Title
Essays on Market-Based Provision of Local Public Services

Author
Clarke, Wyatt

Publication Date
2018

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA,
IRVINE

Essays on Market-Based Provision of Local Public Services

DISSERTATION

submitted in partial satisfaction of the requirements
for the degree of

DOCTOR OF PHILOSOPHY

in Economics

by

Wyatt Clarke

Dissertation Committee:
Associate Professor Damon Clark, Chair
Associate Professor Matthew Harding
Professor Jan Brueckner

2018

DEDICATION

To my wife, Ngoshali. You did this with me.

Of making many books there is no end, and much study wearies the body.

Now all has been heard;
here is the conclusion of the matter:
Fear God and keep his commandments,
for this is the duty of all mankind.

Book of Ecclesiastes

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGMENTS

I am grateful to Damon Clark, Matthew Harding, and Jan Brueckner for their patient instruction, encouragement, and accountability.

I thank the Indiana Department of Education and Zillow Group for providing data used in these essays.

Finally, I thank the University of California, Irvine and California tax payers for funding my studies.

CURRICULUM VITAE

Wyatt Clarke

EDUCATION

Doctor of Philosophy in Economics **2018**
University of California, Irvine *Irvine, CA*

Master of Arts in Economics **2015**
University of California, Irvine *Irvine, CA*

Mandarin Language Studies **2011**
Xi'an Jiaotong University *Xi'an, China*

Bachelor of Science in Industrial Management **2009**
Purdue University *West Lafayette, IN*

RESEARCH

'One Quarter of Hispanic Children in the United States Have an Unauthorized Immigrant Parent', *National Research Center on Hispanic Children and Families* (2017). Research brief.

EXPERIENCE

Teaching Assistant **2013–2017**
University of California, Irvine *Irvine, CA*

Summer Research Fellow **2015**
Child Trends *Bethesda, MD*

English Learner and Migrant Education Specialist **2012-2013**
Indiana Department of Education *Indianapolis, IN*

SOFTWARE SKILLS: Python, ArcGIS, R, Stata, MATLAB

LANGUAGE SKILLS: Spanish (fluent), Portuguese (conversational), Chinese (basic)

ABSTRACT OF THE DISSERTATION

Essays on Market-Based Provision of Local Public Services

By

Wyatt Clarke

Doctor of Philosophy in Economics

University of California, Irvine, 2018

Associate Professor Damon Clark, Chair

This dissertation comprises three self-contained essays. Each essay uses administrative data collected by state and local governments to evaluate a market-based innovation in the delivery of local public services. Chapters 1 and 2 explore school choice programs that were implemented in the state of Indiana starting around 2008, while Chapter 3 is a national study of homeowners associations.

In Chapter 1, student-level standardized test scores are studied to characterize the joint effectiveness of three forms of school choice—public charter schools, private school vouchers, and inter-district open enrollment. Average test scores of all students residing in a school district are shown to improve when the share of students enrolled through one of these options increases.

Chapter 2 narrows the focus to a single school choice policy—inter-district open enrollment. Prior studies show that compelling public schools to enroll non-resident students weakens housing demand in communities with good schools, lowering their home prices. Indiana's policy of allowing public schools the option to enroll non-resident students does not exhibit the same relationship, possibly because schools selectively participate based on the local price elasticity of demand for housing (i.e., the responsiveness of local house prices to changes in demand).

Chapter 3 demonstrates how to detect homeowners associations across most of the country using mortgage riders documented in public real estate records. These private neighborhood governments are shown to involve a fifth of all single-family houses in the U.S. and 80 percent of houses in recently-built subdivisions. Houses with a homeowners association are estimated to cost 4 percent more than nearby similar homes with no homeowners association.

Chapter 1

Charters, Vouchers, and Open Enrollment: The Overall Effects of School Choice in Indiana

1.1 Introduction

School choice programs have proliferated across the United States in the past two decades. Forty-three of the fifty states currently allow charter schools, twenty-four states provide vouchers or scholarships to subsidize private schooling, and forty-four states allow inter-district open enrollment in public schools. Twenty states allow all three of the above programs in some form (Education Commission of the States 2015, 2016, National Conference of State Legislatures 2016).

Most academic studies on the topic of school choice focus on a single type of program, but it is increasingly common for a locality to operate several types of school choice programs at the same time. Focusing on one program to the exclusion of others can bias the results of a study. The solution presented here is to study school choice programs as a package, under the theory that they all affect student performance via the same basic mechanisms.

The state of Indiana is used as a case study for this purpose. From 2005 to 2015, the share of Indiana K-12 students enrolled in a charter school, receiving a voucher, or attending public school in a neighboring school district rose from 1 percent to almost 10 percent. The policies allowing for this change were implemented in the space of a few years, making it impossible to clearly disentangle the effect of any single innovation. All three types of school choice described operate statewide with no enrollment caps and involve similar numbers of students. The authorizing state laws for Indiana's school choice programs adhere closely to model laws promoted by education reform groups (Ziebarth & Bierlein Palmer 2014), helping lessons learned in Indiana to apply in other states too.

The empirical strategy used in this paper is to compare test scores over time in different regions of the state. The makeup of students living in each region (school district) should not have been affected by school choice even as the regions experienced varying degrees of exposure to the policy's effects.[1] As an illustrative example, 44 percent of students living in Gary, Indiana's school district who took the state's annual standardized test (ISTEP+)[2] in 2015 opted into school choice: they attended a charter school, a private school paid with a voucher, or the traditional public school in another district. Thirty minutes' drive away, only 3 percent of test takers residing in Valparaiso's school district opted into school choice. The average test score of Gary students rose from the 23rd percentile of statewide scores in 2005 to the 25th percentile in 2015. Meanwhile, Valparaiso students' average score fell from the 66th statewide percentile to the 62nd percentile. It is impossible to draw conclusions

[1] Hsieh & Urquiola (2006) pursue a similar strategy using data from Chile, but this approach has not yet been used with U.S. data. It should be possible to use this study design in any state since most of the necessary data are collected to satisfy federal reporting requirements.

[2] Indiana's Statewide Test for Educational Progress Plus (ISTEP+) is a state standardized test administered annually to all students in grades 3–8.

from only two school districts, but, if this type of pattern is repeated statewide, it leads one to think that school choice is improving students' test scores.

A salient feature of this study, evident from the example above, is that students are grouped by *residence* in a school district rather than by *enrollment*. This approach addresses a common challenge in studying the effects of school choice: that the composition of a school's enrollment changes when new schools are introduced, making it hard to track whether variation in test scores is due to variation in school quality or a change in the profile of students at the school.

Despite the fact that this study uses only test score averages, they are not the averages released to the public. Indiana and other states release test scores aggregated by enrollment in a school or school system (called a "school corporation" in Indiana) but not by residence in a school district. Before the introduction of school choice, enrollment served as a good proxy for residence. Now, communities with a high proportion of students opting into school choice have no good way of evaluating the testing success of their students with publicly-released data. The average scores used in this paper are calculated by the author on the basis student-level scores provided by the Indiana Department of Education. If states were to release test scores aggregated by students' residence, the type of research performed here could be done more easily.

In the remainder of the paper, Section 1.2 reviews a selection of relevant literature; Section 1.3 presents an overview of Indiana's K–12 educational system; Section 1.4 develops the estimation strategy; Section 1.5 presents the main results; Section 1.6 presents a series of robustness checks and extensions; and Section 1.7 concludes.

1.2 Related Literature

The paper most similar to this one is Hsieh & Urquiola (2006). These authors evaluate the effect of a nation-wide private school voucher program introduced in Chile in 1981. They leverage the differential effects of this policy in Chile's roughly 300 municipalities and find no evidence that greater availability of private school vouchers has led to improved student outcomes. As done in this paper, Hsieh & Urquiola (2006) investigate the performance of all children living in a geographic area rather than comparing outcomes of students who attend different school types. The data used here are at the student level whereas Hsieh & Urquiola (2006) use data aggregated to the school level.

Another related paper is Figlio & Hart (2014), which uses a 1-year lag that occurred between passage and enactment of Florida's voucher law in 2001 to identify competitive effects of vouchers on student test scores. The authors show that traditional public schools (TPS) with many private schools nearby improved relative to TPS's with few private school neighbors after statewide authorization of private school vouchers. Because such a large proportion of publicly funded students enrolled in traditional public schools in the context of their study, this is akin to studying the average performance of all publicly-funded students. The voucher program enacted in Indiana is quite similar to the program studied by Figlio & Hart (2014) in Florida.

The Indiana-specific research that is most related to this study was conducted by Stanford University's Center for Research on Education Outcomes (CREDO). CREDO (2012) uses a form of propensity score matching to measure the test score growth of students enrolled in charter schools relative to the test score growth of comparable students enrolled in traditional public schools on Indiana state tests. The goal of CREDO's research—to find the effect on a student's academic performance of enrolling in a charter school—is different from the goal of this paper, which is to learn how the performance of all students changes after implementation of school choice policies. CREDO (2012) draw from similar data to what is used in this paper.

1.3 Background Information about the Indiana K–12 Education System

The state of Indiana is divided into 289 geographic areas called *school districts*. Each district has a local government called a *school corporation*, with an elected board and taxation power, that is responsible to educate students who live within its geographic boundaries. Each school corporation operates multiple schools. Students that live within a school district are typically assigned by the school corporation to a "school of legal settlement" based on their residence. This is traditionally where the student has attended.[3] Until recently, in order for students to attend a public school outside of their district, they have had to meet the receiving school corporation's acceptance criteria and pay tuition, similar to a private school. In this way, school corporations have acted as local monopoly providers of publicly-funded K–12 education. The avenues for families to choose a different school have traditionally been paying to attend private school, homeschooling, or moving to another school district.

1.3.1 Optional Inter-District Open Enrollment

Recent changes to state law have altered the system described above. In 2008, state law was changed so that funding for Indiana school corporations' General Fund is provided completely by the state government on a per student basis, with only facilities, busing, and debt service

[3] Whether students can attend another school within their district depends on school corporation policy and varies from district to district. These "intra-district open enrollment" policies are not reported in data reported to the state, but it can be shown that they do not align with patterns observed in student test scores reported later in the paper.

allowed to be funded by local property taxes. The change to a "funding follows the child" system has allowed an accompanying increase in inter-district enrollment. Since 2010, school corporations have been allowed to enroll students who are not residents of their district and receive the same funding for those students as they would receive for enrolled residents. Since 2013, school corporations have been barred from selectively admitting non-residents; they must employ a blanket policy of accepting or not accepting non-resident (or "transfer") students up to a limit set each year by the school corporation. Transfer students must provide their own transportation and may be required to pay some tuition to cover the portion of school costs funded by local taxes. This change describes an "optional inter-district open enrollment policy." In 2014–2015, most Indiana school corporations enrolled non-resident students, and 36,000 students, or 3.2 percent of publicly-funded Indiana students, enrolled in a traditional public school ("TPS") outside of their own district.

1.3.2 Charter Schools

Another change has been the introduction of charter schools, which are a new organizational form of public schooling in the state. Each charter school acts as an independent school corporation and is not tied to a geographic district. It is operated by a private, non-profit organization rather than by a local government.[4] Since charter schools do not have taxation power, all funding comes from state and federal sources. Charter schools must accept all students who apply or, if more students apply than can be accommodated, accept students by lottery. Students do not pay tuition but usually must provide their own transportation. Because charter schools operate inside another school corporation's geographic district and can enroll students from anywhere in the state, they are sometimes seen as competitors of traditional public schools.

[4]Because of the private management, charter schools are sometimes described as private. However, Indiana law defines them as public schools.

Charter schools are overseen by a sponsor, which issues (and can revoke) their "charter." The first charters in Indiana were issued in 2002. In 2010, the legislature organized a statewide Charter School Board to issue charters and also allowed universities and city governments to issue their own charters. In 2015, a total of 75 charter schools were operating in Indiana with combined enrollment of 32,500 students, or 2.9 percent of publicly-funded Indiana students. An additional 11,000 students were enrolled in 5 "virtual" charter schools, which allow students to pursue their education off-site over the internet.

1.3.3 Vouchers

The newest form of school choice in Indiana is means-tested private school vouchers. Students who qualify via at least one of seven "pathways" established by the state legislature can apply for a voucher to spend on tuition at a private school of their choosing. Each pathway has a household income limit[5] and requires students to meet one other criterion, including attendance at a public school the prior year, previous receipt by the student or a sibling of a voucher (or scholarship through a related program), residence in the attendance zone of an "F" rated school, or designation as a special education student. Depending on a student's household income, the voucher is worth either 50 percent or 90 percent of the amount the state would reimburse a TPS for educating that student. Private schools must also fulfill some requirements to receive Choice Scholarship funds, including obtaining state accreditation, annually administering ISTEP+ to Choice Scholarship students, and reporting certain data to the state (which is used for this study). In 2015, 29,000 students, or 2.6 percent of publicly-funded Indiana students, used Choice Scholarships to attend 314 private schools.

[5]Depending on the pathway, the household income limit ranges from 100 percent to 200 percent of the limit to receive a reduced price school lunch, which is 185 percent of the federal poverty level. For a 4-person household in 2014–2015, this meant a household income limit of $44,123 or $88,245 (IDOE 2015).

1.3.4 Distribution of Students and Schools

In 2015, 95.4 percent of primary and secondary school children in Indiana used publicly-funded education. Of those, 86.1 percentage points enrolled in a TPS in the school district where they lived. 2.8 percent of Indiana children attended a charter school, 2.5 percent used a voucher to attend a private school, 3.1 percent transferred to a TPS in another school district, and 1.0 percent enrolled in a virtual charter school (Figure 1.1).

While traditional public schools remain the dominant education providers for students in their own districts, the share of students enrolled in choice schools is growing quickly. In the most extreme case (Gary), 44 percent of publicly-funded students enroll outside of the TPS in their district (Figures 1.2 and 1.3).

Population density and district-level average test scores vary geographically across the state as shown in Figure 1.4. See Figure 1.5 for the 2014 geographic distribution of charter schools, private schools that accept vouchers, and traditional public schools that accept transfer students. Note that most charter schools are tightly clustered in densely populated areas with poor school performance. Voucher-accepting private schools are concentrated more generally in densely populated areas, and TPS's which accept transfer students are spread throughout the state.

1.4 Data, Motivation, and Estimation Strategy

1.4.1 Data

The data used for this research include anonymized scores for every student who took Indiana's ISTEP+ test from 2005 to 2015, the school district where each K–12 student in

Indiana lived, the schools they attended, and students' reported race, sex, English language proficiency, and eligibility to receive a free or reduced price lunch.[6] [7]

1.4.2 Motivating Framework

To motivate the discussion, the following framework shows how changes in availability of school choice can affect students' test scores. Similar ideas are assumed or stated in much of the related literature including Barseghyan et al. (2014), Figlio & Hart (2014), Hsieh & Urquiola (2006), and Altonji et al. (2015).

Let us assume that families choose whichever school gives them the highest utility, with quality of instruction being one of the factors they consider. Additionally, assume that school administrators pursue a set of priorities which need not be defined except to say that one priority is to maximize enrollment and thus their school's budget. These administrators allocate the budget between uses that do or do not contribute to quality of instruction. When new school choices become available, some students switch schools, and school administrators may react competitively by shifting their budget toward more academically productive uses. (Alternatively, administrators might spend more on academically non-productive uses like sports teams to attract students.) We expect this to affect students' test scores (1) if they move to another school, (2) if their peer group changes as a result of other students switching school, or (3) if their school shifts more (or less) of its budget toward academically productive uses. In this environment, there should be a relationship between the degree of school choice available to a group of students and their test scores.[8]

[6] These data are provided by the Indiana Department of Education.

[7] The ISTEP+ is administered to all publicly-funded students in grades 3–8 with the exception of some special education students. Eligibility for a free or reduced price lunch will be used as a proxy for low family income since it is an income-qualified program.

[8] A similar, commonly used assumption is that schools try to maximize enrollment while minimizing costly "effort." The author prefers the present environment since it allows for the possibility that students might be drawn to a school via non-academically productive budget allocations (Hastings et al. 2009). Additionally,

The degree of school choice is measured in this study as the percentage of a district's residents in test-taking grades participating in school choice. This quantity is similar to a Herfindahl Index. Students are grouped by residence in a school district since that is the level at which choice enrollment can be measured with the available data.[9]

1.4.3 First Look: Scatter Plot of Changes

The goal of this paper is to find the effect of an increase in school choice on test scores. Following the intuition presented before regarding Gary and Valparaiso, this can be done roughly by comparing the change in percentage of a district's residents who enroll in choice schools to the change in average test score of students in that district. Figure 1.6 shows a scatter plot of these changes over the decade of 2005–2015 for each district in the state. The fitted line suggests a positive relationship, meaning that on average increased school choice is accompanied by improved school performance.[10,11]

this environment imputes motives to school administrators rather than to abstract "school" agents, while remaining agnostic about the motivations of teachers.

[9] A student's decision to switch schools has both private effects and spillover effects for their old and new classmates. This is one reason why the net effect of school choice has to be evaluated for a group of students rather than for individuals.

Other ways of grouping students could also work as long as a group retains students who switch schools and contains all students who receive spillover effects of those students' moves. A problem with grouping students by district is that spillover effects will cross between groups when students from different districts attend school together. For example, students from a district whose traditional public schools accept transfer students might be affected by these new peers even though this is not measured as "school choice" for the receiving district. This problem could be minimized by grouping students into larger geographical areas, but that would decrease the sample size of groups. Hsieh & Urquiola (2006) group students by municipality in Chile.

[10] This relationship remains when districts with the most choice, those on the far right of Figure 1.6, are removed from the sample.

[11] The fitted line in Figure 1.6 is produced by a regression of each district's change in average test score from 2005 to 2015 on the district's 2005–2015 change in proportion of students enrolled in choice schools,

$$Y_d^{14-05} = \alpha + \beta C_d^{14-05} + \epsilon_d.$$

The coefficient β is estimated to be 0.379 (0.113) and provides the slope of the line. Student test scores are transformed so that the scores of all students in the same grade and year have mean zero and standard deviation of one. An interpretation of this regression is that a one standard deviation increase a district's choice enrollment (5.5 percent choice enrollment) predicts a 0.02 standard deviation increase in the test score of all students residing in that district, relative to all other students in the state.

1.4.4 Main Estimation Strategy: Fixed Effect Regression Models

The positive relationship between increased school choice and test scores found in a simple scatter plot (Figure 1.6) persists when tested in a fixed effects regression model that utilizes all years of data, not just 2005 and 2015. This is done in a regression,

$$Y_{dt} = \alpha + \beta_0 C_{dt} + \beta_1 D_d + \beta_2 T_t + \epsilon_{dt} \tag{1.1}$$

where Y_{dt} is the average test score of students residing in district d in school year t; α is an intercept term; C_{dt} is the availability of school choice in district d in school year t; D_d and T_t are indicator variables for district and year; and ϵ_{dt} is an error term.[12] The coefficient β_0 expresses whether a district's best test scores came in years when its students had the most choice.[13]

One would like to use this regression (and the scatter plot before it) to infer whether increased choice drives improvements to test scores. Whether such inference is valid depends on what else was happening at the same time. If districts' scores would have remained constant relative to each other absent the introduction of choice, β_0 represents the causal effect of introducing school choice. Otherwise, β_0 could be expressing the impact of some other change that accompanied the introduction of choice.

Since school choice was not assigned to districts randomly, the causal explanation of β_0 cannot be proven.[14] Instead, we can try to eliminate other credible explanations for why

[12] This formulation resembles a regression difference-in-difference model (Angrist & Pischke 2008) but with a continuous treatment (Acemoglu et al. 2004). As such, the identified quantity resembles the average effect of treatment on the treated (ATT). This means that the effects observed in this study may not be predictive of what would happen were school choice introduced at other levels or in other school districts of Indiana.

[13] β_1 adjusts for a district's overall average score and β_2 adjusts for the average score of all districts in a year. The remaining variation is changes in districts' scores relative to each other. These changes will be explained by β_0 if they are correlated with changes in district-level school choice C_{dt} or by ϵ_{dt} otherwise.

[14] Both supply and demand for school choice are endogenously determined. This means that the effect of school choice on test scores observed in Indiana may be higher or lower than what would be expected if school choice were randomly assigned. One might expect endogenous demand to raise the observed effect since students who enroll in choice schools are presumably those who can benefit most. It is unclear what

test scores might have changed in the hope that this strengthens the suggestive case that school choice caused changes in test scores. A first step is to add other covariates to the regression which could be driving changes in test performance.[15] This is done in Section 1.5 by adding a vector of controls X_{dt} to the regression—the percentage of district residents classified as white, black, Hispanic, limited English proficient (by level), male, and eligible for a free or reduced price lunch. Another solution is to test whether districts which received choice were already improving before choice was introduced. In Section 1.6, this is shown not to be the case. Section 1.6 also checks to ensure that test score shifts are not attributable to changes in the pool of test takers due to the introduction of choice. These precautions still leave risk that changes to test scores could be due to factors other than the introduction of school choice, but only if those factors were correlated with the introduction of choice.

The following headings address two additional aspects of this case that require consideration.

Weighted or Unweighted by Enrollment

School district is the unit of analysis since both measures of school choice are common to all students residing in a district.[16] Some districts are more populous than others, leaving a choice of whether to weight districts by student enrollment in the regression analysis. Estimates from an enrollment-weighted regression will demonstrate the average *student's*

to expect from endogenous supply since the mechanisms by which school choice is supplied are unclear and different for each type of choice. Instrumental variables are often used to learn the direction of endogeneity bias, but it is not obvious what would make a good instrument here. Possible candidates are the pre-treatment proportion of a districts' residents who were designated as black, who qualified for a free or reduced lunch, or who used public transit to commute interacted with the average level of school choice take-up in the state each year. Each of these is strongly predictive of school choice take-up, but is likely to be correlated with changes in student test scores over a 10-year period even absent the introduction of school choice. The results of IV regression using these instruments are imprecise and mixed in magnitude. They are available on the author's website.

[15]Or factors which are correlated to the driver of performance. If districts which saw change in these variables over the study period systematically gained or lost ground relative to other districts, these test score changes will be "soaked up" by the control variables.

[16]Using individual test scores Y_{idt} would allow for individual level control variables like student's own race and gender, but adding these controls has virtually no effect on the regression estimates.

experience. Estimates from an unweighted regression will show the average *district's* experience.[17] Both students' and districts' seem interesting, so weighted and unweighted regression models are both presented.

Averaging the Effect of School Choice Across Test Subjects, Grades, Years of Exposure, and Type of School Choice

The baseline model presented above covers over quite a lot of complexity. The student test scores used are an average of each student's math and English language arts (ELA) scores, standardized so that the scores of all students in the same grade and year have mean zero and standard deviation of one;[18] no distinction is made between scores on the basis of grade or the number of years a student has been exposed to the current level of school choice; and no distinction is made between exposure to the different types of school choice. Thus, the estimates presented here portray the *average* effect of school choice across all of these distinctions. Separate estimates are provided in Section 1.6 by test subject, by grade, and by type of school choice to allow a less aggregated view albeit with less statistical power.

1.5 Main Results

Table 1 presents the coefficient on *% Choice* from the regressions discussed in Section 1.4.4, which are the main empirical findings of this paper. Column 1 shows exactly the regression expressed by Equation 1, wherein the average normalized test score of students residing in one school district and year is regressed on the share of those students who are enrolled in school choice and fixed effects for each district and year. The districts are not weighted by

[17]Interpreting the unweighted estimate as the effect for the average student would overweight the experiences of students in small districts. This is analogous to how votes of citizens in less populous states carry greater weight in American presidential elections because of the Electoral College system.

[18]Figlio & Hart (2014) and other authors make similar standardizations.

enrollment. The coefficient on β_0, the share of students enrolled in school choice, is 0.244 (with standard error 0.116). Because of the way test scores are normalized, this means that a one percentage point increase in choice enrollment in a district predicts an average increase in the test score of all students living in that district of 0.00244 standard deviations, relative to the scores of all other students in the state. Assuming test scores are normally distributed, that would correspond to a student's score moving from the 50.0$^{\text{th}}$ percentile to the 50.1$^{\text{th}}$ percentile when one out of a hundred students in her district enrolls in school choice.

