tell, the results suggest that the market appears to be returning to its pre-shock condition over time.

In addition, it is also possible that in the long run, the introduction of an accountability

system could lead to an increase in housing prices across the board if the new system leads to improved school efficiency and performance. This, however, is not an obvious consequence of higher global performance unless improved performance influences the overall demand for land in the metropolitan area. There could also be additional distributional implications if a grading system impacts, for instance, the willingness of suburban schools to accept low-income students. Therefore, while this study sheds light on the potential short-run distributional consequences of school grading policies, the long-run welfare and distributional implications of the introduction of a system of school grading system are still unknown.

References

Black, Sandra. 1999. "Do Better Schools Matter? Parental Valuation of Elementary Education." *Quarterly Journal of Economics*, 577-600.

Bogart, William and Brian Cromwell. 1997. "How Much More is a Good School District Worth?" *National Tax Journal*, 215-232.

Bogart, William and Brian Cromwell. 2000. "How Much is a Neighborhood School Worth?" *Journal of Urban Economics*, 280-305.

Bradbury, Katherine, Christopher Mayer, and Karl Case. 1999. "Property Tax Limits and Local Fiscal Behavior: Did Massachusetts Cities and Towns Spend Too Little on Town Services Under Proposition 2 1/2?" Forthcoming, *Journal of Public Economics.*

Brown, Byron and Daniel Saks. 1975. "The Production and Distribution of Cognitive Skills Within Schools," *Journal of Political Economy*, 571-593.

Brunner, Eric, Jon Sonstelie, and Mark Thayer. 2000. "Capitalization and the Voucher: An Analysis of Precinct Returns from California's Proposition 174." Working paper, San Diego State University.

Downes, Thomas and Jeffrey Zabel. 1997. "The Impact of School Characteristics on House Prices: Chicago 1987-1991." Working paper, Tufts University.

Hayes, Kathy and Lori Taylor. 1996. "Neighborhood School Characteristics: What Signals Quality to Homebuyers?" Federal Reserve Bank of Dallas *Economic Review*, 2-9.

Table 1
Comparing "A" and "B" Schools in Gainesville, Florida

School:	Schools graded "A" in May 1999		Schools graded "B" in May 1999			
	Norton	Talbot	Finley	Hidden Oak	Little-wood	Wiles
Average Iowa Test of Basic Skills scores, grade 4						
Average ITBS grade 4 reading score, 1999	78	78	78	70	75	69
Average ITBS grade 4 math score, 1999	78	82	77	75	75	73
Average ITBS grade 4 reading score, 1998	75	74	74	75	73	70
Average ITBS grade 4 math score, 1998	77	79	78	77	80	75
Criteria for distinguishing "A" from "B" schools in state accountability system, 1999						
Excessive absentee rate less than state average	yes	yes	yes	yes	yes	yes
Suspension rate less than state average	yes	yes	yes	no	no	no
At least 95 percent took examinations	yes	yes	yes	yes	yes	yes
Substantial reading increase across grade 4 cohorts on FCAT, 1998-99	yes	yes	yes	no	no	no
No substantial math decrease across grade 4 cohorts on FCAT, 1998-99	yes	yes	yes	yes	yes	yes
No substantial writing decrease across grade 4 cohorts, 1998-99	yes	yes	no	yes	yes	no
Three other plausible criteria for distinguishing schools						
Gain from grade 3 to grade 4 reading (ITBS), same cohort	no	yes	yes	yes	yes	yes
Gain from grade 3 to grade 4 math (ITBS), same cohort	no	no	yes	no	yes	no
Ratio of excessive absence rate to mobility rate less than state average	yes	no	yes	yes	yes	yes

Table 2

Suggestive Evidence: Percentage Change in House Prices* Before vs. After 1999 Grade Reports, Representative Neighborhoods

Elementary school	Neighborhood with average sales closest to $150,000		Neighborhood with average sales closest to $100,000	
	Neighborhood name	Percentage change	Neighborhood name	Percentage change
Schools graded "A" in May 1999				
Norton	Robin Lane	11.2%	Rainbows End	3.7%
Talbot	Blues Creek 4B	20.4	Mile Run 3, unit F-2	4.5
Schools graded "B" in May 1999				
Finley	Brywood	-8.2	Anglewood	-21.1
Hidden Oak	Eagle Point	-26.7	Hamilton Heights	11.2
Littlewood	Rock Creek	-3.0	Westwood Estates	-8.0
Wiles	Haile Plantation unit 12, phase 2/3	2.6	Haile Plantation unit 5/6, phase 2	-11.1

Notes: House prices are the residuals of a regression of house prices on month-of-year and neighborhood dummy variables for 1999 sales.

