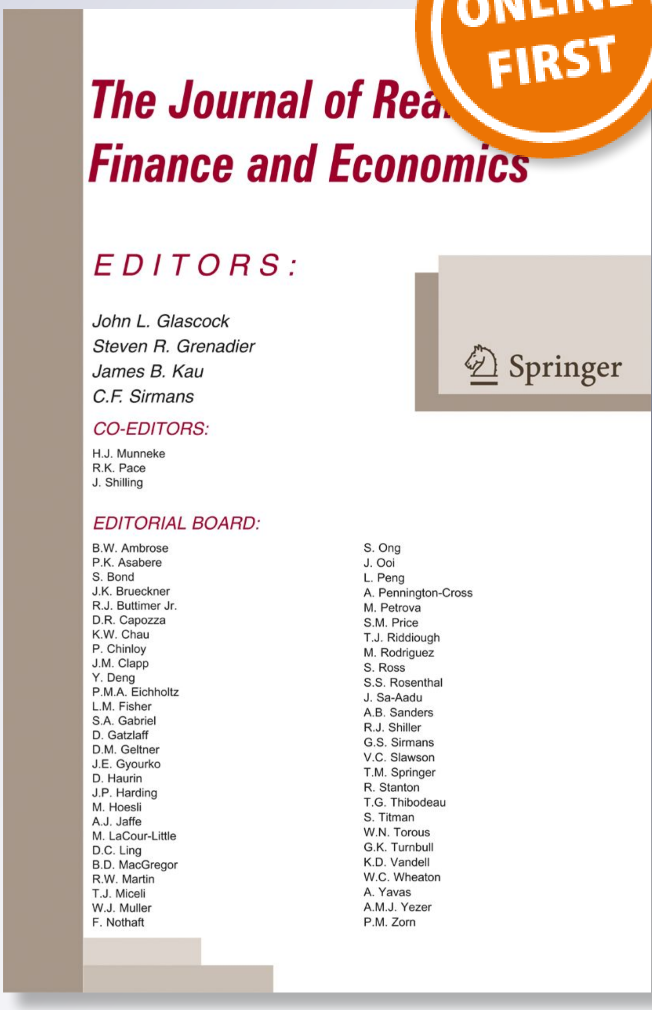
Affordable Housing and the Socioeconomic Integration of Elementary Schools

Keith Ihlanfeldt & Tom Mayock

The Journal of Real Estate Finance and Economics

ISSN 0895-5638

J Real Estate Finan Econ
DOI 10.1007/s11146-018-9665-0



Springer

J Real Estate Finan Econ
https://doi.org/10.1007/s11146-018-9665-0

Affordable Housing and the Socioeconomic Integration of Elementary Schools

Keith Ihlanfeldt[1] · **Tom Mayock**[2]

Abstract Children from poor families achieve more academically if they are enrolled in schools that are socioeconomically integrated, but low-income students are increasingly attending schools characterized by high concentrations of poverty. Providing more housing opportunities for low-income families within the attendance zones of middle- and high-income schools has the potential to reverse this trend, but the link between the housing stock and the socioeconomic segregation of public schools has not been addressed in the existing literature. Using a panel of elementary schools in Florida, we show that increasing the stock of rental and affordable housing units in middle- and high-income neighborhoods has an important effect on the number of poor children attending these schools. Our results also reveal the types of housing units that have the largest impacts on socioeconomic segregation.

Keywords School segregation · Housing affordability · Academic achievement of poor children

✉ Keith Ihlanfeldt
kihlanfeldt@fsu.edu

Tom Mayock
tmayock@uncc.edu

1 Department of Economics and DeVoe Moore Center, Florida State University, Tallahassee, FL, USA

2 Department of Economics, UNC Charlotte and Office of the Comptroller of the Currency, Charlotte, NC, USA

Published online: 26 June 2018

Introduction

One of the most pressing problems confronting the U.S. educational system is the subpar academic performance of children from poor families. The achievement gap between poor and non-poor students is nearly twice as large as that between blacks and whites. Furthermore, the achievement gap between poor and non-poor students has been worsening over time, while the racial gap has been narrowing (Reardon 2011).[1]

As we discuss in more detail below, at least some of this achievement gap can be attributed to the fact that many poor students attend schools where an overwhelming proportion of their peers also come from low-income households. Unfortunately, a large and growing share of students in our public schools are being educated in environments that are socioeconomically segregated. Consider the plot in Fig. 1 that depicts the fraction of all elementary school students attending high-poverty schools in the United States from 1998 through 2013, where a school is alternatively classified as "high-poverty" if at least 50 percent (>50% FLE), 75 percent (>75% FLE), 90 percent (>90% FLE), or 95 percent (>95% FLE) of its pupils qualify to receive a free lunch through the National School Lunch Program.[2,3] Looking first at the left panel, we see that regardless of how we define a high-poverty school, since the early 2000s an increasing share of the nation's elementary school students have been attending such schools. For example, the share of students in schools where the majority of the student body qualified for free lunch rose nationally from 32 percent in 1998 to 48 percent in 2013. Even under our most extreme definition (>95% FLE), the fraction of students in high-poverty schools rose from less than 1 percent to nearly 4 percent over the same time period. Turning to the second panel of Fig. 1, we see that in Florida – the subject of this study – socioeconomic isolation has been increasing at a rate even faster than that of the nation as a whole. In light of the research on the impact of peers' socioeconomic status on a student's academic achievement, these trends do not bode well for the achievement gap.

There is universal agreement that something must be done to improve the academic achievement of poor children, but there is no consensus on how to accomplish this objective. One approach to ameliorating the income-based achievement gap, which has been termed "compensatory education," involves directing additional resources to schools with high concentrations of poverty. A well-known example of this approach is the federal Title One program that provides funding to local school districts to improve the academic achievement of disadvantaged students. Because compensatory education efforts have yielded at best modest returns, policies that aim to narrow the achievement gap through income integration are growing in popularity.

[1] Also illustrative of the poor child's educational deficiency is that the majority of high school dropouts come from families with incomes in the bottom fifth of the income distribution (Carnevale and Strohl 2010).

[2] The eligibility criteria for the National School Lunch Program are described in detail below.

[3] These plots were constructed using data from the Common Core of Data provided by the National Center for Education Statistics. When constructing the plots, a school was classified as an elementary school if it offered third grade. These counts exclude charter schools, magnet schools, and private schools.

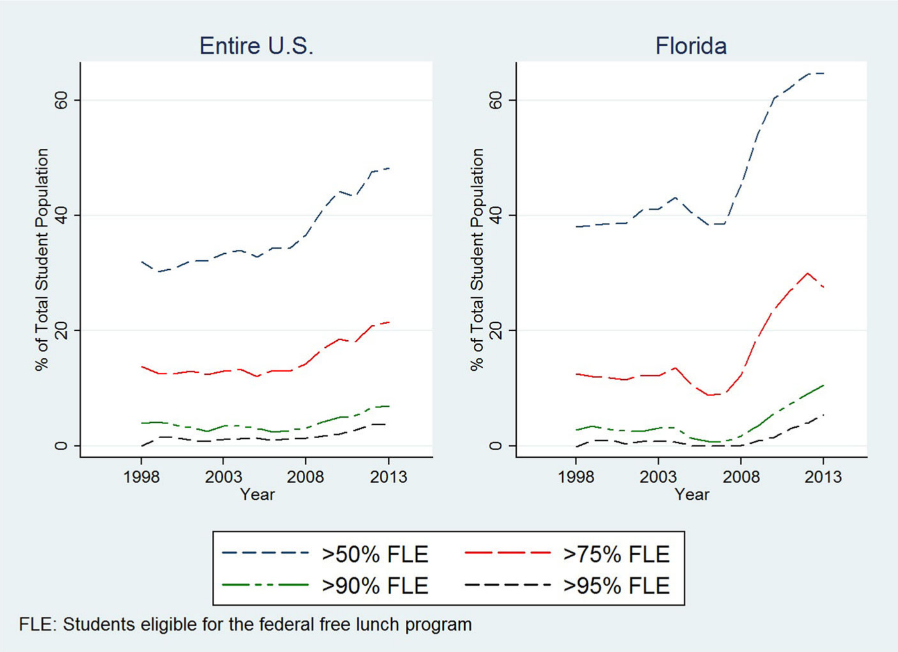


Fig. 1 The Fraction of Elementary School Students in High-Poverty Schools: 1998-2013

A considerable body of research has shown that poor children achieve more academically in income-integrated school environments. Proponents of income integration, however, are divided over how to achieve a desirable mix of poor and non-poor students within schools.

In Florida and generally throughout the United States, public school students are assigned to attend neighborhood schools solely based on where they live. Because poor and non-poor families are residentially segregated, neighborhood school systems oftentimes result in extreme levels of socioeconomic segregation in the public school system.[4] Hence, one option for achieving income integration in public schools is to abandon the neighborhood school model and sever the relationship between a student's residential location and school assignment; such policies are typically referred to as school choice initiatives and generally involve the use of tuition vouchers or the establishment of magnet and charter schools whose enrollments are not dictated by geographic boundaries.[5,6]

There are two ways by which school choice could ameliorate income segregation in schools. First, choice initiatives could be used to encourage students from higher

[4]Fry and Taylor (2012) document the severity and growth of residential segregation by income.

[5]A leading proponent of school choice to achieve income integration is Kahlenberg (2013), who argues for magnet schools in high-poverty neighborhoods and the creation of incentives for non-poor schools to take poor students.