Column 2 of Table 1 adds a list of control variables to the regression—the percentage of a district's enrollment that is black, white, Hispanic, male, limited English proficient levels 1–4, and eligible for a free or reduced price lunch. Adding these regressors helps control for the effect of changes to the demographic make-up of districts that might have happened concurrent with the introduction of school choice. Columns 3 and 4 weight the school districts by enrollment, thus focusing on the average experience of individual students rather than that of school districts. Column 3 does not include control variables; Column 4 does. The point estimates have similar magnitude in all four columns, with regressions that include control variables yielding somewhat larger estimates.

The author's preferred specification is Column 4 which includes control variables and is weighted by district size to capture the average Hoosier students' test performance. In a district where 9.3 percent of students enroll in school choice—the state's overall level of school choice take-up in 2015—the coefficient of 0.443 (0.117) would correspond to a student's test score moving from the 50.0$^{\text{th}}$ percentile to the 51.6$^{\text{th}}$ percentile. Note that this change in expected test score improvement applies to all students in the school district, not just those who participate in school choice.

Stated in terms comparable to related literature, Column 4 indicates that a one standard deviation increase in school choice enrollment of 6.0 percent predicts a 0.026 standard deviation increase in student test scores.[19] Figlio & Hart (2014) report a similar 0.02 standard deviation increase in test scores resulting from schools' competitive response to a one standard deviation increase in exposure to vouchers in Florida. Hsieh & Urquiola (2006) find no evidence of improvement in Chilean test scores after the introduction of universal vouchers.

A separate question from the overall effect of school choice is how the benefits of school choice are distributed. Could school choice be driving better scores for choice participants but worse or unchanged outcomes for students who stay in their assigned neighborhood school? CREDO (2012) reports that students enrolled in Indiana charter schools outperform their "virtual twins" in traditional public school by about 0.04 standard deviations on the ISTEP+. Imagining that test scores of students who stayed in traditional public school were unchanged by school choice, test scores of the 10 percent of students enrolled in school choice in 2015 would have to have improved by 0.26 standard deviations to account for the 0.026 shift in all students' average test score. This magnitude of gains by choice participants far exceeds what CREDO (2012) estimates for charter students, indicating that test scores of students in traditional public schools probably also improved relative to their counterfactual outcome without school choice.

1.6 Robustness Checks and Extensions

1.6.1 Pretrends

A way to check whether the observed effect on test scores could plausibly be caused by the introduction of school choice is to determine whether the districts in question already

[19]The magnitude of a standard deviation change in school choice enrollment is slightly different when the districts are weighted by population (6.0 percent) than when they are unweighted (5.5 percent).

had positive test score trends prior to getting school choice.[20] The worry is that, if school choice took off most strongly in districts that were already improving, we would erroneously attribute their improvement to the effect of school choice. To operationalize this check, compare each district's movement in test scores during the years before school choice was introduced (2001–2005) to the increase in school choice enrollment in that district during the years of this study (2005-2015),

$$Y_d^{05-01} = \alpha + \beta C_d^{15-05} + \epsilon_d.^{21}$$

The resulting estimate of β (-1.422 with standard error 1.512) is negative but statistically insignificant. Thus, no correlation between school choice enrollment and pre-period score growth is apparent. The negative point estimate indicates that if there is a relationship it is probably negative, which would lead the coefficients in Table 1 to be underestimates.

1.6.2 Does School Choice Alter Student Composition?

A key identifying assumption of this research is that the pool of students taking the ISTEP in each school district does not change as a result of the introduction of school choice. This assumption can be violated in a few ways, allowing"composition effects" of school choice. First, the pool of ISTEP takers in a district can be changed if non-public school students—either private or home school students—start taking the ISTEP as a result of school choice. Second, the pool of ISTEP takers in a district can change if students migrate across district lines as a result of school choice. We consider each of these scenarios below.

[20]The change in each district's TPS average ISTEP+ score from 2001 to 2005 is used as the pretrend. This works during 2001–2005 since nearly all publicly-funded students still attended their district's TPS at that time.

[21]Some charter schools operated before 2005 in Indiana, but these were few and limited to a few school districts.

First, what if some district-resident students start or stop taking the ISTEP as a result of school choice? The main concern here is that students might obtain a voucher who would enroll in private school regardless.[22] Closer inspection of voucher recipients' enrollment records reveals that this is indeed a problem. Around half of voucher recipients, or 15,000 students, were enrolled in private school before receiving a voucher for the first time.[23] This problem can be addressed by simply excluding ISTEP scores of students who have never attended a public school. Removing scores from these 6,700 students (of the over 1.3 million students in the sample) changes the point estimates very slightly and does not affect signs or statistical significance levels.[24]

The second channel for "composition effects" to occur is if students migrate across school district boundaries in response to the introduction of school choice. Whether this happens is part of a larger question of Tiebout sorting, whether people "vote with their feet" in response to the quality of local public services (Tiebout 1956, Oates 1969). Papers examining this question find that students do indeed migrate if school choice changes the quality of schooling accessible to residents of a school district (Nechyba 2000, Brunner et al. 2012, Clarke 2016). To illustrate, if a high quality school corporation starts accepting non-resident students because of inter-district open enrollment, that affects the schooling options available to students in nearby districts. Families who would otherwise not want to live in an adjacent school district with poor quality schools might be convinced to move there, planning to enroll as transfer students in the nearby high quality school corporation. Similarly, families might consider moving to take advantage of a new charter school or if they will be able to afford a private school with the help of a voucher.

[22] Another concerning scenario is if students who would otherwise be home schooled enroll in a virtual charter school, but virtual charter school enrollment patterns suggest this is uncommon.

[23] The main way that students already in private school obtain a voucher is by first obtaining a scholarship through a related program with looser eligibility requirements. This is the School Scholarship Tax Credit Program. Students who obtain such a scholarship in one year are eligible for a voucher in the next year if they meet income limits, regardless of prior school attendance.

[24] The estimated effects when these students are removed are 0.422 (0.120) for the weighted specification with controls corresponding to Column 4 of Table 1.

An examination of Indiana transfer students' enrollment and residence patterns sheds light on this process. Of 69,000 first-time transfer students, 50,000 were observed in the year before transferring. 40 percent of those students changed their district of residence but stayed enrolled in the same school corporation by transferring "in reverse." Another 25 percent changed both their district of residence and corporation of enrollment. Only 35 percent of transfer students observed in both years continued living in the same school district but changed the school corporation where they enrolled. While we can see high rates of moving among transfer students taking up school choice, it is difficult to know what this means since we cannot observe the counterfactual. If there were no school choice would these students have moved anyway?

One way to assess how strongly student mobility responses to school choice might affect the regression results is to exclude all students who ever enroll in school choice from the regression sample, under the assumption that only students who take up school choice would move residence because of school choice.[25] While doing this solves the problem of composition effects arising from mobility, it introduces a separate composition effect of students selectively leaving traditional public schools. However, if we think selection into school choice is mostly positive (i.e., better students from within a school opt for school choice), then bias from composition effects would cause an underestimate of the effects of school choice. In fact the result of such a regression, using the preferred specification from Column 4 of Table 1, is a positive significant coefficient on the share of students enrolled in school choice in a district of 0.338 with standard error of 0.131. This suggests that residential mobility is not driving the positive association between school choice and test score improvement.

[25] This approach discounts the possibility of "general equilibrium" effects of school choice on neighborhood composition under which families move who are not interested in school choice.

1.6.3 Is Intra-District Open Enrollment an Important Omitted Variable?

Intra-district open enrollment is a separate type of school choice not addressed in this paper. This is when students are allowed to choose their school within the school corporation of the district where they reside. Referring back to the motivating framework described in Section 1.4.2, the fundamental difference between intra-district open enrollment and the types of choice discussed in the main part of this paper is that it does not cause students to switch between school corporations. The event of students open enrolling within their own district could improve student performance by improving the match quality between student and school, but it is unlikely to provoke a competitive response from school corporations since no students (or funding) are gained or lost.

Although several Indiana school districts have a policy to allow intradistrict open enrollment, how frequently students take up this choice is not directly observable in the state's data. However, one can construct a measure to proxy for intra-district open enrollment and check whether excluding this variable affects the estimates presented so far. Specifically, a measure is constructed of the overlap between students' school-grade peers from one year to the next. Students who have less than 30 percent of the same peers as last year but have not switched corporation are considered to have transferred to another school within the school corporation. Adding an index of this proxy for intra-district open enrollment to the regressions in Table 1 leaves the main coefficient for percent choice enrollment—from charters, vouchers, and *inter*district open enrollment—virtually unchanged. Estimated coefficients for the effect of intra-district open enrollment are small and statistically insignificant. These results are shown in Table 2.

1.6.4 Results Broken Out by Test Subject, Grade, and Type of Choice

Section 1.4.4 noted that the estimates are averaged over test subject (math and ELA), grade (3–8), and type of choice (charter, voucher, transfer, virtual). Tables 3–5 present estimates disaggregated on each of these bases, to ensure that major heterogeneity is not being overlooked. Table 6 additionally disaggregates students who stay in the local TPS and those that enroll in school choice. For brevity, only the specification using *% Choice* with enrollment weights and district-level controls is included—analogous to row 1, column 4 in Table 1.

Table 3 shows that the estimated effects of school choice are statistically significant and of similar size for both math and ELA when these are modeled separately.

Table 4 shows that the effects of different types of school choice may vary though only the estimated effects of charter schools and vouchers have statistical significance. The coefficients for vouchers and virtual schools are notably higher than for charters and transfers. When the same regression is run after removing test scores of voucher recipients who have never attended a public school, the coefficient for vouchers *grows* slightly to 0.928 (0.384) while the other coefficients are virtually unchanged. The coefficient for virtual schools, though only marginally significant, hints at a possible puzzle that could warrant further research since other researchers have found negative effects of virtual schools.[26]

[26] Virtual schools in many states including Indiana have been criticized for having low academic quality and negatively affecting students (e.g., Woodworth et al. (2015)). However, virtual schools are a type of school choice that is available even when the other forms of choice studied here are not locally accessible. If school choice exhibits decreasing marginal returns, the first bit of choice take-up in a district could elicit a potent competitive response from the traditional public schools. Thus, switching to a virtual charter school could hurt the students who switch while helping their peers at the TPS, potentially explaining the highly positive coefficient on virtual enrollment in Table 4.

Table 5 separately estimates the effects of school choice for students in each grade. The estimated effect is positive for all grades and statistically significant for most.[27]

1.7 Conclusion

This paper examines how average test scores of *all* students in a community change when *some* students enroll in school choice. The types of school choice considered are charter schools, vouchers, and inter-district open enrollment. The main takeaway is that average test scores of all students and local levels of participation in school choice are positively linked. This is shown simply with a scatter plot (Figure 1.6) and confirmed by a series of fixed effects regressions. Altogether, this paper provides suggestive evidence that the introduction of school choice caused net improvement to Indiana students' test scores from 2005 to 2015. Any alternative explanation for this growth in test scores would have to be correlated with the introduction of school choice, not accounted for by demographic control variables, and not present in linear test score trends before the introduction of school choice.

The estimated effect is a modest 0.02–0.03 standard deviation increase in average ISTEP+ scores for students living in a district that experiences a one standard deviation increase in school choice. While students who enroll in choice schools improve the most, students who remain in their traditional public school also benefit. Anectdotally, it is interesting to note that, of the 10 districts in the state with the largest growth in choice enrollment, only 1 experienced a meaningful drop in its test scores relative to the rest of the state during 2005–2015. During 2001–2005, test scores in 7 of these same districts fell relative to the rest of the state.

[27] One might expect to see bigger effects for older students if the effects of school choice accumulate over time, but this pattern is not found here. However, the data used here are not well-suited to investigate this question since school choice has only been present in many Indiana districts for a short period of time. That is, if school choice has only been present for three years in a district, the 3rd graders and 8th graders have received equal levels of "exposure" so there is no cause to expect unequal effects of school choice.

As a final note, the main facts established in this paper are observable using average test scores. Nonetheless, these averages are not publicly available: Indiana and other states publish test scores aggregated by school enrollment but not by residence, even though the information needed to do so is already collected. This becomes a problem when substantial portions of a community start enrolling in school choice, leaving local parents and officials no summary measure of educational outcomes in their community. For example, Indianapolis Public Schools' test scores are a poor measure of test score trends for all Indianapolis children because a third of those students enroll outside of Indianapolis Public Schools. States could help residents of districts with high choice take-up and make the effects of school choice more transparent by publishing test scores aggregated by district of residence.

Table 1.1: Coefficient of Interest for Main Difference in Differences Models

	District-Year Average Test Score			
	Unweighted		Enrollment Weighted	
	No Controls	Controls	No Controls	Controls
	(1)	(2)	(3)	(4)
% Choice				
	0.244**	0.321***	0.271*	0.443***
	(0.116)	(0.117)	(0.149)	(0.117)
N	3,178		3,178	
Mean	0.024		0.034	
S.D	0.043		0.062	

Notes: The dependent variable is district-year average test score. The independent variable of interest (coefficients shown) is the share of district resident ISTEP+ takers enrolled in any type of school choice that school year. (Charter and private schools are treated as independent school corporations.) Control variables (not shown) are the percentage of students in a district designated as black, Hispanic, white, free or reduced price lunch, limited English proficiency levels 1-4, and male. There are 289 school districts and 11 years. Robust standard errors are listed in parentheses with significance levels designated by * $p < 0.10$, ** $p < 0.05$, and *** $p < 0.01$.

Table 1.2: Controlling for Intra-District Mobility Measure

	District-Year Average Test Score			
	Unweighted		Enrollment Weighted	
	No Controls	Controls	No Controls	Controls
	(1)	(2)	(3)	(4)
% Choice				
	0.239**	0.332***	0.270*	0.445***
	(0.128)	(0.129)	(0.152)	(0.119)
Proxy for Intra-District Open Enrollment				
	−0.127	−0.034	−0.065	0.054
	(0.128)	(0.127)	(0.150)	(0.132)
N	3,178		3,178	

Notes: The dependent variable is district-year average test score. The independent variables of interest (coefficients shown) are the share of district resident ISTEP+ takers enrolled each year in school choice (meaning charter schools, private schools with a voucher, or as a transfer student outside the district where they reside) and a district-year index of participation in intra-district open enrollment. (Charter and private schools are treated as independent school corporations.) Control variables (not shown) are the percentage of students in a district designated as black, Hispanic, white, free or reduced price lunch, limited English proficiency levels 1-4, and male. There are 289 school districts and 11 years. Robust standard errors are listed in parentheses with significance levels designated by * $p < 0.10$, ** $p < 0.05$, and *** $p < 0.01$.

Table 1.3: Separate Regressions for Each Test Subject

	District-Year Average Test Score (Enrollment Weighted with Controls)	
	Math	ELA
	(1)	(2)
% Choice	0.437***	0.384***
	(0.119)	(0.107)
N	3,178	

Notes: The dependent variables are district-year average ISTEP+ math score in column 1 and English language arts (ELA) in column 2. The independent variable of interest (coefficients shown) is the share of ISTEP+ takers residing in a district who enrolled in any type of school choice that school year. Control variables (not shown) are the percentage of students in a district designated as black, Hispanic, white, free or reduced price lunch, limited English proficiency levels 1-4, and male. There are 289 school districts and 11 years from 2005-2015. Robust standard errors are listed in parentheses with significance levels designated by * $p < 0.10$, ** $p < 0.05$, and *** $p < 0.01$.

Table 1.4: Separately Estimated Effects for Each Type of School Choice

	District-Year Average Test Score (Enrollment Weighted with Controls)
	(1)
% Charter Enrl	0.205
	(0.121)
% Voucher Enrl	1.031***
	(0.253)
% Transfer Enrl	0.319
	(0.288)
% Virtual Enrl	2.206*
	(1.162)
N	3,178

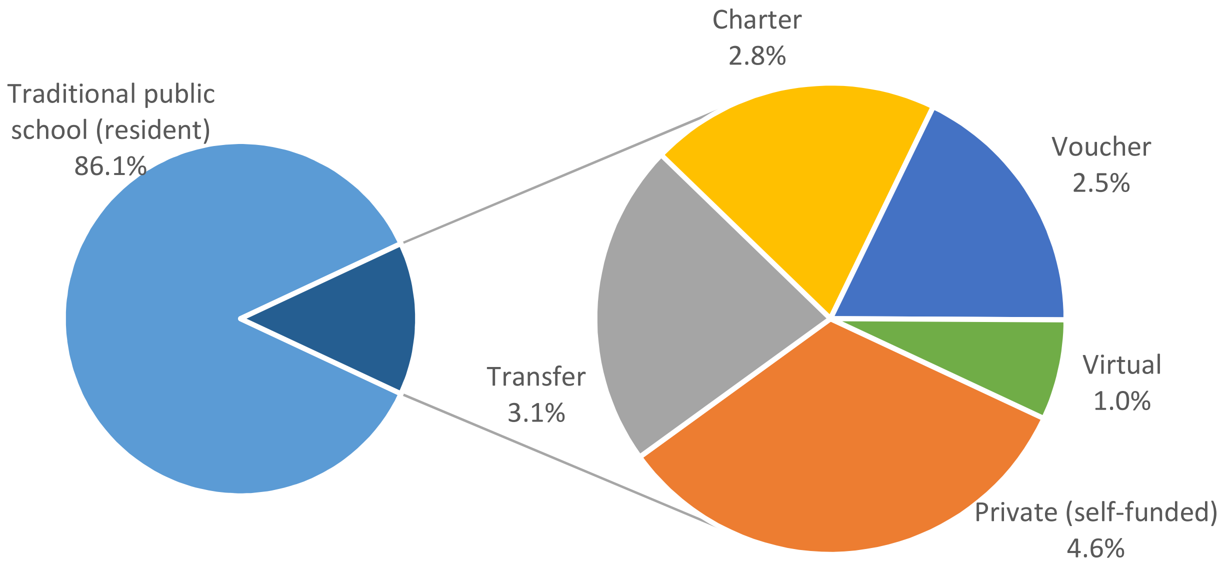
Notes: The dependent variable is district-year average test score. The independent variables of interest (coefficients shown) are the share of ISTEP+ takers residing in a district who enrolled in a charter school, a private school using a voucher, a traditional public school as a transfer student, or a virtual charter school. Control variables (not shown) are the percentage of students in a district designated as black, Hispanic, white, free or reduced price lunch, limited English proficiency levels 1-4, and male. There are 289 school districts and 11 years from 2005-2015. Robust standard errors are listed in parentheses with significance levels designated by * $p < 0.10$, ** $p < 0.05$, and *** $p < 0.01$.

Table 1.5: Separate Regressions for Each Grade Level

	District-Year Average Test Score (Enrollment Weighted with Controls)					
	G3	G4	G5	G6	G7	G8
	(1)	(2)	(3)	(4)	(5)	(6)
% Choice	0.175	0.457***	0.423**	0.534***	0.457***	0.549***
	(0.240)	(0.177)	(0.195)	(0.137)	(0.155)	(0.160)
N	3,178					

Notes: The dependent variable of each column is district-year average test score for students in the indicated grade. The independent variable of interest (coefficients shown) is the share ISTEP+ takers in the indicated grade residing in a district who enrolled in any type of school choice that school year. Control variables (not shown) are the percentage of students in a district designated as black, Hispanic, white, free or reduced price lunch, limited English proficiency levels 1-4, and male. There are 289 school districts and 11 years from 2005-2015. Robust standard errors are listed in parentheses with significance levels designated by * $p < 0.10$, ** $p < 0.05$, and *** $p < 0.01$.

Figure 1.1: Growth of Enrollment in Indiana's School Choice Options, 2005–2015



Note: The proportion of students not appearing in IDOE data is determined by a comparison to ACS 5-year estimates.