Table 3

Estimated Effects of Test Scores and School Letter Grades on House Prices

Specification:	(1)	(2)	(3)	(4)	(5)
Variable					
Average reading test score in 4th grade	225.83 (p=0.20)	-160.49 (p=0.41)	-70.04 (p=0.72)	-465.95 (p=0.12)	-528.44 (p=0.11)
Average math test score in 4th grade	295.92 (p=0.05)	406.23 (p=0.01)	466.68 (p=0.00)	512.06 (p=0.02)	532.78 (p=0.01)
Effect of "B" grade **(compared to A)**		-9179.25 (p=0.07)	28133.33 (p=0.18)	3429.02 (p=0.45)	-4883.56 (p=0.39)
Effect of "C" grade **(compared to B)**		-9875.68 (p=0.06)	7121.87 (p=0.63)	-3200.31 (p=0.58)	-5322.70 (p=0.43)
Effect of B grade x square footage			-22.02 (p=0.11)		
Effect of B grade x average neighborhood price (1000s)				-399.87 (p=0.00)	
Effect of B grade x relative square footage (neighborhood mean=1)					-34650.71 (p=0.00)
Effect of C grade x square footage			-9.44 (p=0.31)		
Effect of C grade x average neighborhood price (1000s)				13.28 (p=0.86)	
Effect of C grade x relative square footage (neighborhood mean=1)					6272.40 (p=0.46)

Note: Regressions control for parcel fixed effects, neighborhood x year fixed effects, and month fixed effects, as well as (except for specification 1) controls for grades of C vs. D and, prior to school grading, levels 1, 2, and 3. In addition, except for specification 1, regressions control for dummy variables indicating the school classification regime present at the time of sale. Specifications involving interaction terms (3-5) include interactions between letter grades and the variable in question (e.g., square footage) and between school level assignment and the variable in question. Standard errors are adjusted for time-specific elementary zone-level clustering.

Table 4
Sensitivity Testing: Estimated Effects of a Grade of "B" Versus "A"

Sensitivity check	Estimated effect of "B" versus "A"
Sensitivity to Changing Sample	
(1) Full sample specification (copied from Table 3, column 2)	-9179.25 (p=0.07)
(2) Restricting sales to greater than 50 percent of assessed value	-8194.54 (p=0.07)
(3) Restricting sales to less than 500 percent of assessed value	-9277.71 (p=0.07)
(4) Restricting sample to 1996-2000 observations	-12136.64 (p=0.03)
(5) Restricting sample to 1997-2000 observations	-20095.93 (p=0.00)
(6) Eliminating multi-zone neighborhoods from sample	-9295.76 (p=0.07)
(7) Eliminating properties rezoned in 1999	-9172.24 (p=0.07)
(8) Eliminating bottom five percent of neighborhoods	-9412.11 (p=0.07)
(9) Eliminating top five percent of neighborhoods	-20359.40 (p=0.00)
Specification Checks	
(10) "Pseudo-grade" assigned in May 1998	-335.31 (p=0.93)
(11) "Pseudo-grade" assigned in November 1998	-2760.53 (p=0.48)
(12) Including all variables used in grade assignment as covariates in the model	-39933.82 (p=0.00)
(13) Controlling for improvements occurring between sales	-9597.19 (p=0.04)

Note: Regressions control for parcel fixed effects, neighborhood x year fixed effects, and month fixed effects, as well as controls for grades of C and D and, prior to school grading, levels 1, 2, and 3. In addition, regressions control for dummy variables indicating the school classification regime present at the time of sale. Specification 12 controls for FCAT reading, math, and writing scores, changes in FCAT reading, math, and writing scores, excessive absenteeism rates, and suspension rates (the variables used to determine grade assignment) simultaneous with the assignment of school grades. Specifications 10 and 11 control for "pseudo-grades" instead of the actual grade assignments. Standard errors are adjusted for time-specific elementary zone-level clustering.