[6]Many economists, starting with Friedman (1955), have long supported the movement away from neighborhood schools towards a choice-based model.

income households to attend schools with a large concentration of poor students. In most cases, this would involve establishing a magnet program that would cater to the educational interests of non-poor students. Critics point out that while a magnet program may help to integrate the high-poverty school, because of educational tracking, the poor children within the school would continue to have peer groups that consisted primarily of other poor children, limiting the potential gains from magnet-induced school integration.[7] Second, school choice could also achieve integration by allowing poor children to attend non-poor schools. A frequent criticism of this approach is that the amount of travel required to get poor children to these schools would be too burdensome for their families.[8]

School choice has been studied extensively by social scientists, but the evidence on whether school choice can be used to reduce segregation has been mixed. The mixed nature of the school choice literature is unsurprising given the significant heterogeneity in school choice policies (Schwartz and Stiefel 2014).

An alternative to school choice and compensatory education policies that may increase income integration in schools is simply creating affordable housing opportunities for the families of poor children within non-poor school attendance zones (SAZs). Many proponents of school income integration favor this approach because it not only improves the poor child's school environment but also her neighborhood environment, which may have its own positive impact on the child's development.[9] However, there are two possible concerns with this approach. First, would poor families be willing to move to a non-poor SAZ if given the opportunity? Other concerns may dominate their locational choice, such as access to public transit, job opportunities, or social services. Low-income families also may not fully realize the benefits that a non-poor school could provide or, if a minority, may be concerned that their child would be uncomfortable in what would likely be a majority white school. Second, even if low-income parents are willing to move to a non-poor SAZ, would poor families be shut out of affordable housing there by racial or class discrimination? Audit evidence continues to show that the goal of equal opportunity remains elusive within local housing markets (Levitt 2014).[10]

[7]Tracking refers to the practice of sorting students into different classes based on their academic ability. Through various assessments, students would be labeled as above average (or gifted), average, or below average. This designation would then determine the student's courses, and teachers and students are grouped only with similarly labeled classmates. Tracking can result from magnet programs because most of the classes offered by these programs coexist in the same public school as regular classes. For a review of the literature see Davis (2014).

[8]Strong criticism of school choice comes from Rothstein (2013), who argues, "too many truly disadvantaged students live too geographically distant from middle-class schools for such schemes to be practical, and too many of their parents are too stressed to make the proactive choices necessary."

[9]For a review of the literature on the relationship between neighborhood quality and school performance visit Harvard's Education Innovation Laboratory at http://edlabs.harvard.edu/neighborhoods-or-schools.

[10]William Fischel's Homevoter Hypothesis (Fischel 1985, 2001), which maintains that homeowners will go to great lengths to protect their property values, adds credence to the idea that poor families may be locked out of non-poor SAZs. Fischel shows that homeowners are aware of the positive effect that a good school can have on their home's value and may believe that low-income students will tarnish the school's reputation.

The purpose of this paper is to test whether reducing the spatial concentration of affordable housing can reduce socioeconomic segregation. Relative to the school choice literature, the empirical work addressing this question is extremely limited. The paucity of studies on the subject is arguably due to data constraints rather than lack of interest in the topic. Identifying the impact of housing affordability on socioeconomic segregation requires data with significant variation in the spatial concentration of affordable housing units, and the slow-moving nature of the housing stock makes such data rare.[11] The volatility in Florida's housing market during the most recent boom-bust cycle in the housing market, however, generated a significant amount of intertemporal variation in the location of affordable housing units within school districts, variation which we use to identify the relationship between the spatial distribution of affordable housing and socioeconomic segregation in public schools. The richness of our data also allows us to identify which types of affordable housing units have the strongest impact on socioeconomic segregation. For example, among affordable units, will more condominiums or mobile homes have the same effect on the number of low-income students attending a non-poor school as more single-family homes? Answering this question could mean the difference between a successful and unsuccessful housing subsidy program.

Our study is based on a unique panel data set at the level of the individual elementary school from 38 Florida school districts that runs from 1998 through 2013. For each school in our panel, we construct counts of the number of students within the school that qualify to receive free lunches through the National School Lunch Program; these free lunch counts serve as our primary measure of the number of economically disadvantaged students that attend a particular school. These student counts are then linked with very detailed data on the housing stock within a school's school attendance zone (SAZ). In addition to structure type (e.g., a single-family unit), our housing stock data includes information on whether a unit is owner- or renter-occupied and whether or not the unit would be affordable to a moderate income household based on federal guidelines.

Due to the unprecedented boom and bust housing cycle experienced in Florida over the course of our panel, a notable feature of our data is the intertemporal variation in the number of affordable housing units found within individual SAZs. In essence, the tremendous shocks to Florida's housing market during the early-2000s housing boom and the bust later in the decade provide a natural experiment to address how changes in the housing stock affect the spatial distribution of students in a neighborhood school system.

We use three alternative estimators – a linear fixed effects model (FE), a Pooled Fractional Probit model (PFP), and a linear instrumental variables model (IV) – to address various econometric challenges associated with identifying the impact of changes in the housing stock on household location decisions. Our results, which are robust across these different estimators, suggest that increasing the fraction of a

[11] Ihlanfeldt (Forthcoming) provides some evidence on the extent to which improving housing affordability in neighborhoods zoned for high-performing schools might reduce racial segregation.

school district's stock of affordable rental units in high-income SAZs reduces socioeconomic segregation. The estimated effects are nontrivial in magnitude, especially for affordable multifamily rental units, where a ten percent increase in the number of apartments within a high-income SAZ is estimated to decrease the number of free lunch students attending low-income schools by as much as three percent.

The remainder of our paper is structured as follows. In "Literature Review", we review the relevant literature, which is sparse since we could not find a single study that directly addressed our research questions. Sections "Data" and "Empirical Methodology" describe our panel data and empirical methodology, respectively. Results are presented in "Results", and in "Conclusion" we conclude.

Literature Review

The Century Foundation has long promoted income integration within schools as an approach toward improving the academic achievement of poor children. Most prominent in this advocacy is Richard Kahlenberg who has written a collection of books and articles on the topic. In his essay "From All Walks of Life" (Kahlenberg 2013), he reviews the literature relating socioeconomic integration to student performance, which he concludes shows an overwhelmingly strong positive relationship between academic achievement and income integration.[12]

Kahlenberg enumerates three reasons for the superior performance of poor students in middle-class schools: peers who, on average, are more academically engaged and less likely to act out than those in high-poverty schools; a community of parents who are able to be more actively involved in school affairs and know how to hold school officials accountable, and stronger teachers who have higher expectations for students. Though these factors are plausible explanations for why poor children exhibit better academic performance in middle-income schools, as Rothwell (2012) notes, there is no empirical evidence on their validity. Thus, while there is little doubt that income integration has a positive impact on a poor child's learning, the current state of the literature is such that we do not know exactly why this is true.

The evidence provided by the Century Foundation and others in favor of income integration is compelling and serves to motivate our analysis. The question that is at the heart of our inquiry is whether income integration in public schools can be achieved by increasing housing opportunities for low-income households in non-poor SAZs. As noted above, there is no prior evidence of direct relevance to this question. Of interest, however, are three recent studies that focus on the contribution that restrictive land use controls, which have been shown to lower housing affordability, play in generating communities that are segregated by income. Rothwell (2012) shows that anti-density zoning makes housing unaffordable to poor families within

[12] Kahlenberg also cites a review of 59 studies by Mickelson and Bottia (2009) that reached the same conclusion.

SAZs containing good schools. The other two studies, by Rothwell and Massey (2010) and Lens and Monkkonen (2016), show that anti-density zoning and residential income segregation are closely related. Despite the growth in school choice programs, the vast majority of elementary school students are still assigned to and attend their neighborhood school. Hence, if housing affordability contributes toward residential segregation, this suggests that it also helps to explain economic segregation within school districts. Each of these three studies is reviewed in turn. Rothwell (2012) estimates the difference in housing costs between the SAZs of low- and high-performing schools for individual metropolitan areas.[13] He finds that homes zoned for high-performing schools cost 2.4 times as much as housing in low-performing SAZs and that this difference can be attributed to exclusionary land use regulations such as density controls. These results lead Rothwell (2012) to conclude that the elimination of exclusionary zoning would substantially lower the cost of housing in high-performing SAZs, allowing poor children to attend better schools. He acknowledges, however, that homeowners' incentives for retaining anti-density zoning are strong and therefore the elimination of exclusionary zoning is unlikely to occur anytime in the near future. Rothwell then goes on to offer a number of second best alternatives, giving particular attention to expanding the portability of housing vouchers. Rothwell and Massey (2010) and Lens and Monkkonen (2016) are highly similar, with the authors of the latter study describing their work as an extension of Rothwell and Massey (2010). Both of these papers use cross-sectional data for metropolitan areas to study the relationship between measures of residential income segregation and land use regulations. Rothwell and Massey (2010) relies upon (Pendall and et al. 2006) survey of local land use regulations, while Lens and Monkkonen (2016) base their analysis on the Wharton Land Use Regulatory Index, along with its eleven sub-indices that describe different dimensions of the regulatory environment (Gyourko et al. 2008). Both studies find a strong relationship between land use regulations that limit density and residential income segregation. Despite their similarity, there are two primary differences between these studies. First, while Rothwell and Massey (2010) address the possible endogeneity of anti-density zoning by estimating two-stage least squares models, Lens and Monkkonen (2016) make no such effort. The second difference is that while Rothwell and Massey (2010) study the relationship between land use regulations and segregation using standard segregation measures, the analysis in Lens and Monkkonen (2016) utilizes measures that register the separate segregation of poor, middle income, and affluent households. The analysis of these alternative segregation measures led Lens and Monkkonen (2016) to conclude that regulations that limit density are associated with the segregation of the wealthy and middle income, but not the poor.