Figure 1.2: Growth of Enrollment in Indiana's School Choice Options, 2005–2015

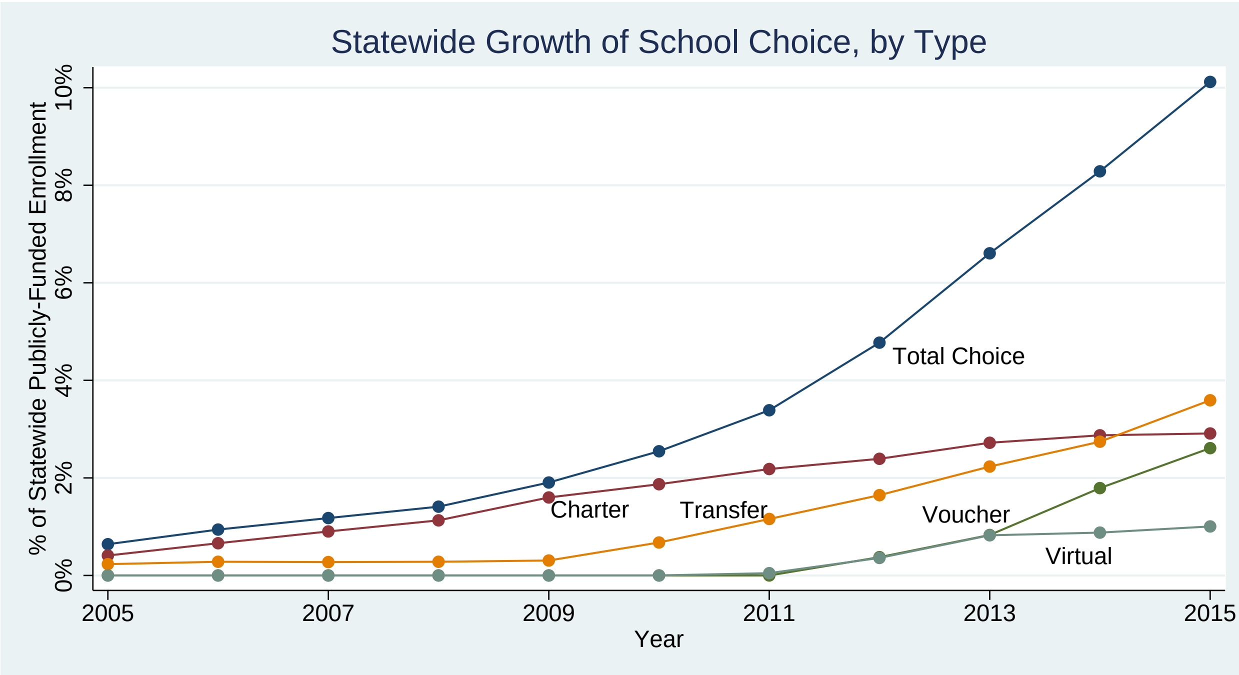
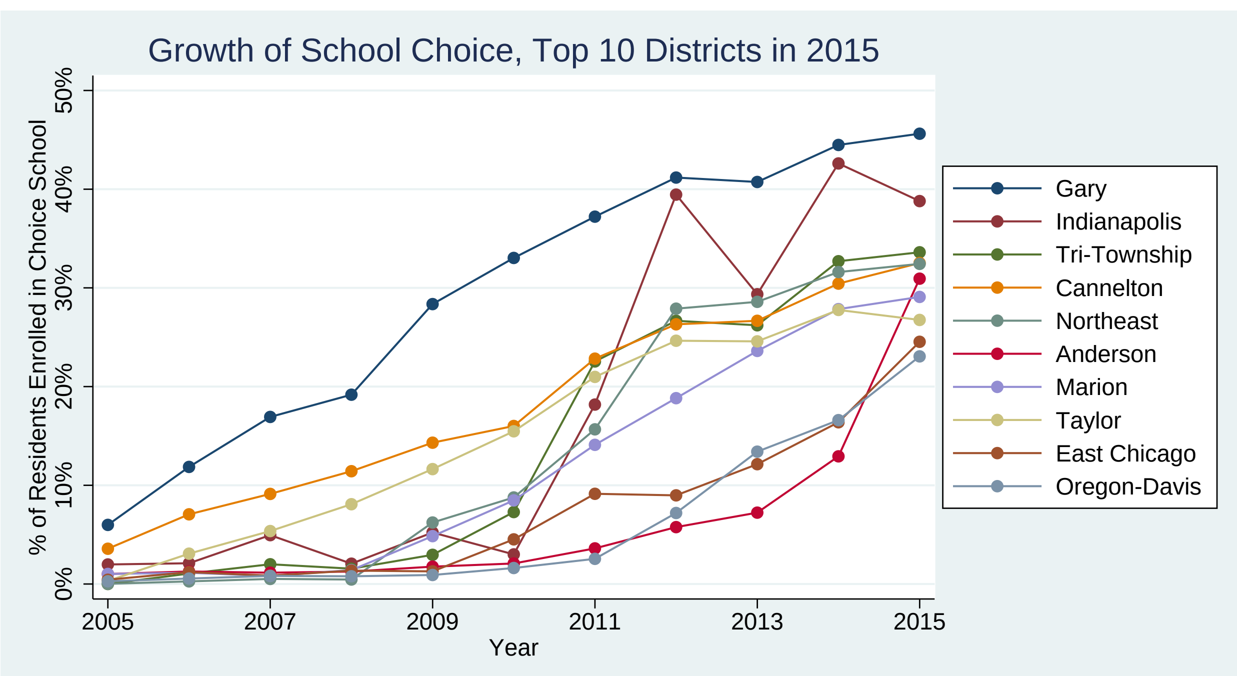
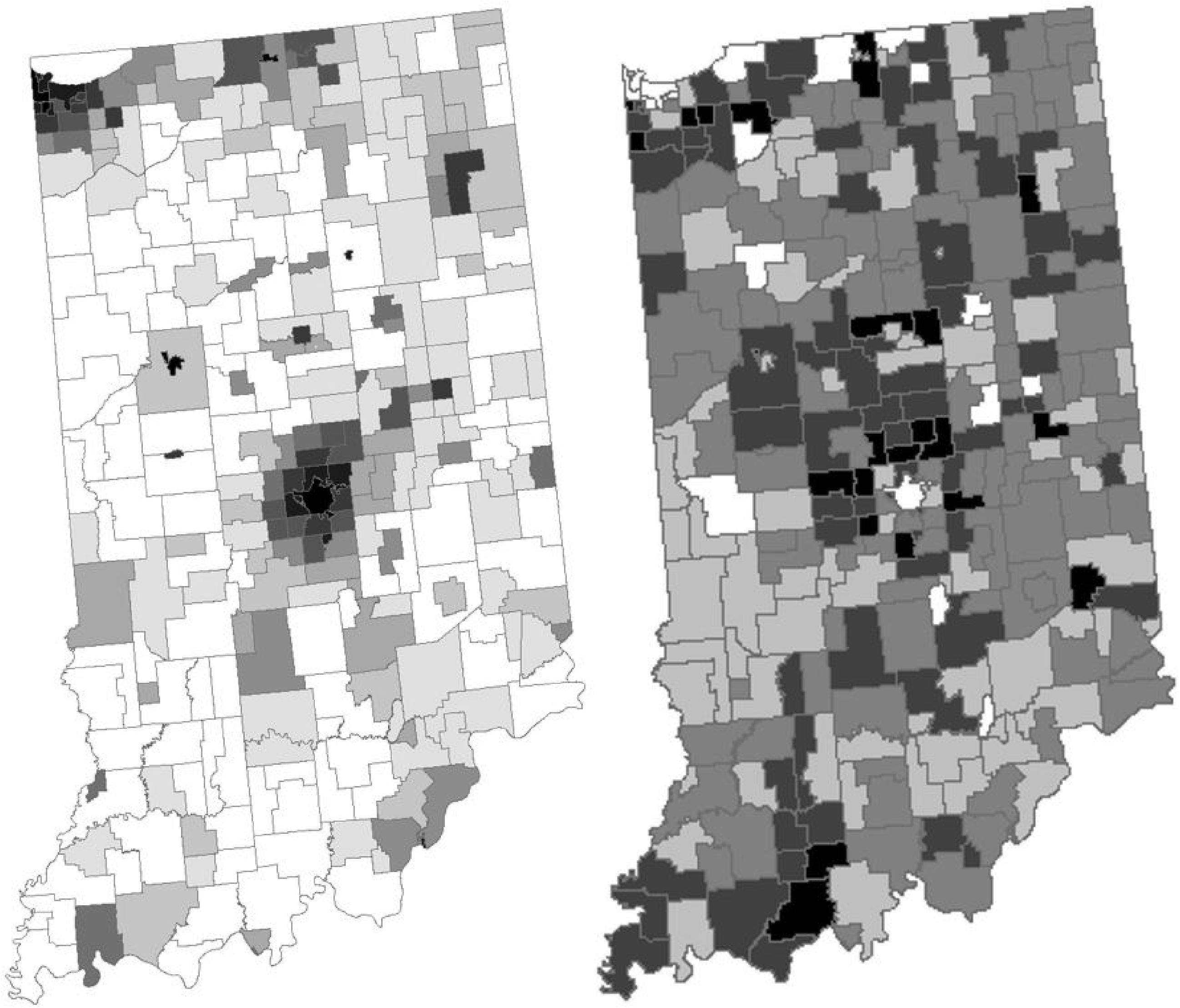


Figure 1.3: School Districts with the Most Choice Take-Up



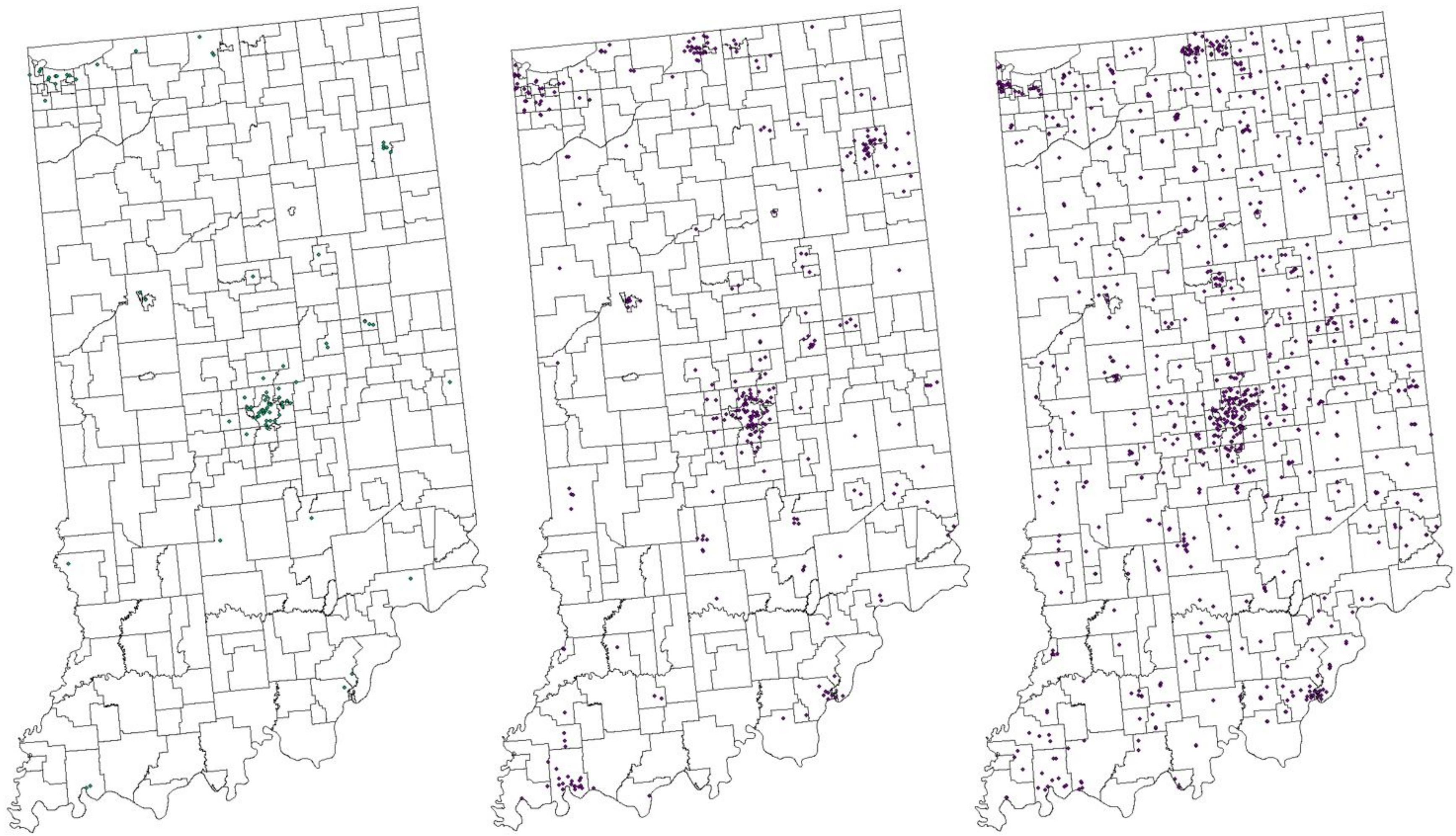
Note: These are the 10 districts with the highest proportion of students enrolled in choice schools in 2015.

Figure 1.4: Population Density and District-Average Test Scores



Note: In the first panel, darker districts signify higher population density. In the second panel, darker districts signify higher average test scores.

Figure 1.5: Geographic Distribution of School Choice Options



Left: charter schools, center: private schools which accept vouchers, right: traditional public schools which enroll transfer students.

Figure 1.6: Scatter Plot of Growth in School Choice vs. Growth in Test Scores

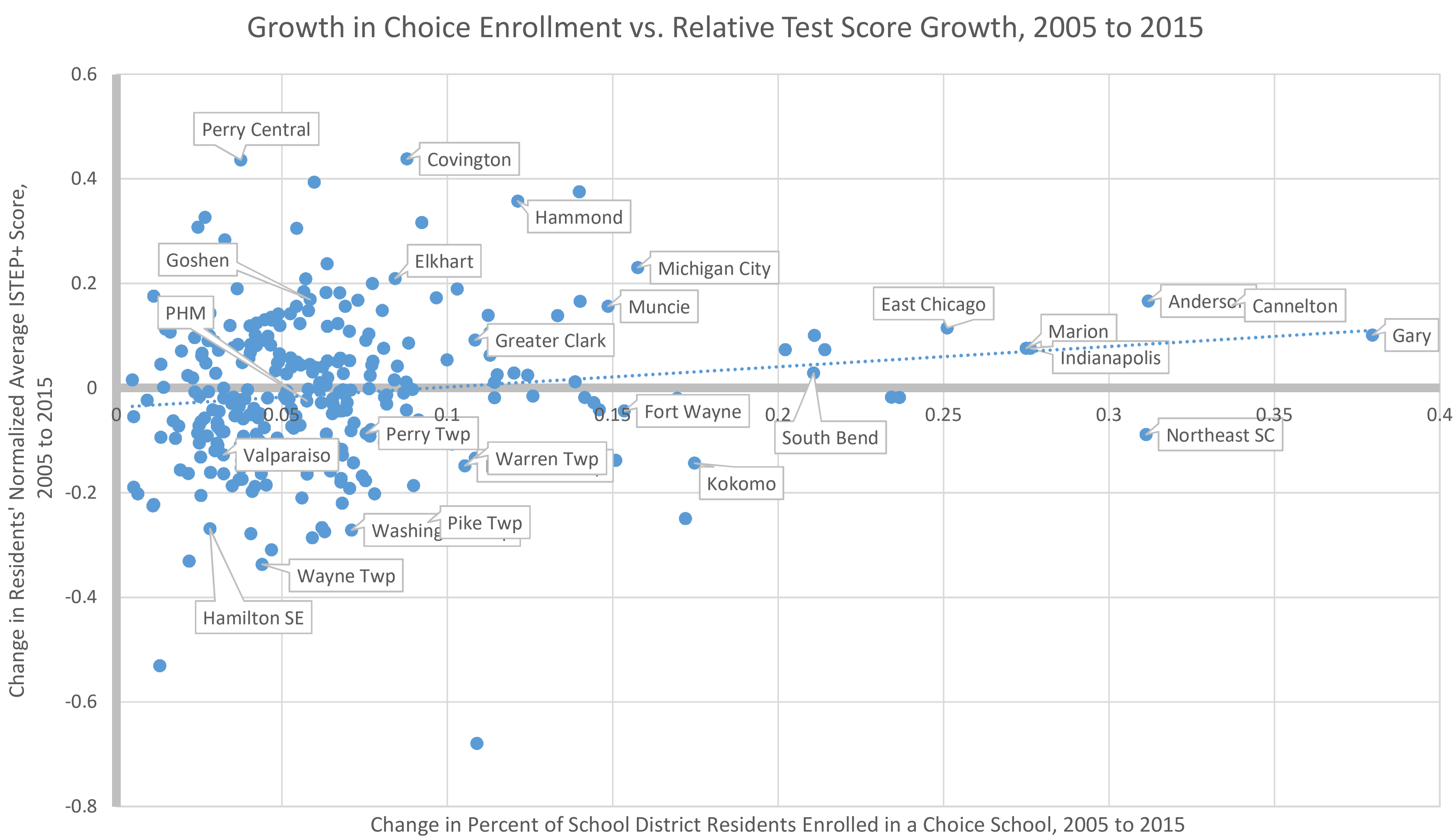
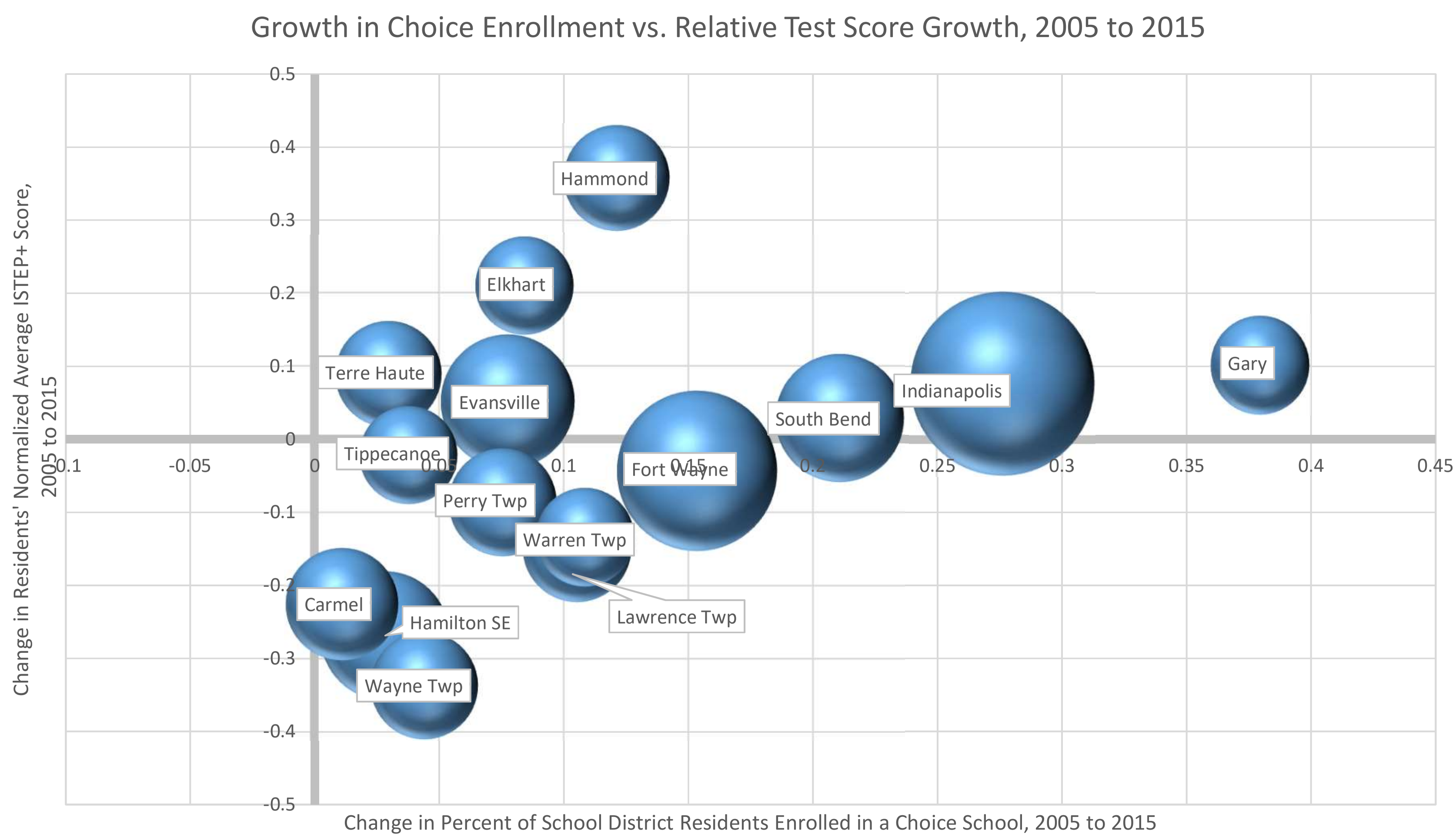


Figure 1.7: Scatter Plot of Growth in School Choice vs. Growth in Test Scores, Weighted by Enrollment



Note: Same as Figure 1.6, but weighted by enrollment and displaying only the 15 most populous districts.

Chapter 2

Optional Public School Inter-District Open Enrollment and House Prices

2.1 Introduction

Eligibility to attend public primary and secondary schools in the United States has traditionally been determined by the location of a family's residence. By buying or renting housing in a particular geographically-defined school district, one receives the right to enroll children in the local public school. Economists have long debated whether and how much the quality of local schools is *capitalized* into the price of housing in such a context.

Over the past several decades, the importance of this right has been altered by a variety of "school choice" initiatives, for which enrollment eligibility is not determined by residency. These initiatives include charter schools, private school vouchers, magnet schools, intra-district open enrollment, and—the subject of this paper—inter-district open enrollment. The idea of inter-district open enrollment (henceforth "IOE") is to allow students to attend school in a district where they do not live, giving families more choices and motivating public schools to improve by fomenting competition (Barseghyan et al. 2014). School districts are reimbursed for educating the additional students. Thirteen states have a general requirement that school districts accept non-resident applicants if the district has capacity. In another 13 states, school districts are allowed but not required to accept non-resident students (Education Commission of the States, 2015). (See Figure 1.)

Economic literature to this point has found that mandatory IOE policies decrease the value of houses in school districts with relatively good public schools. The equally-common situation in which school districts can choose whether to accept non-resident students has received less attention. This paper extends the discussion of how IOE affects property values by considering the decision-making of school districts under optional IOE. Indiana is used as a test case.

Section 2.2 discusses existing findings about the effects of mandatory IOE on property values. Section 2.3 explains the policy setting in Indiana and tests for the same relationship in Indiana, with optional IOE. No relationship is detected. Section 2.4 considers how schools' endogenous selection into accepting transfer students might mute effects on house prices that are found with mandatory IOE and presents empirical evidence that corroborates this story in Indiana. Section 2.5 discusses the empirical evidence, considers policy implications, and concludes.

2.2 Literature on Capitalization Effects of Mandatory Inter-District Open Enrollment

The premise of economic models which predict capitalization—a connection between locally provided public goods like schooling and house prices—is that limited availability of housing effectively rations access to those public goods. As a result, the richest agents afford living in areas with the best schools or other amenities. If the quality of amenities in a location changes, a corresponding migration response occurs: Improved schools attract richer individuals to move in and poorer individuals can be priced out of the market. Worsened schools repel richer residents and the resulting lower house prices attract poorer residents. This idea was introduced by Tiebout (1956) and Oates (1969). Hilber (2015) provides a recent overview of the ensuing literature.

When open enrollment is introduced in such a model, the schooling options in a poor neighborhood improve since residents now have the chance to attend a better school. Schooling options in a rich neighborhood, on the other hand, do not change since residents already attend the best available school. Under the simplifying assumptions that all residents value the relevant schools equally and that transportation costs are nil, all residents of a poor neighborhood will apply to attend school in a rich neighborhood. Some random portion $\alpha \in (0, 1]$ of non-resident applicants are accepted. This chance α of attending a rich school while paying poor neighborhood housing prices convinces some families to move from a rich to a poor neighborhood. The changes in demand for housing implied by the previous sentence cause house prices to fall in rich neighborhoods and rise in poor neighborhoods, until prices reach a new equilibrium. Formal models of how the introduction of school choice can affect housing values are described and estimated by Nechyba (2000, 2003*a*,*b*), Epple & Romano (2003), Ferreyra (2007); a concise, motivating version applied to inter-district open enrollment is presented in Section 2.4 of Reback (2005).

Reback (2005) empirically tests whether housing prices responded as described above upon the introduction of open enrollment in Minnesota, with a difference-in-difference strategy. Using ordinary least squares, he shows that house prices fell in districts which were net receivers of transfer students and rose in districts that were net senders of transfer students. Lest we suspect this result is anomalous, Brunner et al. (2012) conduct a similar study using data from 12 states that implemented mandatory inter-district open enrollment programs.

Their findings regarding house prices are qualitatively the same as those of Reback (2005), though they appeal to instrumental variables to eliminate supposed upward bias caused by families wanting to transfer into school districts where prices are already increasing.[1]

2.3 Optional Inter-District Open Enrollment in Indiana

Consider now the case of Indiana. The goal of this section is to determine whether the shift in house price capitalization detected in Minnesota and 11 other states upon the introduction of mandatory IOE also occurred when Indiana implemented optional IOE. This study is performed using confidential student-level enrollment and residency information from 2008 to 2015 provided by the Indiana Department of Education and publicly-available parcel-level real estate data for the same period provided by the Indiana Department of Local Government Finance.

2.3.1 Description of Optional IOE in Indiana

Optional inter-district open enrollment started in Indiana in 2010 as part of a larger overhaul of the K-12 school funding system.[2] The state government took over full responsibility for schools' "General Fund," which represents the majority of all K-12 education spending for traditional public schools. This shift allowed a "funding follows the child" system under which schools are paid on the basis of the number of students enrolled. Inter-district

[1] Chung (2015) obtains similar results using data from South Korea.

[2] The basic organization of public schools in Indiana mirrors much of the rest of the country. The state is divided into 289 geographic *school districts*. Each district has a local government called a *school corporation*, with an elected board and power to levy property taxes, which is responsible to educate students who live within its geographic boundaries. Students that live within a school district are typically assigned by the school corporation to a "school of legal settlement" based on their residence, which is traditionally where students have attended.

open enrollment was then implemented simply by the state agreeing to reimburse school corporations (Indiana's name for local education authorities) the same amount for resident students, who the school corporations are *required* to accept, and non-resident students, who corporations may *choose* to accept.[3]

The law was amended in 2013 to bar school corporations from "cherry picking" (selectively admitting) non-resident students after complaints of that happening. If corporations choose to participate at all, they must accept any non-resident student up to a limit set by the school board each year for each school and grade. If more students apply than will be accepted, the school board has to choose students randomly via a publicly-held lottery. Priority is given to non-resident students already accepted in prior years, their siblings, and non-resident children of school corporation employees. Unfortunately, these details of the selection mechanism are not reflected in the data used for this study, causing two problems. First, while one can observe which students transfer, a researcher does not know if students were denied admission in a lottery or if the school corporation was willing to accept more transfer students than applied.[4] Second, one does not know which transfer students were given priority for acceptance because they are siblings of transfer students or children of corporation employees. This hidden distinction obscures whether school districts are accepting non-priority transfer students in a given school year.[5] As such, this paper will

[3]Public schools were previously allowed to enroll Indiana-resident students from outside their districts without approval of the sending district, but the school would not be reimbursed. Instead, the school could charge tuition and apply acceptance criteria for non-resident students similar to what a private school might do. Comparatively few students transferred this way, and it should not have impacted property values because tution costs would eliminate any opportunity to save money by living outside an expensive community while accessing its schools. The law also allowed school districts to make bilateral agreements under which one district pays the other to provide services more appropriate for specific children; this arrangement remains in force, and is mostly used to provide special education or vocational training.

Under the new system, transfer students have to provide their own transportation and may still be required to pay partial tuition. This payment covers expenses outside the General Fund—facilities, busing, and debt service—which are still funded via by local property taxes. Anecdotally, many school corporations forgo this partial tuition payment to attract more transfer students.

[4]Transfers are directly observed using an "Enrollment Type" field filled by schools and indirectly observed by noting which students enroll in a different school district from the one in which they reside.