[13] In this study, school performance was measured using test score data from Great Schools, and the housing cost measures were a weighted average of rental and owner monthly expenditures derived from the American Community Survey.

Finally, the literature on the Moving to Opportunity (MTO) project and the housing voucher program suggests that low-income parents may not be willing to relocate to neighborhoods with lower concentrations of poverty even when such moves are financially feasible. MTO was a federally funded demonstration project conducted in five large metropolitan areas between 1994 and 1998 (Sanbonmatsu and et al. 2011). The purpose of the project was to determine the impact that improved neighborhood quality has on the welfare of low-income households. Random assignment was used to place eligible households into one of three groups — the experimental group, the Section 8 group, and the control group. The experimental group was offered a housing voucher that could only be used in a low-poverty neighborhood, the Section 8 group was offered a voucher that could be used anywhere, and members of the control group were not offered vouchers but received normal housing public assistance. Surprisingly, only 48 percent of families in the experimental group used the MTO voucher to move to a lower-poverty neighborhood, and only 63 percent of the Section 8 group moved using the MTO voucher. In the context of our study, the MTO results suggest that low-income families with children attending schools with a high concentration of economically disadvantaged students may fail to move when housing becomes more affordable in better neighborhoods. Additional evidence suggesting that low-income families may not take advantage of housing options in neighborhoods zoned for better schools is provided by Horn et al. (2014). Using confidential data on the residential locations of voucher recipients from the Department of Housing and Urban Development (HUD), the authors find that the children of voucher holders were more likely to attend low-performing schools than the children of households that did not receive housing vouchers. In a follow-up study that investigates why this is the case, Ellen et al. (2016) emphasize the role of distance between the existing neighborhoods of voucher recipients and high-performing schools. Because of the spatial clustering of low-performing schools, for many low-income households relocating to a neighborhood zoned for better schools involves a long-distance move. If information on school quality declines with distance, then many low-income households may have very limited knowledge of better schools, reducing the likelihood of relocation.

In summary, the previous studies provide strong evidence that the educational achievement of poor children can be improved by having them attend schools that are integrated by socioeconomic status, and several papers suggest that land use regulations that limit density contribute to economic segregation. The existing literature thus suggests that counteracting exclusionary zoning could possibly reduce the cost of housing in school boundaries that were previously unaffordable to many low-income families. Given previous research on the location decisions of housing voucher recipients, however, it is unclear whether low-income families will relocate to neighborhoods zoned for better schools when such opportunities arise. The existing literature is thus silent on the extent to which market-driven increases in the supply of affordable housing in high-income neighborhoods can reduce economic segregation in neighborhood school systems. The primary contribution of this paper is providing the first evidence on this question.

Data

Data Construction

The initial step in the construction of our database involved assigning residential properties to SAZs. To do this, we first geocoded the addresses of the residential properties in Florida reported in the DataQuick property tax assessment database. Next, we obtained digital 2013-vintage SAZ maps from the National Center for Education Statistics (NCES) for 58 of Florida's 67 school districts and 2011-vintage SAZ maps for three additional districts from the vendor Maponics.[14,15] The geographic coordinates for each of the addresses in the DataQuick file were then used to identify the school that a student at that address is assigned to attend in every grade between kindergarten and 12^{th} grade. Because the grade coverages of schools are not standardized across or even within school districts, we had to construct an operational definition of an "elementary school." To that end, we defined the elementary school serving a housing unit as the school that a child at that location would be assigned to attend in third grade. Schools that were not classified as elementary schools using this definition were then removed from the data.

To characterize how the housing stock within the elementary school SAZs evolved over time, we linked the residential properties to the standardized tax rolls that each county must submit to the Florida Department of Revenue (FDOR).[16] These tax roll data, which are updated on an annual basis, contain a wealth of information on real property characteristics, including information on the type of property (e.g., single-family, condominium), the year in which the property was constructed, and an estimate of the property's market value. The estimates of market value are made by the property tax assessor in each county. The FDOR requires that these estimates reflect what the property would sell for in an arm's length sale on January 1 of the tax roll year. The assessors construct these estimates using the three standard approaches for estimating the market value of real estate: the replacement cost approach, the

[14] The districts excluded from the NCES and Maponics digital boundary files are mostly small, rural counties.

[15] An admitted limitation of our data is that while school boundaries may change over time, our assignment to schools is based on the boundaries as they existed in 2013 in the case of the NCES files and 2011 in the case of the Maponics data. Because historical SAZ boundaries are not available, we cannot speak to how frequently SAZ boundaries have changed over the course of our sample. The geography that we ultimately use in our empirical work, however, is a grouping of schools based on socioeconomic characteristics and not the individual SAZs. That said, even if the school to which a home was assigned changed over the course of our study, as long as the post-SAZ-change school to which the home was assigned remained in the same socioeconomic grouping, the change in SAZ boundaries would have no impact on our housing stock measures. While we can only speculate, because households are spatially segregated by income and school boundary changes tend to be relatively minor amendments to existing boundaries, we believe that our housing stock measures are likely close to those that would be constructed using SAZs that are updated each year.

[16] Tax roll data for recent years are available online at: http://dor.myflorida.com/dor/property/resources/data.html. FDOR collects these rolls to monitor the performance of the county tax assessors.

comparable sales approach, and the income approach. FDOR requires that two of the three approaches be used to estimate of market value for each property. The assessor's estimates of market value are evaluated over the next two months by FDOR's Property Tax Oversight Division. Sales samples are used to compare the tax assessor's estimates of market values to actual sale prices. In thin markets sales samples are supplemented with appraisals conducted by FDOR staff. Based on these comparisons, FDOR may reject the roll requiring the tax assessor to provide more reliable estimates of market values. The end product of this assessment process is a set of market values that track sales prices quite closely, even during periods of extreme market turmoil.[17,18]

Importantly for our study, the tax roll data also contains fields that indicate whether or not a property was granted a property tax homestead exemption. This exemption is available to "a person who, on January 1, has the legal title or beneficial title in equity to real property in [Florida] and who in good faith makes the property his or her permanent residence or the permanent residence of another or others legally or naturally dependent upon him or her" (Hintermaier 2016). We use the presence of a homestead exemption to classify a property as owner-occupied, and housing units without a homestead exemption are classified as renter-occupied. Because the exemption provides significant tax savings, owner-occupants have strong financial incentives to file for the exemption, and we are thus confident that owner-occupied units will generally be correctly classified based on homestead status. Properties that are not covered by a homestead exemption are primarily either rental units or second homes. The fraction of single-family homes that are second homes is expected to be small because in Florida most vacation homes are condominiums. For condominiums we cannot rule out the possibility that a substantial number of the properties we label as rentals may in fact be second homes not available for rent.

After classifying the housing units as owner-occupied and renter-occupied, we then use the property's estimated market value from FDOR to classify the housing units as affordable and unaffordable. Specifically, we first calculate rent-to-price ratios using the Public Use Microdata Sample (PUMS) for the American Community Survey (ACS) by regressing the rent reported for single-family rentals on variables describing the size and quality of the home. Separate regressions were estimated for each of the 11 market areas in Florida, and the parameters from these regressions were used to construct an imputed rent measure for the owner-occupied single-family homes in the PUMS data. The imputed rent measures were then divided by the owner's estimate of market value to obtain a unit-specific rent-to-price ratio. These ratios were averaged over all of the sampled properties in a market area to obtain the area's rent-to-price ratio. With the rent-to-price ratios in hand, we construct imputed

[17] See Appendix B of Andersson and Mayock (2014) for an analysis of the accuracy of the FDOR values during the 2008 – 2011 period.

[18] In evaluating the county tax rolls, the FDOR computes price related differentials (PRD) to assess the vertical equity of the property tax within each county. The PRD is a statistic for measuring assessment regressivity or progressivity. If the PRD shows vertical inequity in the administration of the property tax, the tax roll may be rejected by the FDOR until the inequity is addressed by the county.

rent measures for every housing unit in the FDOR data between 1998 and 2013 by multiplying the rent-to-price ratio by the property's estimated market value.

A property is designated as affordable in a given year if this imputed rent is equal to or lower than the fair market rent (FMR) for two-bedroom housing units reported by the Department of Housing and Urban Development (HUD), and housing units with imputed rents above the two-bedroom FMR are classified as unaffordable. For each SAZ in our data, we construct a count of each of the following housing types by affordability status: single-family homes, condominiums, mobile homes, multifamily units, and cooperatives.[19]

Lastly, we obtained information on public housing and subsidized rental housing units from the University of Florida's Shimberg Center; we refer to such housing units as "assisted housing units" hereafter.[20] For each residential property containing assisted housing units funded by HUD, the U.S. Department of Agriculture, the Florida Housing Finance Authority, and local housing finance authorities, the Shimberg data reports the year that the structure was built, the address of the structure, the number of units that are subject to rent and/or income restrictions, the primary population served by the housing development (e.g., the elderly, families), and the date on which the housing units enter into and exit from assisted housing status. The addresses in the Shimberg data were used to assign all assisted housing units to elementary schools using the geocoding procedure described above. After this assignment, we utilized the fields reporting the number of assisted housing units in a property and the dates on which such units entered and exited assisted status to generate annual SAZ-level counts of the assisted housing stock. Because there is little reason to believe public housing programs targeted towards housing the elderly and disabled will affect school populations, our assisted housing counts only included those units that were designated for families.

Because we are primarily interested in how the changes in the housing stock affect socioeconomic segregation in traditional neighborhood school systems, we remove from the data school districts that have significant open enrollment policies that break the relationship between a student's residential location and where he or she attends school. Lastly, because intra-district school segregation is only meaningfully defined when a district has multiple schools, we drop the small rural school districts that only contain one elementary school. After imposing these restrictions, we were left with 38 school districts containing 1481 elementary schools.

[19] We exclude retirement homes and institutional housing, such as school dormitories and correctional facilities, from our analysis. We also exclude bank-owned properties, because such properties are not legally inhabitable.

[20] The data can be accessed at http://flhousingdata.shimberg.ufl.edu/a/ahi_basic. We include assisted units as a control variable and as a test of the validity of our models. An increase in the number of assisted housing units located in non-poor SAZs should increase the number of free lunch students in such schools because, by virtue of the fact that they qualified for housing assistance, the families occupying these likely also qualify for the free lunch program. If assisted units do not have this expected effect, it would suggest that our models may be misspecified.

Our analysis requires information on the socioeconomic status of children attending public schools. Because our data does not directly report measures of students' household incomes, we must resort to the use of proxies for socioeconomic status. The proxy most frequently used in previous studies to measure economic disadvantage is based on eligibility data for the Free Lunch Program under the School Lunch Act. Eligibility for free or reduced-price meals is determined by household size and income or through categorical eligibility. The U. S. Secretary of Agriculture sets the income eligibility levels annually. Children in households with incomes at or below 130 percent of the federal poverty guidelines are eligible for free meals. To obtain the number of free-lunch-eligible students ("free lunch students"), we linked our panel to the NCES Common Core of Data. The Common Core files contain a number of variables that characterize the student population of all reporting schools. Most importantly for this study, included among these variables is a count of the number of students in a given school that are eligible to participate in the Free Lunch Program.

The second proxy that we use for socioeconomic disadvantage is based on data from the Florida Department of Education (FDOE) that reports the percentage of students at the school who pass a mathematics proficiency exam administered as a part of the Florida Comprehensive Assessment Test (FCAT). It has been well documented that there is a high correlation between a school's performance on standardized exams and the income level of the students' parents (Dahl and Lochner 2012). If the relationship between average school household income and pass rates is monotonic, then the realized performance on the standardized tests can be used to rank order schools by household income.

In our empirical analysis, we use each of the socioeconomic status proxies to construct neighborhood typologies. The construction of the typology based on the free lunch data proceeded in three steps. First, for each school-year observation in our data we calculated the fraction of students that qualified for the free lunch program. Let PFL_{kit} denote the fraction of students qualifying for free lunch in school k in school district i in year t. Then, for each school district in the data, we calculated the median value of PFL_{kit} for all schools in the district. Let PLF_{it}^{Med} denote this median. If $PFL_{kit} > PLF_{it}^{Med}$, then school k is classified as a low-income school in year t in district i; otherwise school k is classified as a high-income school.[21]

In the second step of our classification process, we defined a school as globally low-income (high-income) if for all years it was classified as low-income (high-income). The remaining schools are those whose percentages of free lunch students tended to be at or close to the district median percentage each year throughout the panel; in some years these schools are classified as low-income and in other years as high-income. We classified these schools as globally middle-income schools. Finally, using schools' global classifications we aggregated them into low-, middle-, and high-income groups, with roughly a third of the schools falling into each of the three groups.

[21] To avoid simultaneity problems, we needed a school income typology in which an individual school remains in the same income group throughout the panel. We found that there was considerable persistence in the annual categorizations. Schools categorized as high (low) income in one year tended to be high (low) income in the other years.

Our school typology based on the performance on the standardized mathematics test was also constructed using a multi-step process. First, we averaged pass rates over the years of our panel at the school and district levels. If the school's intertemporal average pass rate was within 5 percentage points of the district's intertemporal average, it was categorized as a middle-income school. If a school's intertemporal average pass rate was more than 5 points above (below) the district average, it was categorized as a high- (low-) income school. These cutoffs were chosen to produce a roughly equal number of low-, middle-, and high-income schools.

Summary Statistics

Table 1 shows the average proportion of a school's total number of students that qualify for the free lunch program broken down by the three income groups for selected years of our panel — the first year (1998), the last year (2013) and two intermediate years (2005 and 2010). 2005 was near the peak of the housing boom for most

Table 1 Average fraction of a school's students receiving a free lunch by school income group and year

	Year			
Income group	1998	2005	2010	2013
Global free lunch[1]				
Low	0.776	0.792	0.869	0.860
Middle	0.558	0.598	0.713	0.702
High	0.307	0.315	0.424	0.461
Global math pass[2]				
Low	0.748	0.776	0.86	0.839
Middle	0.548	0.573	0.694	0.707
High	0.320	0.323	0.425	0.453
Students eligible for free lunch program[3]	492422	533675	611936	630374
Total student population	909515	983936	966856	978809
School districts in sample	38	38	38	38

[1] High and low income schools are defined as those where the percentage of students qualifying for the free lunch program is greater than the districtwide median percentage of students qualifying for the free lunch program in every year of the panel. The middle income group is comprised of schools where the fraction of students eligible for the free lunch pgoram varied around the district median over the course of the panel

[2] Schools with an average pass rate on the mathematics exam that were more than five percentage points above (below) the districtwide average were categorized as high income (low income). The remaining schools were classified as middle income

[3] Summary statistics are based on the sample that was used to estimate the regression models, not population of all Florida public school students

Table 2 Distribution of pass rates on florida standardized mathematics test over time

	Year 2001	2002	2003	2004	2005	2006	2007	2008	2009	2010
Fraction of schools[1]										
with pass rate										
<20%	1.25%	0.42%	0.00%	0.21%	0.00%	0.07%	0.00%	0.00%	0.00%	0.00%
20%-40%	11.05%	8.96%	5.42%	3.27%	2.64%	1.88%	0.63%	0.14%	0.14%	0.42%
40%-60%	31.20%	28.70%	26.34%	22.45%	19.60%	16.40%	13.34%	10.63%	10.91%	11.26%
60%-80%	33.70%	37.87%	43.36%	44.96%	44.27%	45.73%	46.70%	45.45%	47.32%	47.05%
>80%	22.80%	24.05%	24.88%	29.11%	33.49%	35.92%	39.33%	43.78%	41.63%	41.27%
School districts in sample	38	38	38	38	38	38	38	38	38	38

[1] Pass rates are constructed using data provided by the Florida Department of Education that reports the percentage of students at a school who pass a mathematics proficiency exam administered as a part of the Florida Comprehensive Assessment Test (FCAT)

local markets, and by 2010 all markets in our sample had experienced their crash in housing values.[22]

Table 2 reports the fraction of our schools in our sample that had a pass rate on the math proficiency exam of less than 20 percent, between 20 percent and 40 percent, between 40 percent and 60 percent, between 60 percent and 80 percent, and more than 80 percent in each year from 2001 through 2010.[23] Virtually no schools had pass rates below 20 percent at any point in time. It thus appears that in even the lowest-performing schools in the state at least 20 percent of students are able to pass the mathematics portion of the FCAT exam. In the early years of FCAT testing, a non-trivial number of schools had pass rates between 20 and 40 percent; by the end of the panel, however, the vast majority of schools in our data had pass rates that exceeded 40 percent. Many of the schools with low pass rates in the early years in our sample increased student performance significantly over time. In 2001 roughly 57 percent of schools had a pass rate of 60 percent or higher, while by 2010, more than 88 percent of schools had pass rates of 60 percent or higher.

Unsurprisingly, for both of our socioeconomic status proxies, the fraction of students that qualify for the free lunch program in schools classified as low-income is roughly twice as high as the fraction of students qualifying for the program in schools classified as high-income. Within all three income groups, the fraction of socioeconomically disadvantaged students generally trended upward over the course of the panel, even within the high-income group.

One possible explanation for the increase in low-income students attending the high-income schools is that the collapse in housing values during the Great Recession may have resulted in low-income households sorting into high-income neighborhoods. The figures reported in Tables 3 and 4 – which report the fraction of a district's affordable housing located in low-, middle-, and high-income SAZs – speak to this possibility.

Between 1998 and 2005, the fraction of districts' affordable housing units zoned for high-income schools declined for many housing unit types likely to house school-age children, such as single-family housing units. These dynamics are consistent with housing values appreciating more rapidly during the run-up in high-income SAZs. The Great Recession, however, reversed this trend of declining affordability in high-income SAZs: between 2005 and 2013, the fraction of districts' affordable units zoned for high-income schools increased significantly for many housing types including single-family and condominium units.

It is worth noting that some of the increase in housing the affordability of high-income SAZs may have been driven by the foreclosure crisis. While the foreclosure rate was generally higher in low-income SAZs, the level of foreclosure activity was

[22] Changes in our housing stock measures over time suggest that affordability declined in the high-income SAZs during the rise in housing prices and rose after the crash in housing prices. The differences in these housing stock measures over time, however, are not generally statistically significant, a finding that is largely attributable to very large variances about the means.