[5]This is an opportunity to note three suggestions for state education officials: (1) Publish corporation-level totals of the number of students transferring into and out of school districts to improve transparency. (2) Expand, clarify, and enforce the requirement for corporations to report the number of students they are

treat school districts as resisting transfer students if their transfer enrollment, expressed as a percentage of total enrollment, is in the bottom quartile of school districts; that is if less than 1.1 percent of enrollment is composed of transfer students. This classifier assumes that even "undesirable" school districts can attract some transfer students if they wish. Figure 2 lends credibility to this notion. Of school districts where average test scores are worse than a weighted average of their neighboring districts' scores, 27 percent are in the bottom quartile of transfer enrollment. 23 percent of districts with better test scores than their neighbors are in the bottom quartile of transfer enrollment. One third of all transfer students attend a school district with worse average test scores than the district where they live.

The number of students who open enroll has risen quickly since 2010, reaching 36,000 students or about 3.2% of statewide enrollment in the 2014-2015 school year. The share of enrollment made up by non-resident transfer students varies widely across districts, ranging from 0% transfer students in 22 districts to 35% transfer students in Daleville Community Schools in 2014-2015.[6] Because the mechanism for open enrollment to affect property values in theoretical models is the threat of families changing residence (rather than staying put and changing their school enrollment), it is interesting to classify Indiana transfer students along those lines. Of the 69,000 unique students who have transferred for the first time since the start of the program, 50,000 were observed in the year before transferring. Of those, 35% continued living in the same school district but changed the school corporation where they enrolled upon transferring, and 40% stayed enrolled in the same corporation but changed the district where they reside upon transferring. (The remaining 25% changed both their district

willing to accept in each school and grade. Some corporations do not seem to meet the minimum reporting requirements, and centralized reporting would improve the state's ability to evaluate how widely available inter-district open enrollment is. (3) When school districts report that a student is a transfer student, also include whether the student was given priority as a sibling or child of an employee. These changes will clarify acceptance rates for applicants without special priority.

[6]The focus of this paper is school corporations' decisions, and school corporations cannot directly control whether a student transfers out of the corporation. That is why more emphasis is placed on the number of students transferring "into" a school corporations than the number transferring "out."

of residence and corporation of enrollment.)[7] This breakdown is consistent with theoretical models that predict open enrollment will cause a re-sorting of families into neighborhoods.

2.3.2 Capitalization Effects of Optional IOE in Indiana

As a first step in analyzing Indiana's optional open enrollment program, this section replicates as nearly as possible the main analyses of house prices presented by Reback (2005) and Brunner et al. (2012) using data from Indiana. That is, pre- and post-open enrollment housing prices are compared in districts which received varying levels of transfer students. This analysis does not find the inverse relationship between receipt of transfer students and movement of house prices described by the earlier authors who studied mandatory programs.

Panel 1 of Table 1 presents these results from Indiana in a variety of specifications. The first column presents the basic OLS regression including 92,000 arms-length housing transactions from 2010 and 2015:[8]

$$\text{HousePrice}_{idt} = \alpha_0 + \alpha_1 \%\text{NetTransfers}_{dt} + \alpha_2 \text{Year}_t + \text{HouseChars}_{id}\Gamma + \text{FE}_d\Theta + \epsilon_{id} \quad (2.1)$$

where i, d, and t signify the individual real estate transaction, school district, and year respectively; HousePrice_{idt} is the house's sale price; $\%\text{NetTransfers}_{dt}$ is the number of incoming transfers students minus the number of outgoing transfer students divided by the total number of students residing in the district in each year; and HouseChars_{id} is a vector

[7] The number of transfer students who change enrollment is mostly uncorrelated with the number of moves overall, whereas the number of transfer students who change residence has nearly a 1-to-1 correlation with the number of moves overall.

[8] This number represents about 40 percent of all transactions statewide in those years. County assessors indicate whether a housing transaction is a valid arms-length transaction for use in trending of property tax assessments. Transactions with reported sale price below \$1,000 and above \$1,000,000 are also eliminated as outliers. Similar results are obtained using all real estate transactions.

of characteristics of the transacted houses.[9] The estimated coefficient on $\%\text{NetTransfers}_{dt}$ is statistically insignificant and positive—counter to the models described above which predict a negative coefficient. In his study of Minnesota, Reback's (2005) analogous regression obtained a negative, statistically significant coefficient, which served as his main result. In their multi-state study, however, Brunner et al. (2012) get a similar result to this when using OLS—insignificant with the "wrong" sign.

To explain this finding, Brunner et al. (2012) raise a concern of reverse causality—that students might be more drawn to transfer to schools in districts with upward trending house prices—leading to a positively biased estimate of α_1, the effect of receiving transfer students. Those authors devise two instruments for $\%\text{NetTransfers}_d$ (which the current paper also uses in a 2SLS model in Column 2), variables they do not believe to directly affect short-term trends in house prices and that predict the movement of transfer students:

1. $\text{RacialDifference}_d$, the difference between the percentage of students in district d who are designated as white and the percentage of students in surrounding districts who are designated as white, and

2. $N_d^{neighbors}$, the number of adjacent districts.

Using these instruments, Brunner et al. (2012) obtain the negative, significant result predicted by theory. In Indiana's case, the coefficient becomes negative but is still far from statistical significance, as seen in Table 1, Panel 1.

One may be concerned that something is different between the houses sold in Indiana before and after the introduction of open enrollment that is not captured by housing characteristics included in the model. To address this worry, Column 3 restricts the sample to repeat sales—houses sold once in 2009-2010 and again in 2014-2015. This group comprises about 12,000

[9]House characteristics include living area, age at sale, building frame type, assessor-assigned grades for original construction quality and current condition, number of bedrooms and bathrooms, garage and basement type, and whether the house has air conditioning and a fireplace.

houses or 24,000 transactions. The vector of housing characteristics in the regression is replaced by a fixed effect for each house. Column 4 presents results from the 2SLS regression with repeat sales and the two instrumental variables used earlier. As in the first two columns, the results of Columns 3 and 4 are statistically insignificant. All four specifications fail to provide evidence that the introduction of optional IOE affected capitalization of school quality into house prices in Indiana.

2.4 Selection Into IOE on the Basis of Heterogeneous Capitalization

2.4.1 Theoretical Discussion

Why might receiving transfer students hurt local house prices when it is mandatory but not when districts can choose whether to accept transfer students? One explanation is that districts selectively participate in a way that avoids or minimizes this effect. How would that process work?

The theoretical relationship between quality of schools and house prices posited in Section 2.3 relies on an assumption that the number of houses with access to a school is fixed (i.e., housing supply is perfectly inelastic). This is the mechanism by which access to schools is rationed, provoking a difference in house prices between school attendance zones of differing quality. If an unlimited amount of housing could be build at constant cost in any school attendance zone (i.e., if housing supply were perfectly elastic), nothing would stop residents of a poor neighborhood from moving into a "rich" neighborhood as long as the "rich" school were better.

Debate about the nature of housing supply elasticity traces back to the time of Tiebout and continues today. What seems clear is that housing supply elasticity generally falls somewhere between the two extremes of perfectly inelastic or perfectly elastic and that it varies from place to place (Hilber 2015, Brasington 2002). Research on spatial variations in housing supply elasticity points to factors such as urban density (Brasington 2002, Hilber 2010, Stadelmann & Billon 2012), land use regulations (Hilber et al. 2014), natural land features (Saiz 2010), and whether house prices are below replacement cost (Glaeser & Gyourko 2005) as determinants.

In a school district with relatively elastic housing supply—possibly a steadily growing suburb with ample farm land, low regulation, and a competitive construction industry—house prices should be held close to the cost of new construction. If prices of existing homes were to rise, home buyers would favor new houses over existing houses. Were prices of existing houses to fall below construction cost, the pace of construction would slow until existing home prices rose again. In a school district with relatively inelastic housing supply—due to physical, legal, or market constraints—the quantity response to a demand shock would be muted, "freeing" prices to stray from the price of new construction on the basis of factors like school quality.

This difference between school districts with elastic and inelastic housing supply matters in the implementation of open enrollment because the ensuing demand shock (either positive or negative, depending on the desirability of the district's schools relative to its neighbors's schools) will affect house prices in districts with inelastic housing supply more strongly. (See Figure 3.) To the extent that property owners who want to preserve their housing values have influence over elected school boards, school boards in places with relatively inelastic housing supply will be less inclined to accept transfer students.

2.4.2 Empirical Evidence

Empirical evaluation of the idea that housing supply elasticity affects districts' willingness to accept transfer students requires a way to measure housing supply elasticity. Direct estimation of this quantity requires a large enough sample of variation in school quality (or other local public service) to run a regression. This constraint results in a single estimate over a sizable geographic area, which is not ideal for capturing heterogeneity. More geographically precise measures can be obtained using proxies for housing supply elasticity that are directly observable, like available land for development or proximity to a city center. This section first uses a proxy measure of housing supply elasticity to test whether (1) house prices really react more strongly to inter-district open enrollment in locations with less elastic housing supply and whether (2) school districts enroll less transfer students when housing supply is less elastic. Finally, a direct measure of housing supply elasticity is developed and it is shown that (3) housing supply is less elastic in places that accept few transfer students.

The proxy measure used for housing supply elasticity is the share of developable land in a school district that has been developed (*% Developed*), following Hilber & Mayer (2009). The idea is that housing supply is less responsive to price changes when less land is available on which to build. This measure is constructed using USGS's National Land Cover Database 2006, matched to school districts with GIS software.[10] Direct measurement of housing supply elasticity is done by a hedonic regression, which is explained further below.

Are Housing Prices More Sensitive to Inter-District Open Enrollment in Densely Developed Locations?

The story of this paper is that school districts with inelastic housing supply resist accepting non-resident transfer students because they are afraid of damaging local house values. Most

[10] Land is considered undevelopable if designated as water, wetlands, or barren.

districts that do accept transfer students are those in which house prices are barely affected by the negative housing demand shock that accepting non-resident transfer students represents. For this reason, no statistically significant relationship between housing prices and receipt of transfer students is found for the average school district in Indiana.

The first step in testing this story empirically is to show whether the relatively few transfer students accepted by school districts with inelastic housing supply have a detectable, negative effect on housing prices. Doing so demonstrates whether school districts with inelastic housing supply really have "something to fear" from enrolling non-resident transfer students.[11] The regression from Equation 1—which failed to find a mean effect of transfer students on house prices—can be modified for this test by adding an interaction term that multiplies *% Net Transfers* and *% Developed*:

$$\begin{aligned}\text{HousePrice}_{idt} = \alpha_0 + \alpha_1\%\text{NetTransfers}_{dt} + \alpha_2\%\text{NetTransfers}_{dt} \cdot \%\ \text{Developed}_d \\ + \alpha_3\text{Year}_t + \text{HouseChars}_{id}\Gamma + \text{FE}_d\Theta + \epsilon_{id} \quad (2.2)\end{aligned}$$

Larger *% Net Transfers* means a more negative housing demand shock and larger *% Developed* proxies for less elastic housing supply. Thus, a negative coefficient on their product α_2 would indicate that transfer students bring housing prices down by more when housing supply is less elastic. (See Figure 3.)

Table 1, Panel 2 presents the results of this regression, where the four columns mirror those found in Panel 1, plus the interaction term. The OLS regressions in Columns 1 and 3 are very imprecise just as in Panel 1, perhaps reflecting districts' self selection into receiving transfer students. The IV regressions in Columns 2 and 4 have a negative coefficient on the interaction term as expected, but the coefficients are still fairly imprecise. They are

[11] In their studies of states with mandatory IOE, Reback (2005) and Brunner et al. (2012) find that housing prices respond to IOE more strongly in urban areas than rural areas, which presumably have more elastic housing supply.

"significant" at the 0.12 and 0.10 levels, respectively. Altogether, the second panel of Table 1 provides imprecise evidence that school districts with inelastic housing supply might really have "something to fear" from accepting transfer students.

What Factors Predict Transfer Enrollment?

Another empirical exercise to test the story told in Section 2.4.1—of school districts self-selecting into inter-district open enrollment on the basis of housing price elasticity—is to construct a simple model of how school districts' enrollment of transfer students is determined and check whether *% Developed*, the proxy measure for housing supply elasticity, is a significant determinant. To this end, Column 1 of Table 2 presents results of a regression of the share of each school district's 2014-2015 enrollment made up of non-resident transfer students on the following list of independent variables:

- *% Developed*: The share of developable land in a school district that is developed, serving as a proxy for degree of housing supply inelasticity.
- *% Transfers Sent*: The number of residents of a school district enrolled as transfer students in another district, normalized by the home district's enrollment.
- *% Max Residency*: Number of public school students residing in the school corporation as a percentage of the maximum residency in that district since 2005. This is intended as a measure of schools' capacity utilization.
- *Better Test Scores than Neighbors*: Difference between a district's average standardized test scores and average scores in neighboring districts, weighted by enrollment.
- *Higher Share White than Neighbors*: Difference between proportion of a district's students designated as white and proportion of white students in neighboring districts, again weighted by enrollment.

- *Number of Neighboring Districts*: The intuition is that having more nearby school districts increases the pool of non-resident students who might apply to enroll.
- *% Charter Enrollment* and *% Voucher Enrollment*: Percentage of public school students residing in the school district enrolled in charter schools, virtual charter schools, or private schools using a state-provided voucher.

The coefficient on *% Developed* in Column 1 of Table 2 indicates that having more development in a district (indicating less elastic housing supply) predicts lower transfer enrollment. A one standard deviation increase in *% Developed*—when 26 percentage points more of a school district's land is developed—lowers the predicted share of transfer students by 1 percentage point. For a district with the mean share of transfer students, that would be a one sixth decrease. Other significant predictors of transfer enrollment are *% Transfers Sent* and *% Max* Residency (indicating that school districts enroll more transfer students when their resident enrollment is below capacity), *Better Test Scores than Neighbors*, *Number of Neighboring Districts*, and *Higher Share White than Neighbors*.

A weakness of this exercise is that, even assuming the connection between transfer enrollment and *% Developed* is causal, it is unclear whether the relationship is driven by supply (by schools) or demand (by students) since these two factors are not distinguished in the available data. Fortunately, an asymmetry in the design of Indiana's optional IOE program can be used to help distinguish supply-driven factors from demand-driven factors: School districts are able to choose whether to accept transfer applicants from outside their district but not whether students residing inside their district can transfer out. If the link between *% Developed* and *% Transfers Received* is driven by students' demands, then we can expect to find the opposite relationship between *% Developed* and *% Transfers Sent*, since students influence the number of transfer students both sent and received in a district. However, if the connection between *% Developed* and *% Transfers Received* is driven by schools' supply of transfer spots, there is no reason to expect a relationship between *% Developed* and *%*

Transfers Sent, since schools control the number of transfer students received but not the number sent.[12]

Column 2 of Table 2 presents a regression with the same independent variables as found in Column 1, but using *% Transfers Sent* as the dependent variable (and adding *% Transfers Received* as an independent variable). The coefficient on *% Developed* in Column 2 of Table 2 is close to zero, insignificant, and has a small standard error, consistent with the story that schools, not students, are the ones deciding to limit enrollment of transfer students when housing supply is inelastic.[13]

An obvious threat to the line of analysis pursued in this subsection is omitted variable bias. There are likely numerous factors driving schools' enrollment of transfer students which are not considered here but that would affect the results. It is also unclear how well the degree of land development in a school district proxies for housing supply elasticity. Section 2.4.2.3 presents evidence to address these issues.

[12] This idea can be expressed in slightly more detail with the following illustrative model. Imagine that a student i picks the school s that maximizes his value function:

$$\max_{s \in S} V_{is} = Quality_s + Cost_{is},$$

and that school districts set their enrollment limits for transfer students using some unknown function F_s. If a regressor is linked to $Quality_s$ and does not enter schools' decision function F_s, we expect it to have opposite effects on the number of transfer students received and the number of transfer students sent. That is, for a student transferring from District A to District B, a factor *drawing* him to District B is a factor *driving* him away from District A. If a regressor is linked to $Cost_{is}$ and does not enter schools' decision function F_s, we might expect it to have the same effect in determining the number of transfers students received and sent. For example, if a school district covers a small land area, it is easier for students to both transfer into the school district and out of the school district. Finally, if a regressor does not enter the students' value function V_{is} but enters into school corporations' decision function F_s, we expect it to be predictive of transfers received but not transfers sent, following the logic that school districts can influence the former but not the latter. This informal reasoning does not account for spatial correlation among regressors or the possibility that a regressor is tied to multiple drivers of open enrollment, but it expresses the intuition used to compare Columns 1 and 2 of Table 2.

[13] In addition to the results for *% Developed*, the results in Column 2 are also consistent with *Better Test Scores than Neighbors* being a pull/push factor and *Number of Neighboring Districts* being a factor that brings down costs of transferring. The coefficients on *% Max Residency* could be interpreted as indicating it is a factor associated with costs of transferring or, more intuitively, that it affects students' desire to transfer and districts' desire to enroll transfer students in opposite directions. The coefficients on *Higher Share White then Neighbors* seem to indicate that this is a supply-driven factor and that school districts with more white students than their neighbors are happier to enroll transfer students. The intuition for this last result seems unclear.

Direct Estimation of Housing Supply Elasticity, Separately by Share Transfer Enrollment

A final way of assessing whether housing supply elasticity is associated with school districts' enrollment of transfer students is to directly estimate housing supply elasticity, separately for different groups of school districts. Doing so reveals that school districts which enroll few transfer students have less elastic housing supply than districts which enroll many transfer students.

The message of this exercise does not go beyond what was argued in Section 2.4.2.2—to suggest that school districts resist accepting transfer students when housing supply is inelastic—but it rests on a stronger footing since housing supply elasticity is measured directly. The proxy measure used in Sections 4.2.1 and 4.2.2, *% Developed*, only captures one determinant of housing supply elasticity, physical land availability. Other determinants like land use regulation (Hilber et al. 2014) and prices relative to replacement cost (Glaeser & Gyourko 2005) have so far been ignored.

The estimation procedure begins with dividing the state into districts in the bottom quartile of transfer enrollment, with less than 1.1 percent transfer students (indicating that the district has likely chosen not to participate in open enrollment), and school districts in the top three quartiles of transfer enrollment (indicating more enthusiasm about open enrollment). A hedonic regression is then conducted separately for each group of school districts to test whether school quality is reflected in house prices in that group of school districts. The groups of school districts are mapped in Figure 4. Table 3 shows the result of the hedonic regression:

$$\text{HousePrice}_{id} = \beta_0 + \beta_1 \text{TestScores}_d^{2005} + \beta_2 \text{Year}_t + \text{HouseChars}_{id}\Gamma + \epsilon_{id}$$

where the notation is the same as used in Section 2.3 and $\text{TestScores}_d^{2005}$ is the average standardized test score from 2005 of students enrolled in the neighborhood elementary school.[14] Column 1 shows the results from the whole state. Column 2 includes the bottom quartile of transfer-receiving districts, and Column 3 includes the top three quartiles of school districts for non-resident student enrollment. The results of Column 2 and 3 are strikingly different, with much stronger evidence that school quality is capitalized into house prices in school districts that receive fewer transfer students.

This "hedonic regression" is a relatively crude method of estimating the contribution of school quality to house prices; researchers worry about problems of omitted variables bias in such a regression (Black 1999). However, the intention here is not to produce a reliable estimate of the effect itself, but to contrast the size of estimates for school districts receiving few and many transfer students. Even if the "effect" of good schools detected here is partially attributable to other neighborhood amenities (since good schools are usually present in neighborhoods with many nice attributes), a positive coefficient in this regression still shows that local public goods are capitalized into house prices in those school districts.

Regression discontinuity studies, which look for sudden jumps in house prices at school attendance boundaries, are generally believed to provide better estimates of house price capitalization (Black 1999). A regression discontinuity study was not used in this case because that type of study requires a concentration of houses near each school attendance boundaries to detect any jump in house prices. Most rural school districts would be excluded from such a study. Given that housing density (i.e., lack of land on which to build) is seen as a major driver of housing supply inelasticity, a regression discontinuity study would not represent a balanced sample of Indiana school districts.

[14] School attendance zones are observed using the National Center for Education Statistics' (NCES) School Attendance Boundary Survey (SABS) of primary schools for 2013-2014. For houses located in overlapping school attendance zones (i.e., where multiple schools draw students from the same area), houses are assigned to the school with the smallest ratio of enrolled students to houses in the attendance zone. The idea is that schools with the densest residential spread of students are likely the ones considered to be "neighborhood schools." Note that only houses that have been sold from 2008 to 2015 are considered.

2.5 Conclusion

The thesis of this paper is that a state policy allowing and facilitating school districts to accept non-resident students does not have the same effect on local housing values as a policy requiring school districts to accept non-resident students. Both theoretical and empirical literature find that the introduction of *mandatory* inter-district open enrollment lowers housing values in desirable school districts and raises housing values in undesirable school districts. In contrast, using Indiana as a test case for states where inter-district open enrollment is optional for school districts, this paper fails to find co-movement of housing prices and transfer students.

This contrast can be explained by the combination of heterogeneous rates of capitalization of school quality into housing prices and selective participation in open enrollment by school districts. Theory predicts that places with more flexible housing supply should have a looser link between school quality and housing values. Thus, school districts where housing values are least threatened by open enrollment are more likely to embrace it. A possible mechanism for this selection to happen is that homeowners understand the connection between house prices and school quality and wield influence on the local education authority accordingly. That is, it seems unlikely that a whole neighborhood of property owners would possess something worth several thousand dollars to each—exclusive access to a high quality local school—and that they would fail to realize it and protect it. In places where exclusive access to the local schools is worth more, we might expect homeowners to work harder to protect it. Homeowners have a ready channel to influence their school corporation's actions by voting in the local school board election or even attempting to join the school board. In a similar story, Brunner et al. (2001) and Brunner & Sonstelie (2003) show that California homeowners in high performing school districts consciously opposed a referendum proposing school vouchers out of fear that their property values would be negatively affected. Section 2.4.2.1 gives evidence consistent with this story. Even if residents or school authorities do not consciously select their level of participation in inter-district open enrollment based on the likely impacts on house prices, it is still true that districts with weaker capitalization of school quality into house prices seem to enroll more transfer students. This relationship is shown in Section 2.4.2.3.

Optional inter-district open enrollment policies like the one in Indiana exist in 13 states (Education Commission of the States 2015). That is the same number of states that have mandatory open enrollment policies. One might say that the message to policymakers from earlier research on this topic is that exogenously-assigned inter-district open enrollment can have surprising effects on house prices. This paper offers the counterpoint that, to varying degrees depending on state policy, levels of inter-district open enrollment are actually endogenously determined and that this endogeneity significantly blunts unintended impacts on house prices.[15] Unfortunately, making open enrollment optional also blunts the intended effect of the policy by denying access to open enrollment for many students.

[15]Even in states with nominally mandatory IOE, school districts retain some ability to influence the number of transfer students they accept (Reback 2008).

Table 2.1: Changes in Housing Prices upon the Introduction of Voluntary Open Enrollment

	House Sale Price			
	Arm's-length Sales		Repeat Sales	
	OLS	IV	OLS	IV
	(1)	(2)	(3)	(4)
Panel 1				
% Net Transfers	6447	−25237	18860	37419
	(13276)	(53588)	(21337)	(71985)
Panel 2				
% Net Transfers	9619	243265	2962	602775*
	(14339)	(177197)	(25286)	(346887)
% Net Transfers	−9008	−465191	46988	−1026052*
% Developed	(36263)	(298424)	(116478)	(617688)
House Sales	92676		23888	
School Districts	281		277	
Years	2010, 2015		2008-2010, 2014-2016	
Mean	$139405		$154867	
S.D	$95021		$93251	

Notes: The dependent variable is house sale price. The independent variables of interest are (1) the net inflow of transfer students in a school district, calculated as the number of incoming transfers students minus the number of outgoing transfer students divided by the number of students residing in the school district and (2) the above variable multiplied by the proportion of usable land in school district that has been developed (only included in Panel 2). The regressions in columns 1-2 also control for a vector of house characteristics and a school district fixed effect. The regressions in columns 3-4 control for a house fixed effect. Columns 2 and 4 instrument for net inflow of transfers students using difference in proportion of students who are designated as white between the school district and its neighbors and the number of neighboring school districts, following Brunner et al. (2012). Standard errors are clustered by school district.