[23] The FDOE data begins consistently reporting the pass rates on the mathematics proficiency portion of the Florida Comprehensive Assessment Test (FCAT) in the 2001-2002 school year. In 2011, the FCAT was replaced with a different assessment instrument (FCAT 2.0). Because of this change, we only use FCAT data from 2001 to 2010 in our analysis.

Table 3 Mean proportion of each affordable housing type falling within low, middle, and high income school attendance zones using global free lunch as income proxy

Housing Types	Income[1] Group	Year 1998	2005	2010	2013
Single-family rental[2]	Low	0.337	0.336	0.337	0.348
	Middle	0.410	0.411	0.400	0.397
	High	0.253	0.223	0.262	0.255
Single-family owner	Low	0.294	0.346	0.294	0.298
	Middle	0.428	0.424	0.408	0.408
	High	0.278	0.23	0.298	0.293
Mobile home rental	Low	0.280	0.278	0.263	0.257
	Middle	0.471	0.465	0.495	0.438
	High	0.248	0.256	0.292	0.305
Mobile home owner	Low	0.247	0.254	0.237	0.237
	Middle	0.476	0.464	0.455	0.448
	High	0.277	0.281	0.308	0.315
Condominium rental	Low	0.281	0.291	0.286	0.283
	Middle	0.379	0.359	0.359	0.340
	High	0.314	0.323	0.329	0.351
Condominium owner	Low	0.263	0.287	0.257	0.253
	Middle	0.378	0.35	0.344	0.333
	High	0.333	0.31	0.373	0.387
Multifamily	Low	0.321	0.333	0.343	0.344
	Middle	0.381	0.388	0.378	0.381
	High	0.298	0.279	0.278	0.274
Assisted housing	Low	0.447	0.402	0.414	0.414
	Middle	0.347	0.371	0.340	0.338
	High	0.206	0.226	0.245	0.248
School districts in sample		38	38	38	38

[1] High and low income schools are defined as those where the percentage of students qualifying for the free lunch program is greater than the districtwide median percentage of students qualifying for the free lunch program in every year of the panel. The middle income group is comprised of schools where the fraction of students eligible for the free lunch pgoram varied around the district median over the course of the panel

[2] "Owner" and "Renter" denote housing units that are owner- and renter-occupied, respectively

generally higher in high-income SAZs because these SAZs included a large share of a district's owner-occupied units (Ihlanfeldt and Mayock 2016b). Foreclosures increase affordability in two ways. First, extensive research has shown that foreclosures produce negative spillover effects that depress neighborhood housing values.[24] Second,

[24] Ihlanfeldt and Mayock (2016a) review 12 studies, all of which provide evidence that a foreclosed upon property reduces the values of nearby homes. Although the focus of much of the literature has been on single-family homes, some studies have shown that foreclosures also lower the cost of condominiums (Campbell et al. 2011) and multifamily housing (Schuetz et al. 2008); hence, the evidence suggests that foreclosures bring about a general increase in housing affordability within a SAZ.

Table 4 Mean proportion of each affordable housing type falling within low, middle, and high income school attendance zones using global math exam pass rate as income proxy

Housing types	Income[1] Group	Year 1998	2005	2010	2013
Single-Family eental[2]	Low	0.370	0.409	0.374	0.382
	Middle	0.417	0.411	0.407	0.409
	High	0.213	0.180	0.219	0.209
Single-family owner	Low	0.334	0.387	0.326	0.330
	Middle	0.436	0.425	0.426	0.432
	High	0.230	0.187	0.248	0.238
Mobile home rental	Low	0.324	0.322	0.297	0.294
	Middle	0.470	0.468	0.476	0.474
	High	0.206	0.210	0.227	0.231
Mobile home owner	Low	0.288	0.293	0.265	0.252
	Middle	0.490	0.477	0.498	0.500
	High	0.222	0.230	0.236	0.247
Condominium rental	Low	0.282	0.312	0.297	0.296
	Middle	0.409	0.394	0.399	0.394
	High	0.283	0.268	0.276	0.284
Condominium owner	Low	0.295	0.311	0.283	0.282
	Middle	0.366	0.373	0.351	0.353
	High	0.312	0.264	0.339	0.339
Multifamily	Low	0.361	0.381	0.387	0.388
	Middle	0.400	0.403	0.400	0.403
	High	0.238	0.215	0.213	0.209
Assisted housing	Low	0.506	0.470	0.459	0.458
	Middle	0.352	0.371	0.380	0.381
	High	0.142	0.159	0.160	0.160
School Districts in Sample		38	38	38	38

Schools with an average pass rate on the mathematics exam that were more than five percentage points above (below) the districtwide average were categorized as high income (low income). The remaining schools were classified as middle income

"Owner" and "Renter" denote housing units that are owner- and renter-occupied, respectively

foreclosures most frequently result in a home being repossessed by the lender; during the most recent crisis, it became increasingly common for these bank-owned properties to be purchased by investors who then offered the homes for rent (Ihlanfeldt and Mayock 2016a). These rentals, many of which were single-family units, offered low-income families who could not afford to buy a single-family home the opportunity to rent in neighborhoods that were previously unaffordable. Interestingly, the fraction of affordable mobile homes zoned for high-income schools increased monotonically over the course of our panel. While the stock of affordable mobile homes grew in all three income groups, the rate of growth was faster within the high-income SAZs, resulting in a larger fraction of such housing types being zoned for high-income schools.

To summarize, our data show that public schools in Florida are characterized by high levels of socioeconomic segregation. The spatial distribution of affordable housing within school districts, however, has changed significantly over time, with affordability generally increasing in high-income SAZs during the Great Recession. It is this intertemporal variation in the intra-district distribution of the affordable housing stock that we utilize to identify our empirical models.

Empirical Methodology

In all of our empirical models, the dependent variable is the fraction of a school district's socioeconomically disadvantaged students that attend low-income schools based on the income typologies defined above. Let P_{it} denote this proportion in school district i in year t. Our models express P_{it} as a function of the proportion of a district's affordable housing that is located in the middle- and high-income SAZs in the district. If low-income families relocate to higher income neighborhoods when affordable housing opportunities in those neighborhoods become available, then as the fraction of affordable units located in the middle- and high-income SAZs increases, we expect the fraction of low-income students concentrated in low-income schools to decline. We test this hypothesis using three different empirical models: a standard linear fixed effects model (FE), a linear instrumental variables model (IV), and the Pooled Fractional Probit (PFP) model of Papke and Wooldridge (2008).

P_{it} is bounded between zero and one, a constraint that is imposed by the nature of the PFP model. Furthermore, the PFP model also allows for us to control for unobserved district-level heterogeneity and is straightforward to estimate using quasi-maximum-likelihood (QMLE) methods. The PFP model is thus a natural starting point for our empirical analysis.

As we discuss in more detail below, for the PFP model to be consistent, all of the covariates must be strictly exogenous. Relaxing the strict exogeneity assumption in non-linear models such as the PFP can be econometrically challenging. That said, to check the robustness of our results we also estimate two linear models: a standard fixed effects panel model and an instrumental variables model. While those two models do not bound dependent variable to the unit interval, the coefficients from the FE and IV model can be compared with the average marginal effects from the PFP model to get a sense of how relaxing the strict exogeneity assumption affects our conclusions.

Before expressing these models formally, we must first introduce some notation. Let i index school districts and t index years. H_{it}^{Middle} are H_{it}^{High} are vectors containing the fraction of different types of housing units that are located in the district's middle-income and high-income zones, respectively, in year t. Additionally, let $\overline{H}_{it}^{Middle}$ and $\overline{H}_{it}^{High}$ denote vectors of the intertemporal means of these housing variables in the middle- and high-income zones. Lastly, define θ_t as a year fixed effect and Z_{it} as the full vector of covariates in the model.

The point of departure for the estimation of the PFP model is an assumption regarding the functional form of the conditional mean. Specifically, for the PFP model we assume that

$$E\left[P_{it}|Z_{it}\right] = \Phi\left(\theta_t + H_{it}^{Middle}\beta^{Middle} + H_{it}^{High}\beta^{High} + \overline{H}_{it}^{Middle}\gamma^{Middle} + \overline{H}_{it}^{High}\gamma^{High}\right) \quad (1)$$

where β^{Middle}, β^{High}, γ^{Middle}, and γ^{High} are all parameter vectors and $\Phi\,()$ is the standard normal cumulative distribution function. We estimate Eq. (1) via Bernoulli QMLE. As noted in Papke and Wooldridge (2008), the inclusion of the intertemporal means in Eq. (1) controls for time-invariant district-level unobserved heterogeneity in a manner similar to the inclusion of fixed effects in a linear panel model.

Using the same basic notation as in Eq. (1), our linear fixed effects model (FE) can be expressed as

$$P_{it} = \lambda_i + \theta_t + H_{it}^{Middle}\beta^{Middle} + H_{it}^{High}\beta^{High} + \varepsilon_{it} \quad (2)$$

where λ_i is a school district fixed effect.

The functional form of our instrumental variables (IV) model is identical to that of Eq. 2 but instead of imposing the strict exogeneity assumption, we instrument for variables we believe may potentially be endogenous.

Identification and Robustness Checks

For the PFP and FE models to be consistent, our housing stock measures must be strictly exogenous. Strict exogeneity requires both contemporaneous exogeneity and the absence of feedback effects. We address each of these requirements in turn.

Contemporaneous exogeneity may not hold if there are unobserved factors that influence the fraction of low-income students attending low-income schools and such factors are also correlated with the intra-district distribution of the housing stock across low-, middle-, and high-income SAZs.[25] A potentially important set of unobservables in our case are economic shocks that result in a change in a student's eligibility for the free lunch program. For example, it may be the case that during the Great Recession family incomes in the high-income SAZs fell because of layoffs. If these layoffs resulted in negative income shocks that resulted in more existing students in high-income SAZs becoming eligible for a free lunch, then the unobservable income shocks would raise the proportion of free lunch students in the higher income schools and decrease the proportion of free lunch students attending the low income schools. If these same layoffs also caused a reduction in housing values in the high-income SAZs, then contemporaneous exogeneity would be violated.[26,27]

[25] Unobservables that are time invariant pose no threat to identification because all of our models control for district-level time invariant heterogeneity.

[26] Layoffs in a neighborhood could cause a reduction in local housing values if, in response to the negative income shocks, many households attempted to sell their properties at the same time, increasing inventory and depressing prices.

[27] Even if incomes fell after the crash, within the higher income schools there may have been little if any free lunch switching as there may be relatively few families whose income would have fallen enough to become eligible for the free lunch program. Also, even if formerly high-income students became eligible for the program, their parents may have chosen not to apply to avoid their child possibly being stigmatized.

Even if our covariates are contemporaneously exogenous, strict exogeneity may still fail to hold if values of the covariates in period t affect future values of the idiosyncratic shock or, alternatively, if the idiosyncratic shock in period t affects future values of the covariates. For example, information on the state of public schools is likely received by potential homebuyers with a significant lag. If homeowners are averse to sending their children to schools with a high concentration of low-income students, then a shock that increases the low-income student population in a particular SAZ in period t may reduce prices – and increase affordability – in some time period after t. This type of mechanism would induce a correlation between the error term in period t and the values of the covariates in subsequent periods, violating the strict exogeneity assumption. Papke and Wooldridge (2008) propose a method for addressing endogeneity in the PFP model using control function methods. They found that the average partial effect from the PFP model using the control function methods was very similar to the estimated coefficient on the endogenous variable from a linear instrumental variables estimator. This finding led them to conclude that "in many non-linear contexts, the linear model does a good job of estimating the average partial effect" (Papke and Wooldridge 2008, p. 130). In light of this finding and the relative complexity associated with the control function approach with many endogenous variables, we estimate linear instrumental variables models as a robustness check on our baseline FE and PFP models.

We construct instruments for each of our housing stock measures by interacting the values of the endogenous variables in 1995 – 3 years prior to the start of our panel – with a set of dummy variables for each year of our panel (1998-2013). The key assumption underlying our identification strategy is that after controlling for the distribution of affordable housing across SAZs, housing stock measures from the base year have no direct effect on the distribution of socioeconomically disadvantaged students across low-, middle, and high-income schools in the current year.[28]

It is important to draw a distinction between how we have constructed our instruments and instruments based on lagged endogenous variables the values of which change throughout the panel. As is well known, the use of a lagged endogenous variable as an instrument can be problematic if unobservables are serially correlated (Angrist and Krueger 2001). Because our models explicitly control for time-invariant unobservables through the use of fixed effects, time-varying unobservables are our only threat to identification. These time-varying unobservables will only compromise our identification strategy if the value of the endogenous variable in the base year (1995) does not satisfy the exclusion restriction and has a direct impact on the dependent variable. Although the exclusion restriction cannot be tested directly, tests based on Akaike's information criterion suggested that models without lagged covariates were superior to those that included lags; we interpret this finding as indirect evidence that the exclusion restriction holds.

[28]First-stage diagnostics for the IV models are strong in all cases. For both income proxies, Shea's partial R-squared, which measures the strength of the correlation between the endogenous variable and its instrument, is in all cases respectable, with the lowest value equaling 0.31. For all of the endogenous variables, the first-stage F-statistics are all statistically significant at the 1% level and are, on average, greater than 10.

We implement strict exogeneity tests for all of our empirical models. Wooldridge (2010) proposes a test that involves adding the leading values of suspected endogenous explanatory variables to the estimating equation. The null hypothesis of strict exogeneity is rejected if the leading variables are jointly significant, suggesting the presence of feedback effects. A limitation of this test is that it has low power for testing contemporaneous exogeneity. To test for contemporaneous exogeneity, we employ (Wooldridge 1995) robust score test, which is analogous to a Hausman test, but unlike the Hausman test does not assume that the error terms are independent and identically distributed.

To assess the robustness of our findings, we report the results based on all three estimators. In the discussion of our results, however, in cases where we reject the null hypothesis of strict exogeneity, we emphasize the results obtained using the IV estimator.

Results

Tables 5 and 6 report the results for all three estimators where income groups for the independent variables are defined based on free lunch eligibility and math proficiency, respectively.[29] For the FE and IV models, we report the estimated coefficient and standard error clustered at the district level. The estimated coefficient, the clustered standard error, and the average partial effect (APE) are reported for the PFP model.[30] The APE provides an estimate from the PFP model that can be compared to the estimated coefficients from the FE and IV models.

With no priors regarding potential endogeneity, we treated all of the affordable housing variables as potentially endogenous when conducting our strict exogeneity tests. Based on the results from either the score tests or the feedback tests we rejected the null hypothesis of strict exogeneity for the 14 affordable housing variables.[31] Given these findings, we will focus on the results obtained with the linear IV models. Interestingly, the estimated effects are remarkably stable across all three estimators.

[29]In the interest of brevity, the results tables only include the estimated effects of the affordable housing types. As expected, the unaffordable housing types are largely statistically insignificant.

[30]The estimation of average partial effects from the PFP model is straightforward. If z_{kit} is an arbitrary element of the vector of regressors Z_{it} and the parameter associated with z_{ikt} is denoted Γ_k, then we simply differentiate the conditional mean with respect to z_{ikt} to get the partial effect for observation i in period t

$$\frac{\partial E\left[P_{it}|Z_{it}\right]}{\partial z_{ikt}} = \Gamma_k \phi\left(Z_{it}\right)$$

where $\phi\,()$ denotes the standard normal probability density function. Papke and Wooldridge (2008) show that the average partial effect (APE_k) of z_{ikt} on P_{it} can be consistently estimated by simply averaging over all observations of the above equation. That is,

$$APE_k = \frac{\Gamma_k}{NT}\sum_{t=1}^{T}\sum_{i=1}^{N}\phi\left(Z_{it}\right)$$

[31]There are 16 exogenous variables: the unaffordable housing types and assisted housing, broken down into 8 variables registering the proportion of each type found within high income and middle income SAZs.

Table 5 Estimated effects of increases in the proportions of affordable housing types located in high and middle income school attendance zones on the proportion of free lunch students attending low income schools where income groups are defined by a school's global free lunch status

	Income group[1]					
	High income			*Middle income*		
	Model type			Model type		
Housing types	FE	PFP	IV	FE	PFP	IV
Single-family rental[2]	−0.225**	−0.953*** (−0.308)[3]	−0.289**	−0.058	−0.714 (−0.231)	−0.117
	[0.099][4]	[0.356]	[0.137]	[0.124]	[0.478]	[0.206]
Single-family owner	−0.091	−0.157 (−0.051)	−0.083	−0.320***	−0.843** (−0.273)	−0.251
	[0.082]	[0.284]	[0.127]	[0.108]	[0.339]	[0.233]
Mobile home rental	−0.205***	−0.763*** (−0.247)	−0.144**	−0.063	−0.274 (−0.089)	−0.189
	[0.067]	[0.263]	[0.054]	[0.075]	[0.245]	[0.162]
Mobile home owner	−0.080	−0.001 (0.000)	−0.247*	−0.136	−0.166 (−0.053)	−0.098
	[0.113]	[0.271]	[0.128]	[0.085]	[0.319]	[0.152]
Condominium rental	0.033	0.007 (0.002)	−0.028	0.059	0.075 (0.024)	0.05
	[0.067]	[0.239]	[0.092]	[0.061]	[0.237]	[0.078]
Condominium owner	−0.009	−0.027 (−0.009)	0.046	−0.011	−0.041 (−0.013)	−0.013
	[0.036]	[0.123]	[0.047]	[0.013]	[0.053]	[0.032]
Multifamily	−0.296***	−1.039*** (−0.336)	−0.203***	−0.252***	−0.853*** (−0.276)	−0.206**
	[0.049]	[0.173]	[0.063]	[0.076]	[0.305]	[0.078]
Assisted housing	−0.160***	−0.590** (−0.191)	−0.159**	−0.129**	−0.568* (−0.184)	−0.124*
	[0.055]	[0.279]	[0.062]	[0.058]	[0.346]	[0.064]
Observations	608	608	608	608	608	608

[1] High and low income schools are defined as those where the percentage of students qualifying for the free lunch program is greater than the districtwide median percentage of students qualifying for the free lunch program in every year of the panel. The middle income group is comprised of schools where the fraction of students eligible for the free lunch pgoram varied around the district median over the course of the panel

[2] "Owner" and "Renter" denote housing units that are owner- and renter-occupied, respectively