Table 2.2: Factors that Predict Transfer Enrollment

	% Transfers Received (1)	% Transfers Sent (2)
% Transfers Sent	0.383***	
	(0.114)	
% Transfers Received		0.105**
		(0.044)
% Developed	−0.041**	0.003
	(0.020)	(0.010)
% Max Residency	−0.132**	−0.110**
	(0.064)	(0.043)
Better Test Scores than Neighbors	0.088***	−0.054***
	(0.031)	(0.015)
Higher Share White than Neighbors	0.120**	−0.031
	(0.058)	(0.026)
Number of Neighboring Districts	−0.005**	−0.003**
	(0.002)	(0.001)
% Charter Enrollment	0.215	−0.037
	(0.129)	(0.111)
% Voucher Enrollment	−0.217	−0.265
	(0.239)	(0.173)
N	286	286
Mean	0.063	0.045
S.D	0.078	0.042

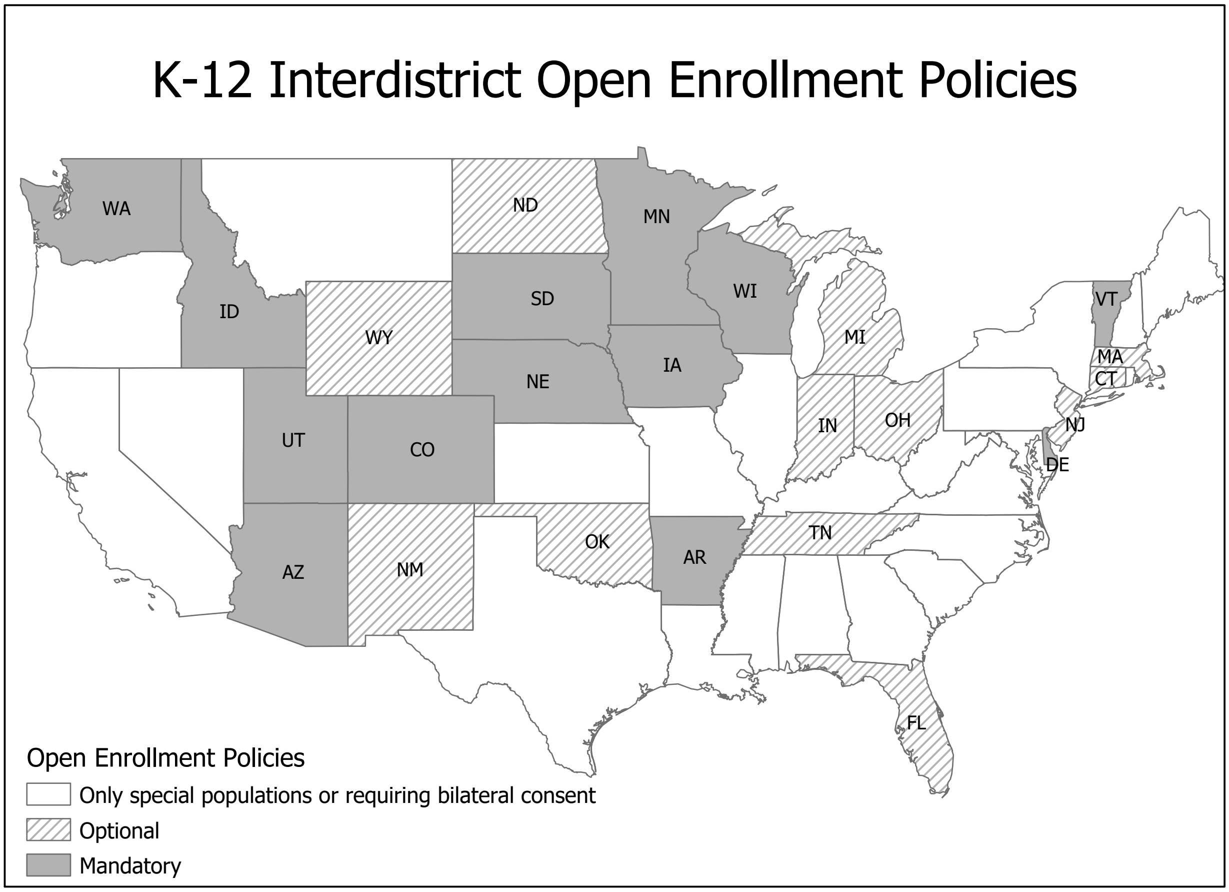
Notes: The dependent variable in Column 1 is percentage of each school corporation's 2014-2015 enrollment composed of transfer students, % Transfers Received. In Column 2, the dependent variable is the number of a district's resident students enrolled as transfer students outside the district divided by the corresponding school corporation's enrollment, % Transfers Sent. % Developed is the share of each school district's developable land that is already developed. % Max Residency is the number of students who resided in the school district in 2015 divided by the maximum number of students residing in the school district from 2005 to 2015. Better Test Scores than Neighbors is the difference between a district's average standardized test scores and average scores in neighboring districts, weighted by enrollment. Higher Share White than Neighbors is similarly the difference between percentage of students enrolled in the school district and the neighboring districts, weighted by enrollment. % Charter Enrollment and % Voucher Enrollment are the share of school district residents enrolled in a charter school or in a private school using a voucher, respectively. Robust standard errors are listed in parentheses with significance levels designated by * $p < 0.10$, ** $p < 0.05$, and *** $p < 0.01$.

Table 2.3: Capitalization of School Quality into House Price

	House Sale Price		
Neighborhood School Quality	Whole State (1)	$\leq 1\%$ Transfers In (2)	$> 1\%$ Transfers In (3)
	16692**	21307***	14420
	(6992)	(3407)	(10154)
House Sales	274199	108841	166055
Schools	690	286	404
Mean	$138646	$150431	$130921
S.D	$94863	$98750	$91406

Notes: The dependent variable is house sale price. The independent variable of interest is the average 2005 standardized test score of students at the associated neighborhood school. The regressions also control for a vector of house characteristics. Transactions from 2008-2016 are used. Column 1 includes transactions from the whole state, Column 2 includes transactions from school districts whose enrollment of transfer students is less than $1/100^{th}$ of the number of students residing in that district, and Column 3 includes transactions from all other school districts.

Figure 2.1: States with General, Statewide Inter-District Open Enrollment Programs



Note: 20 other states have inter-district open enrollment programs (optional or mandatory) for a limited subset of students or which require approval of the "home district" (ECS, 2015).

Figure 2.2: Transfer Enrollment and Test Scores Relative to Students in Neighboring Districts

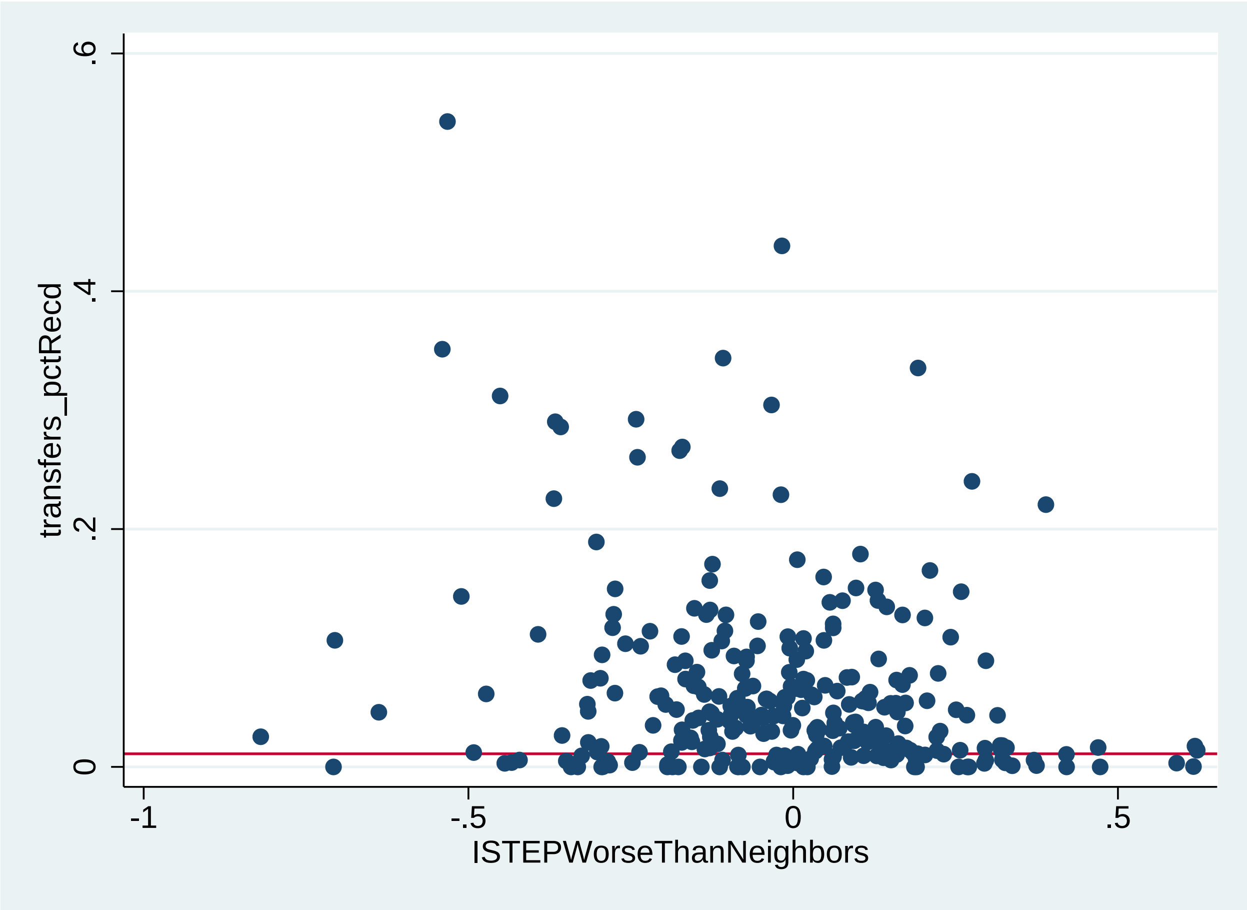
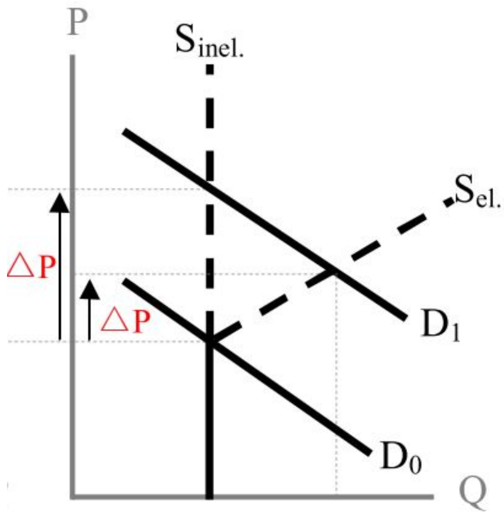
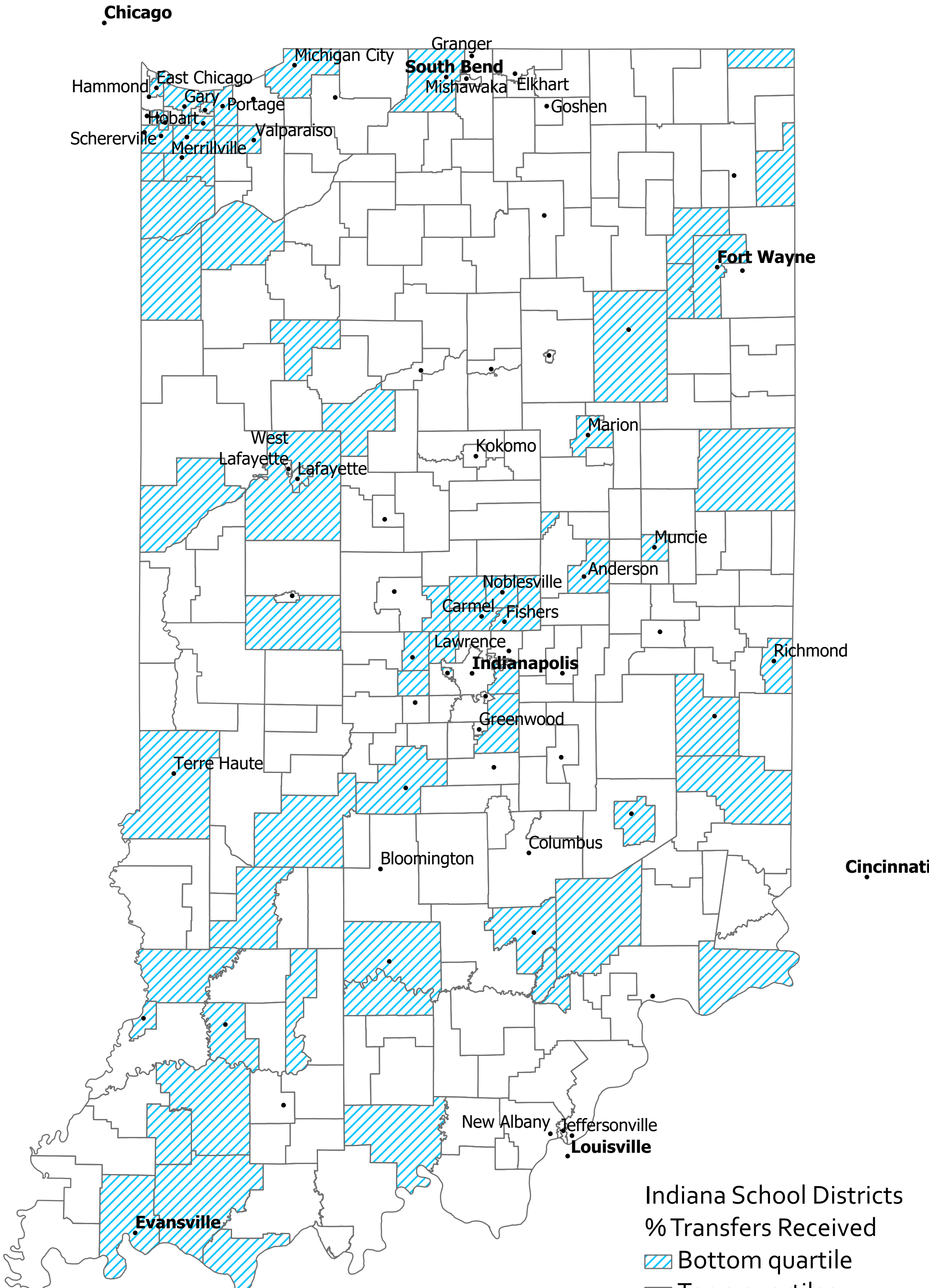


Figure 2.3: Effects of a Housing Demand Shock on Price with Heterogeneous Housing Supply Elasticity



Note: Copied from Hilber (2015) Figure 1, Panel C.

Figure 2.4: School Districts' Acceptance of Transfer Students



Chapter 3

The Rise and Effects of Homeowners Associations

3.1 Introduction

Most new homes constructed in the United States in recent years are part of a *common interest development* ("CID"), governed by a *homeowners association* ("HOA"). HOAs were rare until the 1960s but have come to house a fifth of Americans, a profusion that Guberman (2004) calls "one of the most significant privatizations of local government functions in history." Popular and academic opinions are split over whether homeowners actually like to live in HOAs. Some contend that HOAs represent a sensible market solution to local public goods problems (Foldvary 1994, Barton & Silverman 1994, Nelson 2005, Community Associations Institute 2016*a*) while others see HOAs as a sort of unregulated hostage crisis with unwitting homeowners harassed by busybody neighbors (McKenzie 1994, Lucas 2013, Benson & DeBat 2014).

Economists typically assess residents' opinions about the value of local amenities such as the presence of an HOA, school quality, air quality, or property tax levels by pointing to house prices. We think that, if the presence or quality of one of these features raises the price of a house, it indicates that the marginal buyer values that feature, other characteristics of the house being held constant. Earlier authors have studied whether HOA homes are more expensive than comparable non-HOA homes, but only in one state or city at a time because data on the topic has been scarce. Conclusions of existing studies vary, and their different methods and data samples do not facilitate easy comparison.

This paper presents the first ever (nearly) national estimate of how HOAs affect single-family house prices, using public data curated by Zillow, Inc.[1] Estimates are also presented for eight of the nine Census divisions and for many metropolitan statistical areas—which stand in for cities. The data include prices, HOA membership indicators, and other building characteristics for 34 million housing transactions, observed from 44 U.S. states over 35 years[2]. HOA membership indicators are derived from "planned unit development" and "condominium" riders described in Zillow's collection of public mortgage records. The size of the sample allows estimates for specific locations, which can be usefully compared to each other since they are calculated using consistent data elements and methodology. Having local estimates might be interesting in itself for some readers, and it allows us to explore what local factors seem to influence the value of living in an HOA. The author's estimation approach relies on hedonic theory à la Rosen (1974) and is done with a flexible, semi-parametric generalized additive model (GAM) specification, which is common in the hedonic literature but has not yet been applied to HOAs. Traditional OLS regressions are also presented for the sake of transparency and comparison with earlier work. HOA fees are also scraped from

[1]Although Zillow's data describe both single- and multi-family housing, the relationship between house price and HOA membership is only studied for single-family housing. The reason is that sales of HOA and non-HOA multi-family homes cannot be compared as directly as sales of HOA and non-HOA single-family homes, as discussed in Section 3.3.

[2]Excludes New England, where the requisite data are not reported.

Redfin.com, an online real estate listing site that serves as a complementary data source, another first for this literature.

The paper's overarching conclusion is that single-family houses with an HOA generally sell at a premium—meaning HOAs create value—but with plenty of local variation. On average, buyers pay around 4 percent, or $13,500 extra for a house that is subject to an HOA. This is near the middle of the range of existing estimates from economic literature. Next to the $120 per month that an average HOA member on Redfin.com pays in dues, this price premium implies that each dollar paid to an HOA buys about $1.37 worth of benefits.[3] The HOA price premium varies by location: Southern cities tend to have higher HOA premiums while New York, cities in the Rust Belt, and cities on the West Coast exhibit lower premiums. City-level HOA price premiums are shown to correlate with a city's degree of land use regulation, level of economic inequality, implicit racial bias, and several other factors.

Along with price premiums, the data used for this paper reveal an unprecedentedly detailed picture of where and when HOAs were established, what the houses look like, and who lives in them. HOA membership is most common where land was developed recently. Accordingly, HOA houses are newer and slightly bigger. HOA neighborhoods are also found to be less racially homogeneous than non-HOA neighborhoods.

Section 3.2 of this paper explains in more detail what an HOA is, reasons why HOAs are an important institution to study, and what is known about them from earlier empirical study. Section 3.3 presents the local government real estate records used for this paper, explains how those records can be used to observe HOA membership, and why the analysis must be limited to single-family homes. Section 3.4 depicts the people, places, and structures found in HOAs. Section 3.5 uses hedonic regression to evaluate whether and how HOAs affect the

[3] Net present value of $120 per month at a 4 percent annual return is $36,000. See Section 3.6 for an explanation.

value of single-family homes contained in them. Section 3.6 discusses implications of the geographic variation found in HOA price premiums, and Section 3.7 concludes.

3.2 Background and Literature Review

3.2.1 What is a Homeowners Association?

Buyers of property in a common interest development purchase a *divided interest* that belongs only to the individual buyer plus a share in a *common interest* that is owned corporately with the owners of nearby homes. Common interests in residential housing can range from pools and parks to streets and sewers—or the building itself in the case of condominiums. Owners in CIDs are also subject to a set of contractually enforceable rules, known as *covenants, conditions, and restrictions* (CC&Rs). By law, each CID must have a homeowners association to manage shared property and enforce the CC&Rs. The owner of every home in a CID automatically becomes a voting member of its HOA. This is analogous to buying shares of stock in a corporation; in fact, most HOAs are organized as non-profit corporations with a board elected from their membership. Homeowners associations are empowered to collect dues and fines, which are enforceable in civil court (McKenzie 1994, Esquivel & Alvayay 2014).[4]

Questions surrounding homeowners associations are broad, while the questions answered in this and earlier empirical studies are relatively narrow. We will first consider some wide-ranging questions about HOAs discussed in academic literature—to demonstrate the importance of studying HOAs and the potential uses of better data—before outlining previous empirical studies on HOA house price capitalization.

[4] For simplicity, the rest of the paper will mostly ignore the distinction between CIDs and HOAs and refer to both the physical housing and governing organization as an "HOA."

3.2.2 Broad Questions

Common interest developments evolved in England during the 19th century and first made it to America with the establishment of Gramercy Park in Manhattan in 1831: houses were built around a private park and charged with the park's perpetual maintenance (McKenzie 1994). Reflecting this first use, Foldvary (1994) presents a case that HOAs are able to provide "territorial public goods" more effectively and efficiently using "contractual government" than can be accomplished by municipalities via "coercive government." Helsley & Strange (1998) use a theoretical economic model to predict that the existence of private governments (e.g., HOAs) reduces provision of goods by the public sector but increases aggregate welfare. The intuition for this result is that private governments allow tailoring of government service levels to individuals' heterogeneous preferences. Total welfare of non-members, who demand less services, increases since their taxes fall, while the welfare effects for members is indeterminate since they pay for extra services through membership dues.

McKenzie (1994) portrays HOAs more negatively. He argues that HOA buyers are often unaware of extensive CC&Rs before buying or lack a non-HOA alternative in their local market for the type of housing they want. He also characterizes HOAs as an instrument of exclusion and a successor to racially restrictive covenants (as in the first "C" in CC&Rs). The timing of HOAs' explosion in popularity around 1960 is suspect: it followed a 1948 Supreme Court decision that racially restrictive covenants (intended to prevented ethnic minorities from living in a neighborhood) were unenforceable and preceded the Fair House Act of 1968, which outlawed racially restrictive covenants entirely. McKenzie (1994) also explores how HOAs' basis in contract law allows them to legislate details of residents' lives and limit speech in ways that exceed the police power of municipal governments. In fact, the U.S. Congress felt compelled to pass the Freedom to Display the American Flag Act of 2005 in response to HOA restrictions against display of the American flag. Far-reaching authority of HOAs is especially worrisome given that only owners may vote in HOA elections, not renters.

Helsley & Strange (2000) formally model welfare effects when private social organizations allow individuals to "secede" from the larger society (e.g., HOAs or private schools). The model predicts that individuals with high socioeconomic status will secede, leaving those who do not secede unambiguously worse off.[5] Nelson (2009) postulates a rich political economy of HOAs wherein existing residents extract rents from new residents by forcing new housing to be built in HOAs. New residents pay for services they receive via HOA fees and contribute to public goods they do *not* receive via local taxes. Nelson (2009) also asserts that the popularity of HOAs has halted the creation of new municipalities, since HOA residents in unincorporated areas want to avoid the problem of double taxation noted above. Foldvary (1994) wonders whether HOAs will be able to flexibly adjust when structures age and optimal land uses change, noting that the relative newness of most HOAs has so far kept obsolescence from being a big issue.

3.2.3 Empirical Studies

A number of empirical HOA studies have focused on how HOAs affect house prices, the topic of this paper. It is an important question because home price capitalization of local public goods, like neighborhood or school quality, is a measure of how the market values that good (Foldvary 1994). Hopkins (2017) provides an up-to-date summary of empirical literature on home price capitalization of HOAs.