[3] Average partial effect reported in parentheses

[4] Standard error clustered at the school district level reported in brackets

*, **, *** indicate statistical significance at the 10%, 5%, and 1% levels, respectively

Estimated coefficients only displayed for affordable housing measures. Model also included unaffordable housing measures as well as year fixed effects

Table 6 Estimated effects of increases in the proportions of affordable housing types located in high and middle income school attendance zones on the proportion of free lunch students attending low income schools where income groups are defined by a school's global math exam pass rate

	Income group[1]					
	High income			*Middle income*		
	Model type			Model type		
Housing types	FE	PFP	IV	FE	PFP	IV
Single-family rental[2]	0.302**	−1.165	−0.355**	−0.030	0.016	0.058
		(−0.395)[3]			(0.005)	
	[0.117][4]	[0.396]	[0.149]	[0.132]	[0.374]	[0.216]
Single-family owner	0.087	0.462*	0.137	−0.273**	−1.279**	−0.408*
		(0.157)			(−0.434)	
	[0.077]	[0.270]	[0.140]	[0.128]	[0.581]	[0.229]
Mobile home rental	−0.453***	−1.336***	−0.514***	−0.270***	−0.825***	−0.261***
		(−0.453)			(−0.280)	
	[0.102]	[0.312]	[0.135]	[0.041]	[0.130]	[0.069]
Mobile home owner	−0.017	0.006	0.02	−0.038	−0.070	−0.046
		(0.002)			(−0.024)	
	[0.070]	[0.163]	[0.092]	[0.066]	[0.168]	[0.091]
Condominium rental	0.053	0.176	0.029	−0.012	0.044	−0.123
		(0.060)			(0.015)	
	[0.049]	[0.165]	[0.063]	[0.042]	[0.143]	[0.075]
Condominium owner	−0.025	−0.056	−0.037	0.070**	0.197**	0.146**
		(−0.019)			(0.067)	
	[0.019]	[0.076]	[0.027]	[0.030]	[0.099]	[0.058]
Multifamily	−0.313***	−0.965***	−0.358**	−0.334***	−1.122***	−0.436**
		(−0.327)			(−0.381)	
	[0.090]	[0.252]	[0.163]	[0.107]	[0.345]	[0.199]
Assisted housing	−0.294**	−0.866***	−0.288**	−0.175**	−0.501**	−0.153**
		(−0.294)			(−0.170)	
	[0.115]	[0.307]	[0.110]	[0.081]	[0.229]	[0.069]
Observations	608	608	608	608	608	608

[1] Schools with an average pass rate on the mathematics exam that were more than five percentage points above (below) the districtwide average were categorized as high income (low income). The remaining schools were classified as middle income

[2] "Owner" and "Renter" denote housing units that are owner- and renter-occupied, respectively

[3] Average partial effect reported in parentheses

[4] Standard error clustered at the school district level reported in brackets

*, **, *** indicate statistical significance at the 10%, 5%, and 1% levels, respectively

Estimated coefficients only displayed for affordable housing measures. Model also included unaffordable housing measures as well as year fixed effects

With but one exception, parameter estimates that were found to be statistically significant in the IV model were also statistically significant in the FE and PFP models, and the magnitude of the estimated effects were generally similar across all three models.

Table 7 Estimated effects of a 100 unit increase in the affordable housing type for the median school district where income groups are defined by global free lunch status

Housing types	A	B	C	D	E	F
High Income Schools[1]						
Single-family rental[2]	1950	5.13	2175	15	0.70	0.14
Mobile home rental	724	13.81	2175	25	1.13	0.08
Mobile home owner	784	12.75	2175	38	1.73	0.14
Multifamily	1737	5.76	2175	12	0.53	0.09
Assisted housing	710	38.61	2175	48	2.21	0.06
Middle income schools						
Multifamily	4171	2.40	2175	8	0.38	0.16
Assisted housing	1298	16.67	2175	29	1.33	0.08

[1] High and low income schools are defined as those where the percentage of students qualifying for the free lunch program is greater than the districtwide median percentage of students qualifying for the free lunch program in every year of the panel. The middle income group is comprised of schools where the fraction of students eligible for the free lunch pgoram varied around the district median over the course of the panel

[2] "Owner" and "Renter" denote housing units that are owner- and renter-occupied, respectively

A: Number of units of housing type within income group

B: Percentage increase in housing type within income group from 100 more units

C: Number of free lunch students attending low income schools

D: Decrease in the number of free lunch students attending low income schools from 100 unit increase in housing type within high income SAZ

E: Percentage decline in free lunch students attending low income schools from 100 unit increase in housing type within higher income SAZ

F: Elasticity of E with respect to B

Our results provide strong support for the hypothesis that improving the affordability of housing in higher income SAZs results in fewer poor students attending low-income schools. The coefficients on many of the affordable housing variables are negative and statistically significant, suggesting that increasing the affordable housing stock in middle- and high-income neighborhoods reduces socioeconomic segregation in public schools.[32] According to the results based on both income proxies, increasing the proportion of affordable single-family rentals, affordable mobile

[32] A limitation of our affordable housing stock measures is that because we only observe the free lunch eligibility status and not the income levels of the students that we have classified as low-income, we cannot directly account for the impact of changes in the income levels on affordability. If incomes fell for low-income families whose children attended the schools that we have classified as low-income, then these families may have been unable to afford to relocate to housing units that we have classified as affordable. If none of the units that we classify as affordable could be purchased or rented by low-income households, then changes in the stock of affordable housing in the high-income SAZs should have no impact on the location of low-income households. The fact that we find that increases in the stock of affordable housing in high-income SAZs reduces the concentration of low-income students in low-income schools suggests that, if anything, our estimated effects are understated because our affordable housing stock variables are measured with error.

home rentals, affordable multifamily housing units, and assisted housing units found in the high-income SAZs reduces the proportion of low-income students attending low-income schools. We also find a similar effect for an increase in the proportion of affordable multifamily housing and assisted housing found in the middle-income SAZs.

While our results are generally similar across the models based on the two different income proxies, there are, however, a few differences. For instance, when using the SAZ typology based on the eligibility for the free lunch program (Table 5), we find that increasing the proportion of affordable mobile homes located in high-income SAZs reduces the concentration of low-income students in the low-income schools. In contrast, when we utilize the math proficiency measure to construct the typology (Table 6), we find no such effect. Likewise, the coefficients on the affordable single-family owner-occupied home and mobile home rental variables are statistically significant in Table 5 for the middle-income SAZs but statistically insignificant in Table 6.[33]

While our results provide strong statistical evidence that the spatial distribution of the affordable housing stock influences the extent of socioeconomic segregation in public schools, the paramount question from a policy perspective is whether our findings are economically significant. To put context around the estimated magnitude of our results, we used our empirical models to predict the change in the dependent variable associated with a 100-unit increase in each of the housing types that were found to be have a statistically significant impact on the fraction of low-income students that attend low-income schools.[34]

In these calculations, we use the median values of the covariates when constructing the fitted values. Tables 7 and 8 report the results of this exercise for each of the income proxies. Two columns (D and F) are most relevant for assessing the economic significance of the results. Column D reports the predicted decline in the number of free lunch students attending low-income schools from the 100-unit increase in

[33] To investigate the extent to which our findings are sensitive to our choice of affordability cutoff, we reconstructed the affordable housing stock measures using two different cutoffs for the implied rent on the property: the one-bedroom and three-bedroom FMRs reported by HUD. The one-bedroom-FMR cutoff mechanically decreases the affordable housing stock relative to the stock measures based on the two-bedroom-FMR cutoff as the one-bedroom-FMR is lower than the two-bedroom-FMR. By similar logic, the affordable housing stock based on the three-bedroom-FMR cutoff is mechanically larger than the stock based on two-bedroom-FMR cutoff. We used these new affordable stock measures to re-estimate the models that are reported in the main text. The results are remarkably similar across the models based on the three different affordability measures. In all but a few cases, if the parameter associated with one of the affordable housing types is statistically significant (at the 10 percent level) for the two-bedroom-FMR definition of affordable housing, it is also statistically significant for the one- and three-bedroom FMR affordability definitions. The only notable exception is that the proportion of single-family rentals located in high-income SAZs, while significant across the board using the two-bedroom FMR to define affordability, is insignificant in the models using the one-bedroom-FMR cutoff and the exam scores to define the school groupings. The coefficient on the same single-family rental variable is also statistically insignificant (but only marginally, with p-values ranging between .11 and .12) in the models where affordability is based on the three-bedroom FMR and the school groups are defined by the Global Free Lunch Status.

[34] To illustrate, we added 100 units of single-family rentals to the high-income group of SAZs and then calculated the effect this would have on the proportion of the district's single-family rentals located in high-income SAZs.