The most relevant prior study of HOA price effects is Meltzer & Cheung (2014). Rachel Meltzer and Ron Cheung obtain a list of home addresses for Florida HOA board members

[5] The seemingly contradictory conclusions of Helsley & Strange (1998) and Helsley & Strange (2000) highlight the truth that conclusions of formal economic modeling depend critically on what factors are taken into account.

from a marketing company, likely derived from the Florida Department of State business registry. Board member addresses are then matched to publicly recorded transaction and assessment data and subdivision boundaries. Subdivisions with board member residents are then treated as HOAs, under the twin assumptions that board members live in the HOAs they serve and that HOA boundaries match subdivision borders. This deductive process yields a set of 583,000 sales of single-family homes in Florida suburbs occurring from 1960 to 2008 that are used for analysis. Applying hedonic regression, Meltzer & Cheung (2014) estimate that HOAs raise house prices by nearly 5 percent. They also conclude that the HOA price premium falls as a subdivision ages, that larger HOAs offer a smaller premium, and that homes located nearby an HOA also exhibit a price premium, indicating positive spillover effects.[6] The current paper improves upon Meltzer & Cheung's (2014) method of observing HOAs, even within the state of Florida, by utilizing HOA indicators for each house. House-level data mitigates the risk of entire subdivisions being mis-classified by one wrong address, allows observation of HOAs that might not align with subdivision boundaries, and is more transparent in its flaws since a researcher can look at individual houses in a computer mapping program for patterns that do not make sense. Additionally, the data used for the current paper were provided free of charge by Zillow, instead of requiring purchase (see Section 3.3).

[6] Rachel Meltzer and Ron Cheung also use these data for a number of other papers. Cheung & Meltzer (2013) show that a higher density of HOA housing in a city is correlated with fewer forms of land use regulation. Cheung & Meltzer (2014) show that HOAs formed earliest in Census tracts that were predominantly white, higher income, further from the city center, and with high vacancy rates. HOAs formed more slowly in cities with high expenditures on public services. Cheung et al. (2014) show that Florida home prices react less negatively to rising mortgage delinquency rates in cities where HOAs are more prevalent, suggesting that HOAs mitigate neglect of foreclosed homes. Meltzer (2013) explores whether increased presence of HOAs over time is associated with cities becoming more segregated. She concludes that a 10 percent increase in HOA share leads to a 1–2 percent increase in local racial segregation, but has no impact on economic segregation. In a slightly older paper with different data, Cheung (2008) tests whether growth of HOAs leads to relative decreases in spending by California cities, as predicted by Helsley & Strange (2000). Using a list of HOAs derived from the Secretary of State of California business registry, he cleverly estimates the number of HOA homes that are located in 110 California cities over time and finds that a 10 percent increase in prevalence of HOA housing leads to a 1.5 percent fall in local government spending.

An earlier group of papers estimates the effect of *private covenants* on single-family house prices, an overlapping but wider category than HOAs.[7] These studies have small sample sizes and rely on data provided by local multiple listing services ("MLS"). Speyrer (1989) performs hedonic regression with 230 single-family home sales in Houston, TX, a city known for lacking any zoning laws. She finds that private covenants raise home prices by nearly 9 percent relative to houses with no zoning. Hughes & Turnbull (1996) run hedonic regression using 1,314 transactions in Baton Rouge, LA and find positive price premiums which decrease over time for homes with private covenants. A 10 year-old covenanted home in his sample is expected to enjoy a 6 percent premium over comparable non-covenanted homes, while a 20 year-old covenanted home would sell at only a 2 percent premium. Rogers (2006) uses 1,487 single-family home sales in Greeley, CO to consider the effect of different forms of covenants that can be implemented by HOAs. He finds price premiums of 2–3 percent for membership in an HOA and attributes the gains to *use restrictions*, rather than *building restrictions* which are not found to alter prices. Cannaday (1994) compares pet restrictions in 1,061 high-rise condomium units in Chicago to conclude that units in buildings that ban dogs but allow cats sell for a 16-17 percent premium over units in buildings that allow large dogs.

Groves (2008) is the first study to focus explicitly on whether HOAs (or "Residential Community Associations") raise house prices. He is also the first to introduce a larger data set, based on state government sale records and an extensive search of local deed restrictions: 124,878 sales of single-family homes in St. Louis County from 1992 to 2001. Groves' regression specifications that are most comparable to other studies find a 1–3 percent price premium for living in an HOA.[8]

[7] Houses can be subject to private covenants but not be members of an HOA as long as there is no common interest, i.e., property owned jointly with other home owners (Esquivel & Alvayay 2014).

[8] Groves's (2008) more complicated, preferred regression specification accounts for "spatial lag" of home price by including the average sale price of previously sold nearby homes as an independent variable. This addition alters the proper interpretation of his regression, since identification of an HOA price premium now relies only on houses at the edge of an HOA, which have non-HOA neighbors. With this specification,

For comparison to some of the studies mentioned, the approach used in this paper estimates a 6.0 percent HOA price premium for the state of Florida, which is statistically different and slightly higher than the 4.9% estimated by Meltzer & Cheung (2014)[9]; a 5.1 percent price premium in the Greeley, CO metropolitan statistical area (MSA), which is higher than the premium found by Rogers (2010), and a 6.9 percent price premium in the St. Louis MSA, which is higher than results by Groves (2008).

3.3 Data

Barbara Coyle-McCabe writes in *Homeowners Associations as Private Governments: What We Know, What We Don't Know, and Why It Matters*, "Data concerning the number and location of HOAs are inexact because most of the information about them resides in local property records, where their founding is recorded. They are not included as discrete entities in the U.S. Census, tracked by state governments, or comprehensively mapped by local governments" (Mccabe 2011).

That is, public records of the details of homeowners associations exist, but are "buried" in municipal records. It is easy to find founding documents for an individual HOA, but Herculean to compile them for the whole country. Fortunately, an industry has grown up around compiling and harmonizing local property records for business purposes. Zillow, a company that compiles property data for use in their online real estate viewing platform, has recently offered to share data with qualified researchers, calling it the Zillow Transaction and Assessment Dataset (ZTRAX).[10] Most founding documents to which Coyle-McCabe refers,

Groves's (2008) estimated HOA price premium falls to a precise 0 percent, indicating that the value of living in an HOA may be linked to a house's seclusion from other neighborhoods.

[9] State-level estimates are not presented elsewhere in this paper. A special regression was run for the state of Florida for the sake of comparing to Meltzer & Cheung (2014).

[10] More information on accessing ZTRAX can be found at http://www.zillow.com/ztrax. The results and opinions are those of the author(s) and do not reflect the position of Zillow Group.

such as parcel maps and text of deed restrictions, are not captured in ZTRAX; but one key indicator is captured: mortgage riders.

Mortgage lenders require that an addendum (or "rider") be added to publicly recorded loan documents for properties that are included in a common interest development (CID), since that membership abridges the lender's ownership rights in case of foreclosure. Thus, ZTRAX includes two relevant flags for mortgage transactions, indicating whether a mortgage has either a *condominium rider* or a *planned unit development rider* attached to it. The presence of either of these riders indicates that the home is bound by an HOA, as explained more fully later in this section.[11] (For simplicity, this paper references a notional "HOA rider" to indicate when either of these mortgage riders are present.) Along with mortgage records, ZTRAX also includes past sale prices, details of each property's physical features, and each parcel's geographic coordinates.[12]

3.3.1 Geographic Coverage

ZTRAX is an amalgamation of local databases, mostly collected at the county level. Presentation of the data is harmonized across jurisdictions, but there is substantive local variation in the data that are reported. At the most basic level, ZTRAX includes at least one hundred residential real estate transactions from 2,619 of the 3,142 U.S. counties. Of those, 2,021 counties have at least one mortgage with an HOA rider. It can be difficult to know whether the counties that lack any reported mortgage riders simply have no HOAs or if they are not reporting mortgage riders. (In the case of six New England states, no mortgage riders are

[11] The distinction between these two most common types of CID is a legal construct, rather than a difference of building type, so this paper combines them to form a single HOA indicator. Condominiums are typically built as multi-family developments, but can be composed of single-family homes (Esquivel & Alvayay 2014). It seems more informative to distinguish between physical form of housing developments rather than legal forms, as described below.

[12] ZTRAX geocodes are "enhanced Tiger coordinates." This means each block face is located using Census Bureau data, and locations of individual rooftops are interpolated based on their house number. Viewed on a map, many ZTRAX geocodes are slightly shifted from their true location but correctly capture the distance between neighboring properties.

reported for the whole state.) When estimating the effect of HOA membership on house values, transactions from counties with no HOA riders can be left out since comparisons should be between similar houses to the extent possible. When reporting descriptive statistics, I use a simple, admittedly arbitrary decision rule: any county with at least 5,000 new houses built since 1960 that doesn't report any HOAs is treated as a non-reporter and left out of the calculation. Counties with fewer than 5,000 houses built since 1960 are left in the calculation with the understanding that, even if they really do have HOAs, the counties' moderate size means omitting their HOAs will cause only modest downward bias in the estimated share of houses with an HOA. By these criteria, ZTRAX contains information about the HOA status of homes in 2,526 counties, containing 90 percent of the U.S. population.

Figure 3.1 provides a visualization of which years' real estate transactions are included in ZTRAX, by state.

3.3.2 Inferring HOA Status from Mortgage Riders

As stated above, mortgage origination rules dictate that loan documents for houses bound by an HOA should include either a planned unit development rider or a condominium rider, depending on the legal type of HOA. Thus, an idealized version of ZTRAX would yield a simple binary indicator of HOA inclusion for each house that has a recorded mortgage.

This simple idea gets complicated by real data. Apart from the limitations of geographic coverage discussed above, mortgage riders are also inconsistently reported within localities. Wrong HOA rider flags can arise at several points: Mortgage originators might mistakenly omit (or include) an HOA rider in the actual document, local government might record the loan documents with an error, or Zillow might err when adding public records to the ZTRAX database. In any case, inconsistency between multiple mortgages on the same property and visual inspection of ZTRAX data displayed on a map suggests non-trivial levels of both Type

1 and Type 2 errors. That is, some mortgage records of homes that truly *do* belong to an HOA lack the appropriate HOA flag and mortgages for some homes that truly *do not* belong to an HOA have an HOA flag. (See Figure 3.2.)

If a large number of mortgages were observed for each house, occasional wrong mortgage rider flags would not be a big problem. The modal HOA status of mortgages tied to each house would reliably indicate whether the house was in an HOA, thanks to the law of large numbers. (e.g., We could safely assume a house was in an HOA if 80 percent of its mortgages had an HOA rider.) Unfortunately, most houses have record of being mortgaged, at most, a few times.

Thus, this paper uses two separate approaches for determining whether a house is in an HOA. The first approach is simple: houses are designated as "HOA" if any mortgage on that house has ever included an HOA rider. Houses with no recorded mortgage are disregarded. The second approach works by finding groups of houses that probably all share the same HOA status, and deciding their HOA status together, based on the larger number of mortgages.

When should neighboring properties share the same HOA status? The legal process of establishing an HOA requires that all involved land owners agree to restrict their own property rights. Such complete agreement is usually only achieved when all the land is owned by one entity—a real estate developer. Developers subdivide larger plots of land, establish an HOA, and build houses that future HOA members will occupy. The houses are built in physically connected clusters, either all at once or in phases (Esquivel & Alvayay 2014). Thus, if a house is included in an HOA, it is a good bet that a neighboring house built in the same subdivision or in the same year is also in the HOA. A large enough group of neighboring houses built at the same time will be tied to enough mortgage records to reliably conclude whether or not the whole cluster is bound by an HOA.

With the above real estate development process in mind, two options arise for associating nearby houses that likely share the same HOA status. The paper will use both. First, houses can be grouped by subdivision whenever that information is included in ZTRAX. Second, even when a house is not in a recorded subdivision, clusters of houses that were built near each other at the same time can be observed using an algorithm. Once houses are grouped—by subdivision or by cluster—the whole group's HOA status can be judged based on how many of the houses have an HOA rider. The following decision rule is used: if less than 20 percent of ever-mortgaged houses in a subdivision or cluster have been flagged with an HOA rider, none of the houses in the group is considered part of an HOA. If over 60 percent, all houses in the group are treated as belonging to an HOA. Between 20 and 60 percent, no determination is made.

The details of the clustering algorithm used, analysis of how well subdivisions and clusters work in finding groups of houses with homogeneous HOA status, and justification for the cutoffs to decide HOA membership are presented in the appendix. Both methods—grouping by subdivision or by cluster—work similarly well and each produce a classification for nearly half the houses in the full sample. By using both grouping methods—accepting the result of either classification scheme as long as they do not contradict each other[13]—we can assign an HOA status based on neighboring houses for 33 million houses, or 61 percent of the full sample.

To summarize, a house's HOA status will be judged in this paper based on [1] whether any mortgage on the house has ever included an HOA rider or [2] whether most nearby houses' mortgages include HOA riders, in cases where we think the nearby homes all share the same HOA status. The reason for creating a neighborhood-level measure of HOA status is that some snapshots of the simpler house-level measure arrayed on a map appear to show the wrong HOA status for a substantial number of houses, in both directions (see

[13]For the 14 million houses where both methods assign an HOA status, the two methods agree over 99 percent of the time.

Figure 3.2). How different are the house-level and neighborhood-level results? Assuming the neighborhood-based method is correct, the house-based assignment method has a Type 1 error rate (false positives) of 5.5 percent and a Type 2 error rate (false negatives) of 4.5 percent.

3.3.3 Other Data Elements in ZTRAX

In addition to houses' HOA status and sale prices, ZTRAX provides a wealth of information detailing the characteristics of land parcels, the structures on the land, and transactions involving the parcels. Table 3.1 lists and briefly explains the data elements from ZTRAX used in this paper. Table 3.2 presents descriptive statistics involving those data elements.

3.3.4 Focus on Single-Family Houses

Homeowners associations govern both single- and multi-family housing, but this paper only looks at single-family housing. Why the omission?

Remember that property in a common interest development consists of a *divided interest* bundled together with a *common interest*, which must be governed by an HOA. A neighborhood of single-family homes without any common interests needs no HOA. Thus, one can imagine comparing sale prices of identical single-family homes, where some of the homes include a common interest *cum* HOA. Systematic differences between sale prices of HOA and non-HOA homes would be attributable to the value of the common interest + HOA bundle.

With multi-family homes, comparing sales of otherwise identical units with and without an HOA is impossible. A multi-family building must have an HOA if the units are to be owned by multiple parties, since the walls are shared. The building may not have an HOA if all units have one owner, but, in that case, any sale is of the whole building. That cannot be compared to the sale price for a single unit. Lastly, sales of multi-family HOA homes cannot be compared to sales of single-family non-HOA homes, since there is no way to separate the value of HOA membership from the value of being in a multi-family building. Sales of multi-family HOA units have no suitable comparison group.

Apart from this conceptual problem, ZTRAX has data limitations that make study of HOAs in multi-family housing difficult. For single-family homes and multi-family condominium units, each property tax record corresponds to one dwelling. But for multi-family apartment buildings, one tax record can correspond to many units. The same is true for cooperatives, a less common form of multi-family HOA popular in New York City. Even though ZTRAX includes a field for the number of units in a building, it is hard to gauge how consistently that field is populated. Thus, it is difficult to even count what share of multi-family dwellings in the U.S. have an HOA using ZTRAX.

3.3.5 Excluding Home Sales with No Mortgage

ZTRAX draws from public data sources that record *all* real estate transactions. Many transactions occur between related parties and do not reflect a market price. Anecdotally, the author has noted in prior work that Indiana tax assessors utilize less than half of all real estate transactions to estimate market prices. Non-market transactions decrease the precision of regression results and can also bias results if non-HOA homes are more frequently transacted below the market price than HOA homes.

The logical step is to eliminate apparent non-market transactions. The author does this partially by dropping transactions with an *intra-family transfer flag* and sales priced below $1,000. However, exploratory regression results still indicate a likely problem: the average home sale that includes a mortgage is priced 30 percent higher than the sale of an observationally equivalent home where they buyer does not take a mortgage. Such a large, unjustifiable difference raises a worry that many homes sold without a mortgage are not sold at market prices. For this reason, estimates of the value of HOA membership presented later in this paper only use sale records that are accompanied by a mortgage, representing around 60 percent of sale records.

3.3.6 Redfin Data

A secondary, also novel source of data for this paper is HOA fees described in online house listings. Standard real estate listing forms include a field for monthly HOA fee. Redfin.com is a real estate listings site that allows visitors to download a text file with details from up to 350 house listings in a given ZIP code at a time. Those details include the monthly HOA fee, if any. Using a computer program, the author gathered a list of 900,000 house listings appearing on Redfin.com on June 5, 2017.[14]

3.4 Descriptive Statistics About Common Interest Developments

The data used in this paper reveal HOA neighborhoods in unprecedented detail. Before estimating how membership in an HOA shapes house price, it seems interesting to explore where HOAs are, who lives in them, and what the buildings look like. This section focuses on our simplest measure of HOA membership—whether any mortgage on each house has ever included an HOA rider—in order to capture a representative sample of homes. The neighborhood-based method of eliciting HOA membership restricts focus to subdivisions and clusters of homes built together.

[14] Code was adapted by the author from the GitHub repository of Lopez (2017).

3.4.1 Profusion of HOAs

Time Trend

Figure 3.4 indicates the share of single-family homes built in each year that are governed by an HOA. Note the steady increase in popularity of HOAs beginning in the 1960s. Almost 60 percent of new single-family home construction today is in an HOA. What about land being developed for the first time? The top trend line in Figure 3.4 shows single-family homes built in new subdivisions and utilizes the neighborhood-based measure of HOA membership. We see that around 80 percent of single-family homes built in new subdivisions today are bound by an HOA.[15]

Current Housing Stock

The bottom panel of Figure 3.5 shows the share of houses that are in an HOA by county, demonstrating that HOAs are mostly found in and especially around cities. This fact reflects patterns of new residential land development since 1960. The top panel of Figure 3.5 shows the share of houses in an HOA by Census division. We see that HOAs are most common in the Mountain States, the states around Texas, and the southern Atlantic coast. Figure 3.6 mirrors Figure 3.5, but restricts to houses built in new subdivisions since 2000 to show the ubiquity of HOAs in recent new land development.

3.4.2 Who Lives in an HOA?

This subsection ties ZTRAX data to 2010 decennial Census estimates to learn about the people living in HOAs (Manson et al. 2017). The smallest area for which the Census estimates

[15] This group includes homes whose age is within 5 years of the mode in their subdivision, constituting more than half of single-family homes constructed in recent years.

demographic statistics is a block group. The US is divided into over 200,000 block groups, generally housing from 600 to 3,000 people. A necessary assumption for this exercise is that residents of HOAs are representative of the Census block group in which they live.[16]

Table 3.3 presents estimates of the income (median annual household income) and ethnicity of HOA residents in single-family homes. This table is calculated by weighting each Census block group by its population multiplied by the share HOA. That is, a block group with 1,000 residents with 60 percent HOA homes will contribute a weight of 600 residents in the HOA column and 400 residents in the non-HOA column. This method of weighting reveals that HOA households earn over one third more than non-HOA residents. Comparing the ethnicity of HOA- versus non-HOA residents, HOAs contain slightly more white residents, less Hispanics, less blacks, and more Asians.

Table 3.4 presents isolation indices by ethic group and HOA status, again for residents of single-family homes. This commonly-used measure of residential segregation calculates the average percentage of residents in one's own block group that share one's own race. To illustrate, the figure for white HOA residents in Table 3.4 means that the average white non-Hispanic HOA-resident lives in a Census block group with 73.6 percent other white residents. White non-HOA residents live among 78.1 percent whites, on average. Thus, white HOA residents have more racially diverse neighbors than white non-HOA residents. The same is true for Hispanics, blacks and Asians.

One might suspect that this finding is due to HOAs being built in less segregated parts of the country. However, even if counties are re-weighted according to their HOA population,

[16]One could soften this assumption by leveraging the names of property holders, which are included in ZTRAX, as an indication of race.

we still find that Census block groups with high HOA shares are less segregated than block groups with low HOA shares.[17]

3.4.3 Building Characteristics of HOA Housing

Table 3.2 shows average characteristics of single-family homes with and without an HOA. Homes in an HOA are much newer on average, have around 20 percent more floor space although they are situated on smaller land parcels, and sell for about 25 percent more than non-HOA homes. Additionally, property taxes on homes in an HOA are around 0.4% higher, and they are three times as likely to sport a tile roof.

3.5 Hedonic Valuation of HOA Membership

Proponents argue that HOAs raise property values. In other words, it is nicer to live in an HOA than not, at least in the marginal buyer's view. This section will test that assertion by checking whether the average selling price of a home marked with an "HOA" flag in the

---

[17]The isolation index is calculated as,

$$\frac{\sum_i \frac{x_i}{t_i}\left\{x_i \frac{c_i}{h_i}\right\}}{\sum_i \left\{x_i \frac{c_i}{h_i}\right\}} = isol,$$

where $\frac{x_i}{t_i}$ is the ethnic percentage of the block group (i.e., white, Hispanic, black, or Asian), calculated as the ethnic population x_i divided by the total population t_i of block group i; and $\left\{x_i \frac{c_i}{h_i}\right\}$ is the ethnic population of block group i adjusted for the share of houses that are part of an HOA, with c_i being the number of HOA houses and h_i being the total number of houses in block group i.

The re-weighted isolation index is calculated as before, but computed within county and weighted by ethnic HOA population χ_j of county j (the county's ethnic population multiplied by share of houses in an HOA):

$$\frac{\sum_j \chi_j * isol_j}{\sum_j \chi_j}.$$

ZTRAX database is higher than the average price of an observationally-equivalent home without the flag. Results indicate that, yes, HOA homes cost a few percentage points more than non-HOA homes.

The econometric challenge of studying how HOA membership affects house price is that HOA membership is always sold in a bundle with a private home. There is no market price of HOA membership, which is the same dilemma faced when studying the value of other features of a home, like solar panels. I approach the problem as other authors have done by using hedonic regression in the tradition of Rosen (1974) to estimate the contribution of HOA membership to house prices.[18] Different from prior authors, estimation is done using both a readily-interpretable log-linear OLS model and also a more flexible semi-parametric regression model, intended to better fit the presumably non-linear house price function. Additionally, regressions are presented using both measures of HOA membership described in Section 3.3, and broken out by city and region of the country (i.e., metropolitan statistical area (MSA) and Census division).

The largest regression sample used includes 10.1 million transactions. As discussed in Section 3.3, this group includes sales of single-family homes built since 1960, sold since 2005, and recorded along with a mortgage loan. When restricted to the neighborhood-based measure of HOA membership, the national regression includes 8 million transactions. Census-division level regressions give an idea of where those transactions took place, with the largest numbers coming from the South Atlantic division (3.2 million, mostly from Florida) and the Pacific division (2.7 million, mostly from California).

[18] A controlled field experiment for this question is unrealistic. One might imagine asking housing developers to design two plans for each of a set of building sites, one with and one without an HOA; randomly choosing which plan to develop at each site; then waiting many years to watch how the neighborhoods evolve. Quasi-experimental methods, which leverage randomness in past events to interpret them as experiments, are also ill-equipped for studying HOAs because of the complexity of housing investment decisions and the long time periods involved. For example, it would be difficult to convincingly tie even a perfectly random policy change regarding HOAs from 30 years ago to housing prices today.

3.5.1 Description of Hedonic Regression Models

The first regression model presented is log-linear OLS. Log sale price is regressed on the property and transaction characteristics described in Table 3.5 plus a vector of Census-tract indicator variables. That is:

$$\mathrm{lnP}_{ijt} = \beta_0 + \beta_1 \mathrm{HOA}_i + \mathrm{X}_{it} B_2 + d_j + \epsilon_{ijt} \tag{3.1}$$

where lnP_{ijt} is the log sale price of home i, located in Census tract j, and sold at time t; HOA_i is an indicator for whether home i is part of a common interest development; X_{it} is a vector of house- and transaction-specific characteristics described in Table 3.5; d_j is a vector of Census tract dummies; and ϵ_{ijt} is an error term.

The resulting regression coefficients can be interpreted as the percentage change in sale price expected in response to a change in the independent variable, other things equal. This interpretation is deliberately straight forward, but the log-linear model implicitly supposes that house prices are determined as a log-linear function of the included variables. Hedonic price theory gives no reason to think the function is log-linear. Thus, we also consider the following generalized additive model ("GAM") for the hedonic price function.