Table 8 Estimated effects of a 100 unit increase in the affordable housing type for the median school district where income groups are defined by global math test pass rates

Housing types	A	B	C	D	E	F
High income schools[1]						
Single-family rental[2]	1741	5.74	2385	19	0.80	0.14
Mobile home rental	397	25.19	2385	102	4.29	0.17
Multifamily	1594	6.27	2385	21	0.86	0.14
Assisted housing	674	37.31	2385	85	3.58	0.10
Middle income schools						
Mobile home rental	1127	8.87	2385	35	1.46	0.16
Multifamily	3512	2.85	2385	19	0.78	0.27
Assisted housing	937	19.55	2385	38	1.58	0.08

Schools with an average pass rate on the mathematics exam that were more than five percentage points above (below) the districtwide average were categorized as high income (low income). The remaining schools were classified as middle income

"Owner" and "Renter" denote housing units that are owner- and renter-occupied, respectively

A: Number of units of housing type within income group

B: Percentage increase in housing type within income group from 100 more units

C: Number of free lunch students attending low income schools

D: Decrease in the number of free lunch students attending low income schools from 100 unit increase in housing type within high income SAZ

E: Percentage decline in free lunch students attending low income schools from 100 unit increase in housing type within higher income SAZ

F: Elasticity of E with respect to B

a particular type of affordable housing. Focusing first on the results in Table 7, we see that the impact of changes in the distribution of the assisted housing stock is quite large: a 100-unit increase in these units in high-income SAZs results in 48 fewer low-income students in low-income schools, while if the increase is in middle-income SAZs, the number of socioeconomically disadvantaged students attending low-income schools is predicted to decline by 29. These results are not surprising as these subsidized units are frequently targeted at low-income families with children.

We also find economically large impacts associated with changes in the distribution of mobile homes across SAZs.[35] Specifically, we estimate that adding 100 renter-occupied (owner-occupied) mobile homes to a high-income SAZ would result in 25 (38) fewer low-income students attending low-income schools.

In Table 8 assisted housing and mobile home rentals again show relatively large effects, but in comparison to Table 7 the mobile home effects are substantially

[35] In Florida, mobile homes are an important part of the housing stock. If we define the total single-family stock as the sum of detached single-family units, condominiums, and mobile homes, on average mobile homes constituted 14 percent of the total number of single-family units. Condominiums, on the other hand, comprised only 13 percent of this total.

larger, with a 100-unit increase in the stock mobile home rental stock in high-income and middle-income SAZs estimated to reduce the number of low-income students attending low-income schools by 102 and 35, respectively.

Another approach to gauging the magnitude of effects implied by our empirical models is to estimate the percentage decline in the number of low-income students attending low-income schools when the number of affordable housing units located in a high- or middle-income SAZs is increased by 1 percent. These elasticities, which are reported in column F, range from 0.06 (assisted housing, high-income SAZ) to 0.16 (multifamily housing, middle income SAZ) in Table 7 and from 0.08 (assisted housing, middle income SAZ) to 0.27 (multifamily housing, middle income SAZ) in Table 8. The latter estimate suggests that a 10 percent increase in the number of apartments located within a middle-income SAZ would result in a 3 percent decline in the number of free lunch students attending low-income schools.[36]

Regardless of whether we focus on changes in levels or elasticities, our results demonstrate that increasing affordable housing within higher income SAZs has nontrivial impacts on the concentration of free lunch students within low-income schools.

Conclusion

The educational underachievement of children from poor families represents a major crisis within America's educational system. Many top scholars view the socioeconomic integration of schools as the best strategy for addressing the crisis, and providing more affordable housing in high-income neighborhoods zoned for high-performing schools is frequently suggested as one way to achieve this integration. There are reasons to believe, however, that using housing policy to reduce socioeconomic segregation in schools may be unsuccessful because poor families with school-age children may fail to occupy the affordable units because such families are unwilling to relocate or they find that they cannot take advantage of these housing opportunities because of discrimination.

In this paper we have offered the first evidence on whether providing more affordable housing units within high-income neighborhoods will in fact decrease the number of low-income students attending low-income schools. Our results are based on a natural experiment provided by Florida's volatile housing market over the past decade and a half. The results show that a range of affordable housing types can make a difference, especially apartments and mobile homes. Our findings are at odds with previous research that suggested that low-income families do not take the opportunity to improve their child's school environment when the opportunity presents itself. These results suggest that housing policy can play a significant role in

[36]The affordable housing types with relatively large elasticities and those yielding relatively large effects from an increase in their number differ because the percentage change is dependent on the base, which we measure as the median number of existing units found within the middle- or high-income SAZs.

addressing socioeconomic segregation in public schools and, potentially, the yawning achievement gap between low- and high-income students.

References

Andersson, F., & Mayock, T. (2014). Loss severities on residential real estate debt during the great recession. *Journal of Banking & Finance*, *46*, 266–284.

Angrist, J., & Krueger, A. (2001). Instrumental variables and the search for identification: From supply and demand to natural experiments. *The Journal of Economic Perspectives*, *15*(4), 69–85.

Campbell, J. et al. (2011). Forced sales and house prices. *American Economic Review*, *101*(5), 2108–31.

Carnevale, A., & Strohl, J. (2010). How Increasing College Access is Increasing Inequality, and What to Do About It. In Kahlenberg, R. (Ed.) *Rewarding Strivers: Helping Low-Income Students Succeed in College* (pp. 71–190). New York: Century Foundation.

Dahl, G., & Lochner, L. (2012). The impact of family income on child achievement: Evidence from the earned income tax credit. *The American Economic Review*, *102*(5), 1927–1956.

Davis, T. (2014). School choice and segregation: 'Tracking" racial equity in magnet schools. *Education and Urban Society*, *46*(4), 399–433.

Ellen, I. et al. (2016). Why don't housing choice voucher recipients live near better schools? insights from big data. *Journal of Policy Analysis and Management*, *35*(4), 884–905.

Fischel, W. (1985). *The economics of zoning laws: a property rights approach to american land use controls*. Baltimore: Johns Hopkins University Press.

Fischel, W. (2001). *The homevoter hypothesis: How home values influence local government taxation, school finance, and Land-Use policies*. Cambridge: Harvard University Press.

Friedman, M. (1955). *The role of government in education rutgers university press*. NJ: New Brunswick.

Fry, R., & Taylor, P. (2012). The Rise of Residential Segregation by Income. *Pew Social and Demographic Trends*.

Gyourko, J. et al. (2008). A new measure of the local regulatory environment for housing markets: The wharton residential land use regulatory index. *Urban Studies*, *45*(3), 693–729.

Hintermaier, T. (2016). Exemption of Homesteads. Florida Statute 196.031.

Horn, K. et al. (2014). Do housing choice voucher holders live near good schools?. *Journal of Housing Economics*, *23*, 28–40.

Ihlanfeldt, K. (Forthcoming). The Deconcentration of Minority Students Attending Bad Schools: The Role of Housing Affordability within School Attendance Zones Containing Good Schools. *Journal of Housing Economics*.

Ihlanfeldt, K., & Mayock, T. (2016a). The Impact of REO Sales on Neighborhoods and their Residents. *The Journal of Real Estate Finance and Economics*, *53*(3), 282–324.

Ihlanfeldt, K., & Mayock, T. (2016b). The variance in foreclosure spillovers across neighborhood types. *Public Finance Review*, *44*(1), 80–108.

Kahlenberg, R. (2013). From all walks of life: New hope for school integration. *American Educator*, *36*(4), 1–14.

Levitt, R. (2014). Evidence Matters: Paired Testing and the Housing Discrimination Studies. U.S. Department of Housing and Urban Development Report.

Lens, M., & Monkkonen, P. (2016). Do strict land use regulations make metropolitan areas more segregated by income?. *Journal of the American Planning Association*, *82*(1), 6–21.

Mickelson, R., & Bottia, M. (2009). Integrated education and mathematics outcomes: a synthesis of social science research. *North Carolina Law Review*, *88*(3), 993–1089.

Papke, L., & Wooldridge, J. (2008). Panel data methods for fractional response variables with an application to test pass rates. *Journal of Econometrics*, *145*(1), 121–133.

Pendall, R., et al. (2006). From Traditional to Reformed: A Review of the Land Use Regulations in the Nation's 50 Largest Metropolitan Areas. Brookings Institution Metropolitan Policy Program Report.

Reardon, S. (2011). The Widening Academic Achievement Gap between the Rich and the Poor: New Evidence and Possible Explanations. In Duncan, G., & Murnane, R. (Eds.) *Wither Opportunity? Rising Inequality, Schools, and Children's Life Chances* (pp. 91–15). Chicago: Russell Sage Foundation.

Rothstein, R. (2013). For public schools, Segregation Then. Segregation Since. Economic Policy Institute Report.

Rothwell, J., & Massey, D. (2010). Density zoning and class segregation in U.S. Metropolitan areas. *Social Science Quarterly*, *91*(5), 1123–1143.

Rothwell, J. (2012). Housing Costs, Zoning, and Access to High-Scoring Schools. Brookings Institution Metropolitan Policy Program Report.

Sanbonmatsu, L., et al. (2011). Moving to Opportunity for Fair Housing Demonstration Program. U.S. Department of Housing and Urban Development Report.

Schuetz, J. et al. (2008). Neighborhood effects of concentrated mortgage foreclosures. *Journal of Housing Economics*, *17*(4), 306–319.

Schwartz, A., & Stiefel, L. (2014). Linking housing policy and school reform. In Lareau, A., Goyette, K., Schwartz, A., Stiefel, L. (Eds.) (pp. 295–314). New York: Russell Sage Foundation.

Wooldridge, J. (1995). Score Diagnostics for Linear Models Estimated by Two Stage Least Squares. In Maddala, G., Srinivasan, T., Phillips, C. (Eds.) *Advances in Econometrics and Quantitative Economics: Essays in Honor of Professor C.R. Rao* (pp. 66–87). Oxford: Blackwell.

Wooldridge, J. (2010). Econometric Analysis of Cross Section and Panel Data, MIT press, Cambridge.