GAMs are a relatively popular form of semi-parametric estimator in which at least some of the individual terms are estimated non-parametrically, but they are combined additively with other terms (Hastie & Tibshirani 1990, Wood 2017). The GAM model used is:

$$\mathrm{lnP}_{ijt} = \beta_0 + \beta_1 \mathrm{HOA}_i + \sum_p f_p(x_p) + \epsilon_{ijt} \tag{3.2}$$

where f_p is a non-linear data driven function of x_p, known as a *smooth*.[19] A limitation of GAM models is that only continuous variables can be smoothed. Discrete variables, such as the number of bedrooms, can either enter linearly or be combined into a continuous index that can then smoothed. The model presented here does the latter by first predicting log sale price as a linear function of the discrete variables to create an index, then modeling log sale price as a function of HOA status and a vector of continuous variables. All continuous variables are de-meaned within Census tract, which is equivalent to including a vector of Census tract dummies but is more computationally tractable. The vector of continuous variables x_p includes all continuous variables used in the OLS model, the index of discrete variables, and an interaction of latitude and longitude which helps the model account for spatial autocorrelation (Wood 2017).

3.5.2 Results

National Results

Table 3.5 presents national regression coefficients for six model specifications: The basic OLS model in Column 1 estimates that an HOA raises prices by 7.1 percent on average and provides plausible estimates for the effects of other housing characteristics. As noted, the OLS model is likely to provide a poor fit of the true function, but it is comforting to see that estimated effects of the housing characteristics included in the OLS model align generally with economic intuition and earlier hedonic housing studies. The OLS model is also useful for comparison to earlier HOA studies that estimate a linear model. Column 3 shows results of the same OLS model, but the neighborhood-based measure of HOA membership is used. Note that the neighborhood-based measure of HOA membership is applicable to a smaller sample. For transparency, Column 2 provides results using the simple house-based measure

[19] *Smooths* work the same as non-parametric functions applied to histograms to provide a smooth, local estimate in one variable.

of HOA membership, but on the smaller sample of homes covered by the neighborhood-based measure.

Columns 4–6 of Table 3.5 mirror columns 1–3, but they use the GAM model with discrete variables combined into a continuous index. Coefficients are provided only for the effect of HOA membership, since a GAM model fits a curve to continuous variables rather than producing a coefficient. Column 4 indicates that houses that have been flagged with an HOA mortgage rider sell for 4.4 percent more, fitted with the GAM model. Similarly, Column 6 says that houses in neighborhoods where most homes have been flagged with an HOA mortgage rider sell for 3.6 percent more, again, fitted with the GAM model.

Having established a near-national estimate of how HOA membership affects house value, it is also interesting to explore variation among U.S. regions and cities within a unified framework. Doing so can be intrinsically interesting, opens the door for direct comparison to related work that is based on more localized data, and allows us to consider how local conditions might affect the value of living in an HOA (see Section 3.6).

Regional Results

We first break the analysis out by Census division to show regional differences in how HOA status co-varies with house price. There are nine Census divisions, each comprised of a few states, as shown in the top panel of Figure 3.7. For brevity, only one regression specification is be presented for each division. The semi-parametric GAM specification is used since it should provide a better fit and HOA membership is indicated by the simpler house-based measure, corresponding to Column 4 of Table 3.5.

Referring to Figure 3.7, we see the smallest price difference between HOA and non-HOA single-family homes in the mid-Atlantic (including New York) with a 1.8 percent boost to the price of HOA-member houses, and along the West Coast, with a 3.2 percent price premium. The highest HOA price premiums are in the southern U.S., peaking in the East South Central division with a boost to property prices of 5.6 percent. Census division results are listed along with standard deviation in Table 3.7 in the Appendix.

Figure 3.7 also shows the distribution of where observations used in the national regression come from, with disproportionate shares coming from the states of California and Florida.

City-Level Results

The smallest geography for which results are presented is the metropolitan statistical area. MSAs are a Census geography that corresponds roughly to cities. The same regression specification used with Census divisions is used here—a GAM with HOA membership measured at the house level. The bottom panel of Figure 3.7 presents regression coefficients for 162 MSAs that have a sample size of at least 10,000 transactions or a statistically significant coefficient at the 99% level. Table 3.7 in the appendix presents regression results for MSAs with at least 1,000 observations, including the standard error and sample size.

Some of the smallest HOA coefficients are found in Detroit and New York City, with regression coefficients of -2.7 percent and -1.3 percent. Places with remarkably large HOA coefficients include Huntsville, AL (18.1 percent); Greenville, SC (17.8 percent); and Cape Coral, Naples, Port Saint Lucie, and Ocala, FL with coefficients of 16-21 percent.

3.6 Discussion of Results

HOA fees are analogous to property taxes, in that they are a mandatory ongoing stream of payments required of property owners, set through a democratic process of voting. It is also reasonable to think that HOA fees are determined via a market process wherein property owners can "vote with their feet," similar to local property taxes (Tiebout 1956). The consensus view among academic studies is that property taxes are usually partially capitalized into real estate prices (Sirmans et al. 2008). This means that the net effect of raising both property taxes and public expenditures is to lower real estate prices, but by less than their full value. Simply put, prices go down when taxes go up.

This paper and most prior related papers argue that HOAs have the opposite effect. Prices of HOA housing are higher, despite the increased costs of ownership from HOA fees. Redfin.com listings show that HOAs charge an average of $120 per month in fees (Figure 3.3). If prices went down by the full value of those fees (i.e., full capitalization), we might expect the average HOA house to cost $36,000 less than an observationally equivalent non-HOA house.[20] Nevertheless, results from Section 3.5 show that HOA houses are priced four percent higher than their non-HOA cousins, or $13,500 more for the average sale in an HOA subdivision from our sample (from Table 3.2). The above contrast implies that homeowners obtain $1.37 in benefits for each $1 paid in HOA fees, and that HOAs are efficient providers of "territorial public goods" (Foldvary 1994).

In a poll conducted by Zogby Analytics in 2016 for the Community Associations Institute, 87 percent of respondents reported a positive or neutral experience with their HOA (Community Associations Institute 2016*a*). This positivity contrasts to the negative sentiment of much news reporting about HOAs and the steady stream of invective against HOAs on internet forums. What aspects of HOA living do home buyers like? Unfortunately, the data used for this paper give little insight about the conditions of individual HOAs. We only know if a house is in an HOA.

Overall, HOAs range from the barely-existant to the hyper-active and can provide a diverse mix of amenities and services. An annual survey of large HOA communities conducted by

[20] $36,000 is the net present value of a $1,440 perpetual annuity, paid annually and discounted at a 4 percent interest rate. Do & Sirmans (1994) estimate that homeowners discount the value of property taxes at 4 percent and note that most studies use discount rates ranging from 3–6 percent.

the Community Associations Institute indicates that common amenities and services include sports and recreational facilities, enforcement of building restrictions, disaster preparedness and mitigation, gated access to the community, ownership of streets, social programming, landscaping, and maintenance of natural areas (Community Associations Institute 2016*b*). An HOA in the author's home town recently voted to purchase a nearby financially distressed golf course rather than allowing it to deteriorate further or pass to an alternative land use.

A separate question from what individual buyers value in an HOA is why HOA price premiums vary systematically across localities and regions of the country. Such differences are likely driven by factors that influence supply and demand for HOA housing. Table 3.6 shows results of a series of simple linear regressions relating MSA-level HOA price premiums to MSA-level conditions, one at a time. The coefficients in Table 3.6 are equivalent to the slope of the line of best fit for a scatter plot with HOA premiums on the vertical axis and the other MSA-level characteristic on the horizontal axis.

The metric in Table 3.6 with the most statistically significant relationship to MSA-level HOA price premium is the zoning index, or the Wharton Residential Land Use Regulation Index. HOAs are less highly valued where land use regulations abound. This finding is consistent with the idea that HOAs provide "private zoning," which is most valuable in the absence of public zoning. Speyrer (1989) presents this hypothesis when showing that HOA premiums are higher in Houston—which lacks zoning regulations—than in nearby municipalities that have zoning regulations.

Helsley & Strange (1998) model HOAs as a means of providing differentiated levels of public services within the same municipality. Thus, demand for HOAs could be higher when residents' demands for public services are more varied. Contexts with greater income inequality might feature a larger customer base willing to pay for higher service levels. In partial support of this idea, the coefficient for GINI index in Table 3.6 is positive but not quite statistically significant at the 10 percent level. The GINI coefficient is a measure of wealth inequality, valued at 0 if wealth is distributed evenly and at 1 if all wealth is held by one individual. A true positive effect of GINI coefficient on HOA price premium would mean that HOAs are more valuable in cities where wealth inequality is greater.

Another statistically significant coefficient in Table 3.6 is for racial preference, as measured by the "Black-White Implicit Association Test." This is an online test administered by a Harvard University-affiliated group called Project Implicit. Test subjects' reaction speed is measured when asked to identify positive and negative adjectives associated to pictures of European- and African-origin faces. The positive coefficient in Table 3.6 indicates that HOA premiums are highest in cities where the average white resident who takes the IAT has a harder time associating good adjectives with "black" faces, relative to their speed at the same task with "white" faces. This correlation is consistent with McKenzie (1994), who asserts that demand for HOAs is driven by a desire for exclusion and status.

Table 3.6 also shows that HOAs are more highly valued in places with cheaper real estate prices, higher rates of HOA affiliation, newer buildings, higher rates of homes being occupied by their owner, more individual householders, smaller MSA population size, more black residents, and fewer Hispanic residents. Links between HOA premium and the following metrics are statistically insignificant: the MSA's average education level, poverty rate, share of married householders, occupancy rates, median age, and share of white residents.

3.7 Conclusion

In November 2017, U.S. senator Rand Paul was tackled and severely injured by his retired neighbor, allegedly over a dispute about landscaping. Newspapers quoted the neighborhood's developer, saying the senator "was probably the hardest person to encourage to follow the (homeowners association rules) of anyone out here, because he has a strong belief in property rights" (Postrel 2017).

This paper set out to ask whether Americans like living in HOAs. The above story, and many others like it that appear in popular press, seem to indicate the answer is "No!" Senator Paul, for one, apparently bristles at HOA rules that limit his gardening—but something still prompted him to buy a home in his neighborhood. Perhaps he values the privately-owned streets around his house that prevented reporters from taking pictures at the scene and that stop Google Maps from displaying street-level images of his neighborhood.

Unquestionably, HOAs have become common. HOAs govern 80 percent of houses built in new subdivisions today, and a fifth of all existing single-family homes (Figure 3.4 and Table 3.2). Some have argued that this popularity is driven by land developers and local governments pushing HOAs on home buyers, as a way to increase profits and strengthen municipal budgets (e.g., McKenzie (1994)). If it were the case that HOAs were favored by suppliers but not consumers, we would expect prized HOA-free homes to sell at a higher price, after accounting for the size and quality of home. Instead, this paper finds the opposite, that homes with an HOA sell for roughly 4 percent more than observationally equivalent homes with no HOA. Love them or hate them, Americans are willing to pay a little extra to live in an HOA.

The data behind this estimate are public real estate records of counties housing 90 percent of Americans, furnished by Zillow. Estimates for smaller portions of the country reveal that HOA price premiums are highest in southern states and lowest in New York, the Rust Belt, and the West Coast. Since the data used do not reveal amenities and services provided by individual HOAs, the paper does not have much to say about what buyers value in an HOA, but the paper is able to show what local qualities correlate to a higher HOA price premium. Some qualities associated with large city-level price premiums for HOA homes are few public zoning regulations, lower house prices, and a greater degree of implicit racial bias in favor of Caucasians.

Apart from studying how HOAs affect house values, this paper also examines the features of HOA houses and who lives in them. Homes in an HOA are around 400 square feet larger on average, occupy smaller plots of land, and are three times more likely to be covered with a tiled roof (Table 3.2). The average HOA fee listed on Redfin.com is $1,400 per year (Figure 3.3), compared to $2,200 per year paid by the average home owner in property taxes (Kiernan 2018). HOA residents' annual household income is found to be $21,000 higher and the residents slightly whiter, less Hispanic, less black, and more Asian than non-HOA residents (Table 3.3). Interestingly, though, HOA residents of all races live in more racially diverse neighborhoods than do their non-HOA counterparts (Table 3.4).

Table 3.1: Data Elements Used from ZTRAX

Property Characteristics	**Description**
Building area	Usable area of primary home, in square feet
Lot area	Land area, in square feet
Garage area	Usable area of the garage, in square feet
Other structure area	Usable area of any secondary buildings, in square feet
Implied property tax rate	Property tax bill divided by assessed value in most recent year
Owner occupied	Whether the property is occupied by its owner
Year built	Year the main building was constructed
Bedrooms	Number of bedrooms
Baths	Number of bathrooms
Total rooms	Total number of rooms in the house
Topography	Note about the property's terrain (e.g., hilly/level ground)
Tile roof	Indicator for a tiled roof
Near golf course	Indicator for proximity to a golf course
Waterfront	Indicator for having a view of water
Flooring	Type of interior flooring: carpet, wood, or other
Exterior wall	Type of exterior walls: brick, siding, stucco, or wood
Fence	Indicator for having a fence
Fireplace	Indicator for having a fireplace
Pool	Indicator for having a pool
Deck	Indicator for having a deck
Census block	Most basic Census geography; includes state, county, and tract
Legal subdivision name	Links groups of parcels previously subdivided from one larger parcel
Property address latitude	TIGER block face coordinates, interpolated based on house number
Property address latitude	TIGER block face coordinates, interpolated based on house number

Transaction Characteristics	**Description**
Planned unit development rider	Indicator for inclusion of a PUD rider, on mortgage records
Condominium rider	Indicator for inclusion of a condominium rider, on mortgage records
Sale price amount	Price paid for property deed transfer
Sale price amount type	Indicator of how sale price was observed
Transaction date	Date the transaction was finalized
Intra-family transfer flag	Indicator for buyer and seller having a family relationship
Deed type	Type of property deed: warranty, foreclosure, or other
Public record type	Deed transfer, mortgage, deed transfer with concurrent mortgage

Notes: This table lists and briefly explains the data elements relating to land parcels and real estate transactions used for analysis from the Zillow Transaction and Assessment Dataset (ZTRAX).

Table 3.2: Descriptive Housing Statistics, by HOA Status

	All single-family parcels				Built since 1960, transacted since 2005, with accompanying mortgage			
	HOA by house		HOA by neigh.		HOA by house		HOA by neigh.	
	No	Yes	No	Yes	No	Yes	No	Yes
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Unduplicated parcels								
Building area (sqft)	1,907	2,299	1,793	2,274	1,907	2,364	1,828	2,355
Lot area (sqft)	15,954	11,241	12,816	10,121	16,584	10,954	13,151	9,750
Garage area (sqft)	279	321	279	324	294	339	299	343
Other structure area (sqft)	6,262	622	1,368	35	9,437	1,697	51	31
Implied property tax rate	2.5%	2.9%	2.5%	3.0%	2.5%	2.8%	2.5%	2.9%
Pool	11%	14%	12%	14%	8%	10%	9%	10%
Deck	11%	10%	9%	9%	12%	8%	10%	7%
Bedrooms	3.4	3.5	3.3	3.5	3.4	3.5	3.4	3.5
Baths	2.2	2.6	2.2	2.6	2.2	2.6	2.2	2.6
Total rooms	6.8	6.9	6.7	6.9	6.8	7	6.8	7
Tile roof	5%	15%	5%	16%	4%	16%	5%	17%
Golf	0.1%	0.4%	0.0%	0.4%	0.1%	0.3%	0.0%	0.3%
Waterfront	1.1%	1.2%	1.1%	1.2%	0.7%	1.3%	0.8%	1.3%
Flooring—carpet	16%	19%	17%	20%	15%	21%	16%	21%
Flooring—wood	2%	1%	2%	1%	2%	1%	2%	1%
Fence	4%	6%	5%	6%	3%	6%	4%	7%
Sample size (millions)	36.6	9.6	24.1	9.1	2.2	1.5	1.5	1.4
Built since 1960	21.6	9.1	14.7	10.3	-	-	-	-
+ transacted since 2005	4.0	2.3	2.8	2.2	-	-	-	-
+ with mortgage	2.2	1.5	1.5	1.4	-	-	-	-
Percent HOA	79%	21%	73%	27%	59%	41%	52%	48%
Transactions								
Sale Price ($1,000, nominal)	-	-	-	-	$287	$352	$282	$352
Transaction Date (year)	-	-	-	-	2009	2009	2009	2009
Age at sale (years)	-	-	-	-	24	6	25	5
Deed type—warranty	-	-	-	-	61%	73%	60%	74%
Deed type—foreclosure	-	-	-	-	3%	7%	3%	8%
Deed type—other	-	-	-	-	36%	20%	37%	19%
Sample size (millions)	-	-	-	-	5.7	4.4	4.2	3.8

Notes: Selected property and sale characteristics of single-family houses recorded in the Zillow Transaction and Assessment Dataset (ZTRAX). Columns 1 and 2 include all houses for which there is record of a mortgage, allowing a house-specific determination of HOA status. Columns 3 and 4 include houses for which a neighborhood-specific HOA status can be determined, based on mortgage riders for that house and neighboring houses as described in Section 3. Columns 5-8 mirror columns 1-4, but the sample is limited to houses used for regressions in Section 5: those built post-1960 and sold since 2005, concurrent with a new mortgage.

Table 3.3: Income and Race, by HOA Status

	HOA	Non-HOA
Median annual household income	$81,173	$59,702
Percent white	63.6%	62.3%
Percent Hispanic	15.6%	17.8%
Percent black	9.6%	12.3%
Percent Asian	7.4%	4.5%

Notes: Calculated as the weighted average of Census block group characteristics, where weights are population multiplied by the percentage of single-family houses in (not in) an HOA. House-specific HOA status is used.

Table 3.4: Racial/Ethnic Isolation Indices, by HOA Status

	HOA residents	Non-HOA residents	Difference
White	73.6%	78.1%	−4.5%
Hispanic	35.3%	50.9%	−15.6%
Black	30.7%	49.8%	−19.1%
Asian	23.0%	24.7%	−1.7%

Notes: Average experience of the percentage of one's "neighbors" (within Census block group) who share one's race/ethnicity. Footnote 17 contains the exact formula.

Table 3.5: National Regression Results

	Log price					
	OLS	OLS	OLS	GAM	GAM	GAM
	HOA house	HOA house	HOA neigh	HOA house	HOA house	HOA neigh
	(1)	(2)	(3)	(4)	(5)	(6)
HOA	0.0710	0.0716	0.0782	0.0441	0.0391	0.0365
	0.0004	0.0004	0.0005	0.0003	0.0003	0.0003
Log Building Area	0.5751	0.5879	0.5853	–	–	–
	0.0006	0.0007	0.0007			
Log Lot Area	0.0947	0.1120	0.1136	–	–	–
	0.0003	0.0003	0.0003			
Transaction Date	−0.0086	−0.0104	−0.0104	–	–	–
(days since start, x365)	0.0000	0.0000	0.0000			
Age at Sale (years)	−0.0044	−0.0048	−0.0047	–	–	–
	0.0000	0.0000	0.0000			
Implied Prop. Tax (%)	−0.0039	−0.0042	−0.0041	–	–	–
	0.0154	0.0162	0.0162			
Deed Type—Foreclosure	−0.0094	0.01	0.0102	–	–	–
	0.0012	0.0017	0.0017			
Deed Type—Warranty	0.0896	0.1108	0.1108	–	–	–
	0.0013	0.0017	0.0017			
Owner Occupied	−0.015	−0.0219	−0.0211	–	–	–
	0.0003	0.0004	0.0004			
Bedrooms—1	0.0636	0.079	0.0779	–	–	–
	0.0027	0.0033	0.0033			
Bedrooms—2	0.024	0.0276	0.0266	–	–	–
	0.0011	0.0013	0.0013			
Bedrooms—3	0.0089	0.007	0.0073	–	–	–
	0.001	0.0011	0.0011			
Bedrooms—4	0.0013	−0.0036	−0.0031	–	–	–
	0.001	0.0011	0.0011			
Bedrooms—5	−0.0165	−0.0253	−0.0243	–	–	–
	0.0012	0.0013	0.0013			
Bedrooms—6+	−0.0556	−0.0593	−0.0581	–	–	–
	0.0018	0.002	0.002			
Pool	0.1063	0.1038	0.1041	–	–	–
	0.0005	0.0005	0.0005			
Waterfront	0.247	0.1958	0.1946	–	–	–
	0.0019	0.0021	0.0021			
Flooring—Carpet	−0.0097	−0.012	−0.0116	–	–	–
	0.0007	0.0008	0.0008			
Exterior Wall—Brick	0.0561	0.0463	0.0449	–	–	–
	0.0012	0.0013	0.0013			
Exterior Wall—Siding	−0.0227	−0.0262	−0.0272	–	–	–
	0.0012	0.0013	0.0013			
R^2	0.31	0.32	0.32	0.36	0.39	0.39
N	10.1M	8.0M	8.0M	10.1M	8.0M	8.0M

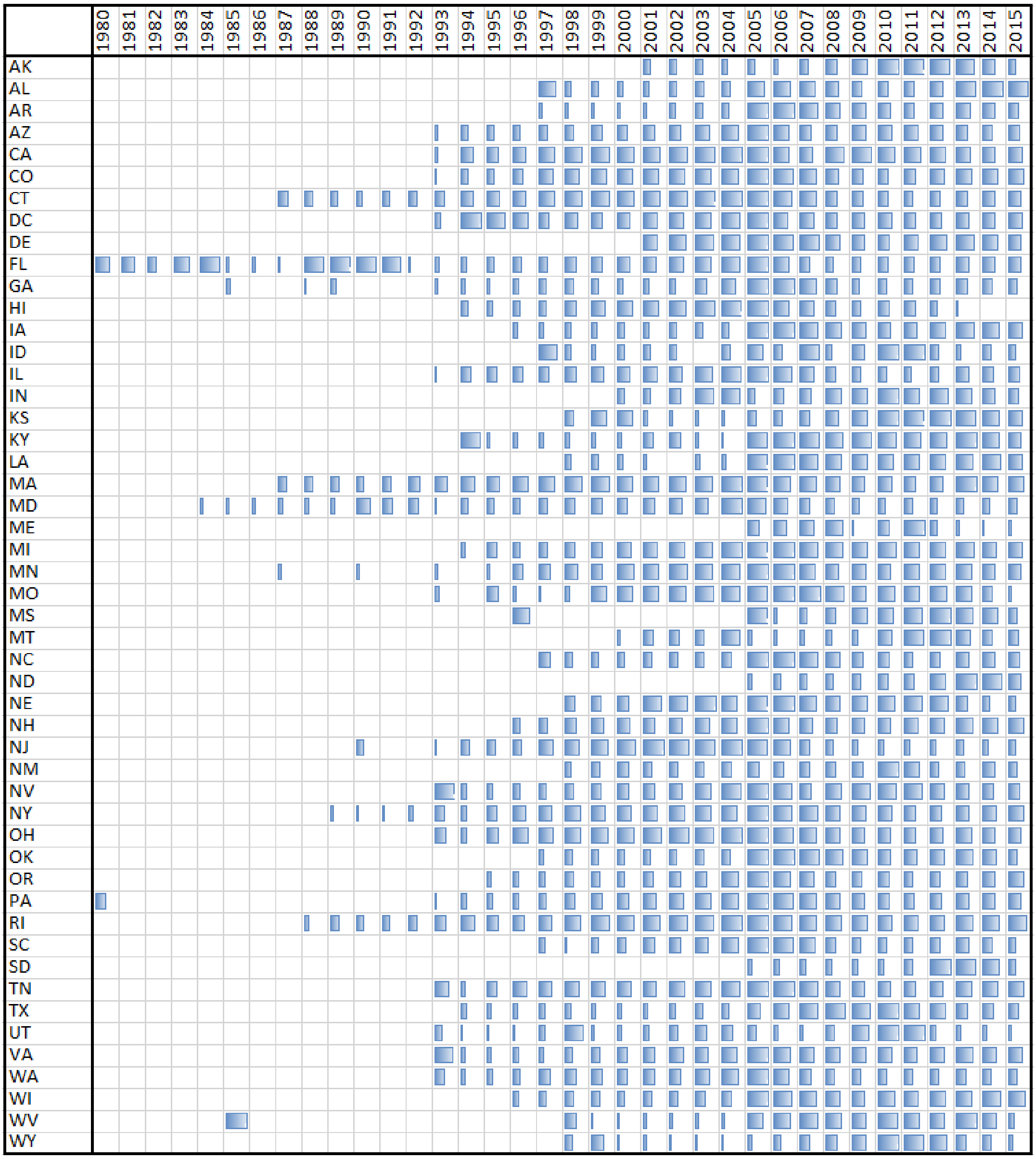
Notes: Regression includes single-family home sales since 2005, of homes built since 1960, that are recorded along with a mortgage. Regressors not displayed include garage area, other structure area, fireplace, deck, tile roof, fence, golf course nearby, baths (10 levels), total rooms (10 levels), topography (10 levels), flooring type (wood), exterior wall (stucco and wood), and Census tract. Columns 4-6 use a generalized additive model (GAM) with discrete variables other than HOA status combined into a single continuous index. The effects of continuous variables are estimated by functions not reported here, rather than coefficients.

Table 3.6: Relationships Between MSA-Level Characteristics and HOA Price Premiums

	Coefficient	Std deviation	t-Statistic	Sample size
Zoning strength, WRLURI	−2.10	0.49	−4.28	107
Median Price ($1,000)	−0.01	0.00	−3.43	159
Average year built	0.26	0.09	2.78	185
White preference, IAT	29.60	12.05	2.46	143
Population size (millions)	−0.39	0.17	−2.31	159
Owner occupancy rate	14.29	6.28	2.28	159
Percent black	8.37	3.75	2.23	159
Percent HOA	5.49	2.49	2.21	185
Percent hispanic	−6.82	3.12	−2.19	159
GINI coefficient	35.14	18.20	1.93	145
Percent solo householders	22.12	13.69	1.62	159
Poverty rate	18.52	11.45	1.62	159
Housing occupancy rate	−6.92	5.27	−1.31	159
Median resident age	0.07	0.08	0.88	159
Education—% HS grads	6.80	9.80	0.69	159
Education—% bachelors grads	−2.99	5.00	−0.60	159
Percent married householders	−4.70	10.30	−0.46	159
Percent white	1.05	2.56	0.41	159

Notes: Each line represents a bivariate OLS regression with MSA-specific HOA price premium as the dependent variable and the corresponding MSA-specific characteristic as the independent variable. Coefficients are multiplied by 100 for readability.

Figure 3.1: Geographic and Temporal Extent of ZTRAX Transaction Records



Note: Box size in each cell represents the number of single-family home sales recorded in that state and year, relative to the year with the most sales.

Figure 3.2: Example of Apparent HOA Status Mis-Classification, when Measured at House Level

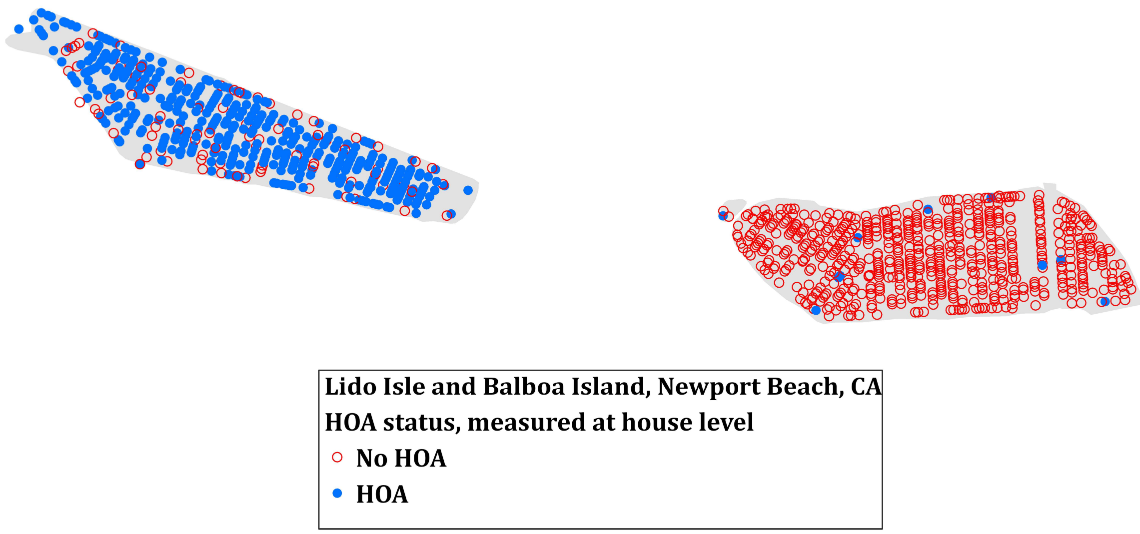


Figure 3.3: HOA Fees, Scraped from Redfin.com

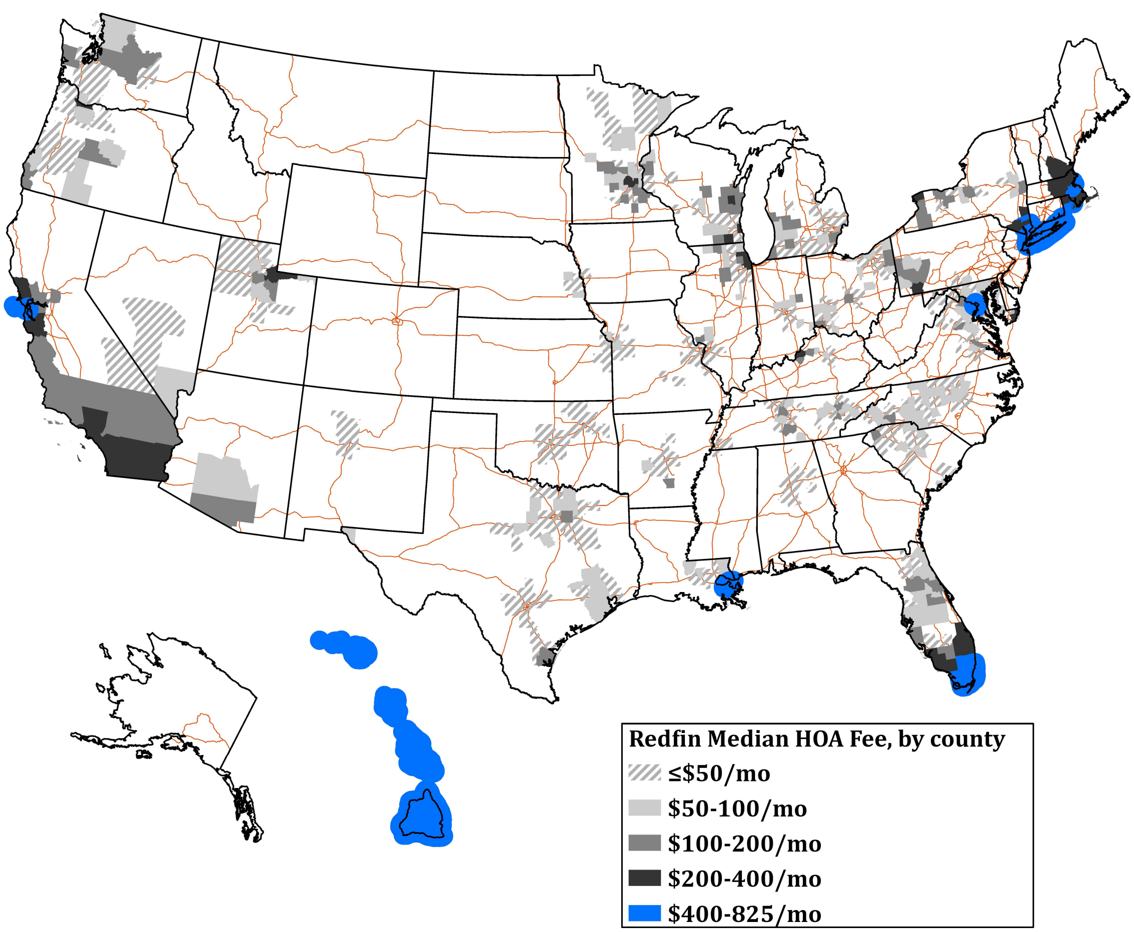


Figure 3.4: Percentage of U.S. Single-Family Homes Built with an HOA, by Year

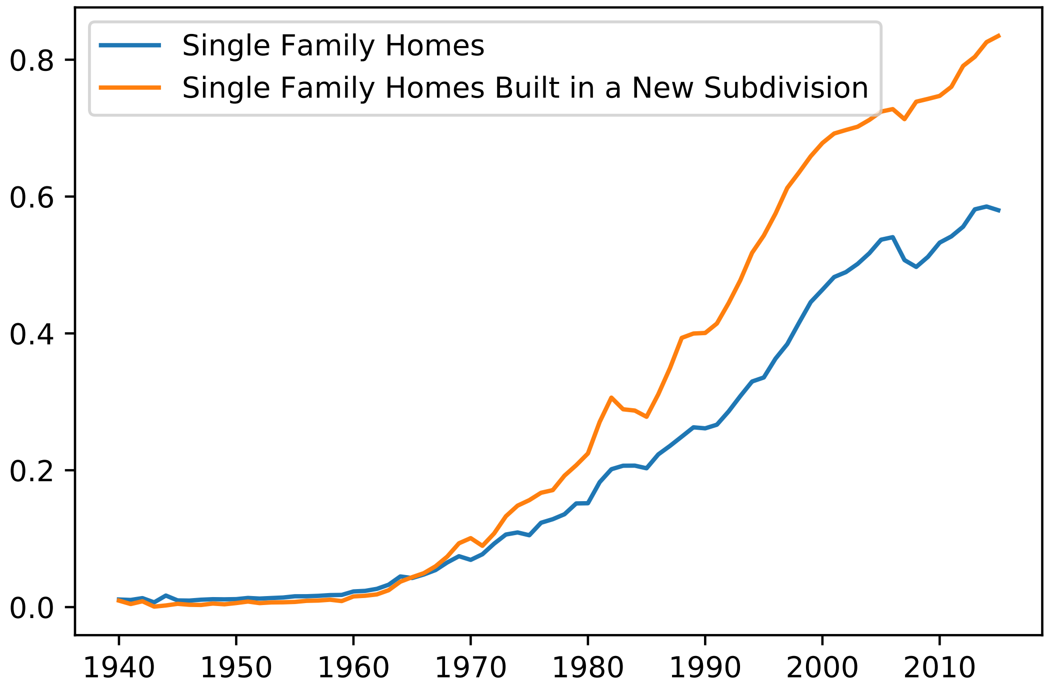
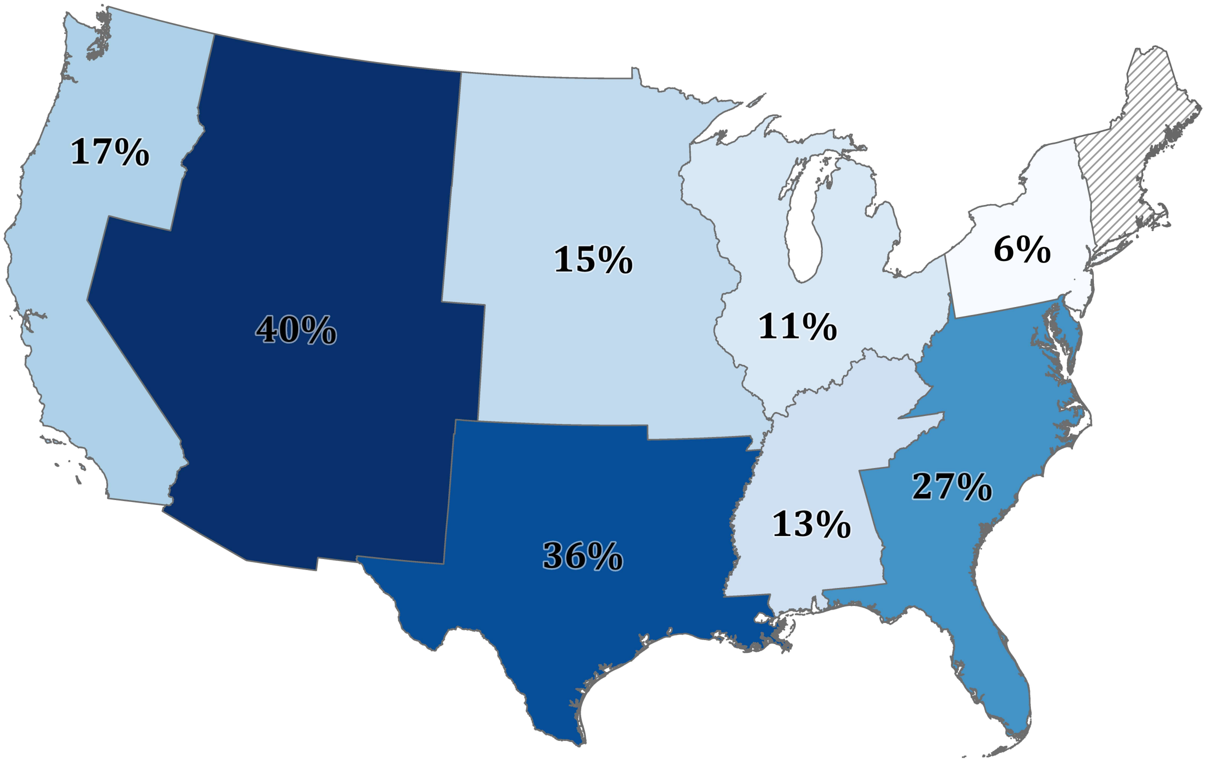


Figure 3.5: Percentage of Single-Family Housing with an HOA, by Census Division and County.



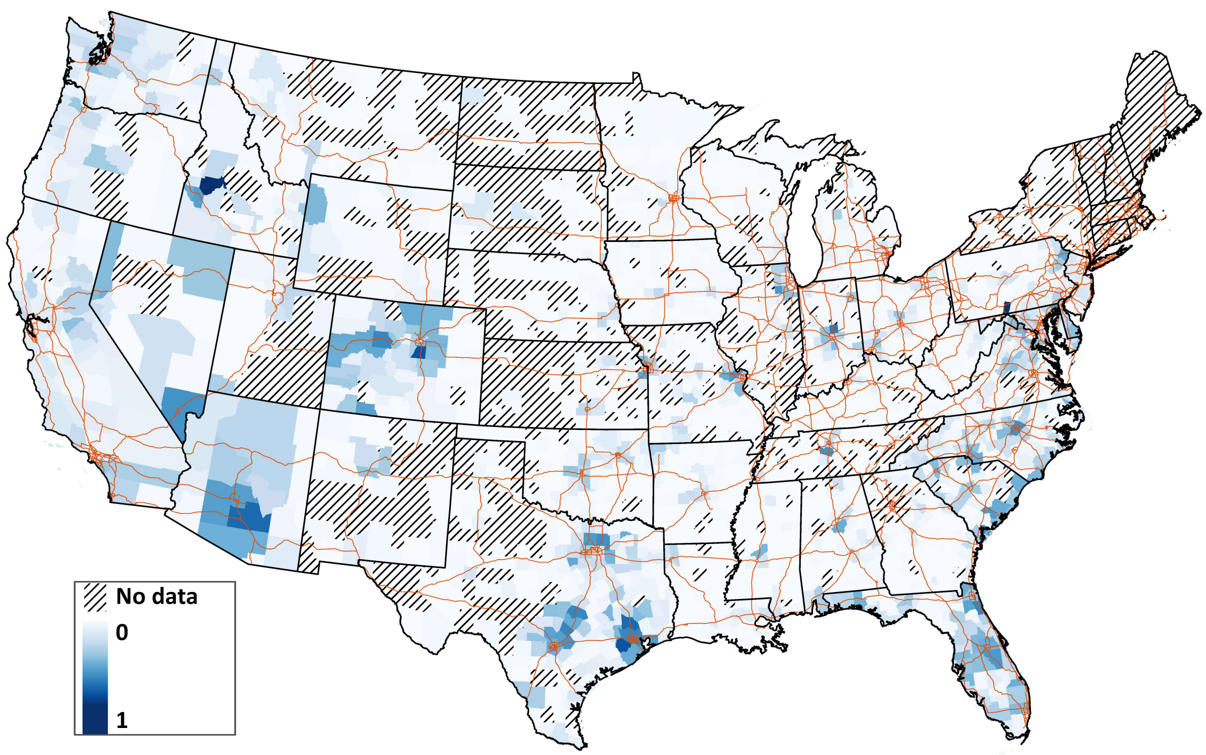


Figure 3.6: Percentage of Single-Family Homes Built in New Subdivisions Since 2000 with an HOA, by Census Division and County.

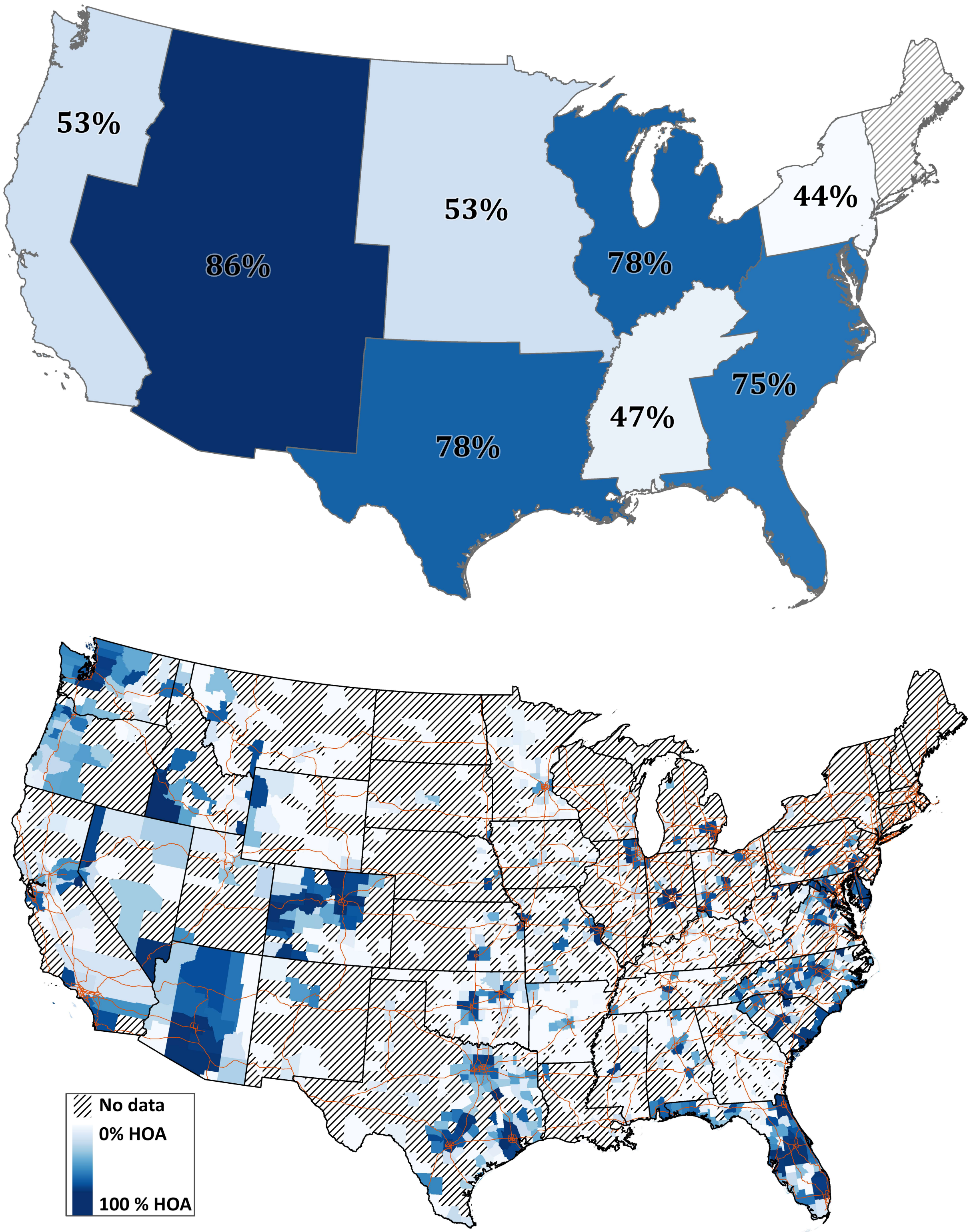
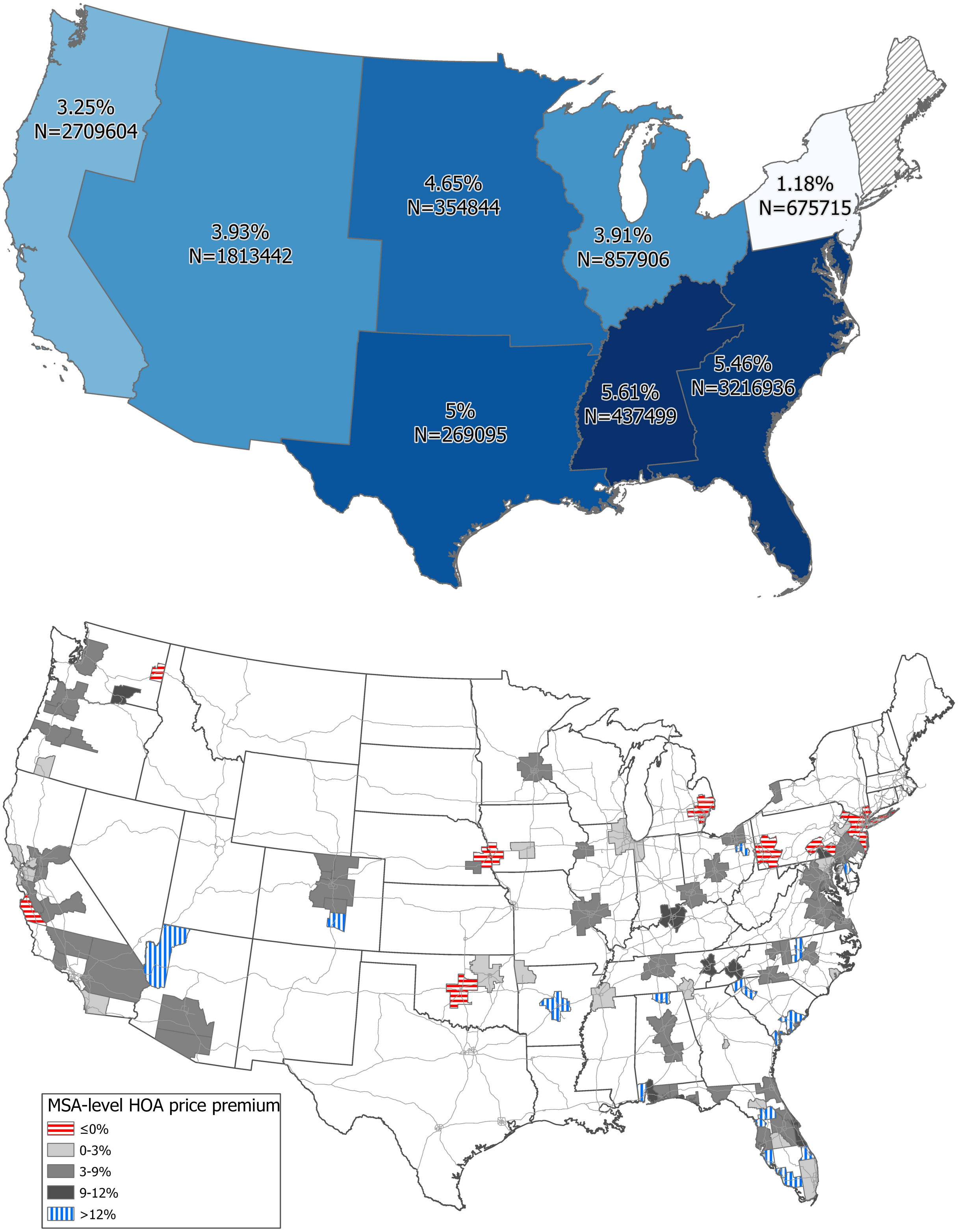


Figure 3.7: HOA Premium, by Census Division and Metropolitan Statistical Areas



Note: Displayed MSAs have at least 10,000 transactions or a statistically significant coefficient at the 99% level.

Bibliography

Acemoglu, D., Autor, D. H. & Lyle, D. (2004), 'Midcentury Women, War, and Wages: The Effect of Female Labor Supply on the Wage Structure at Midcentury', *Journal of Political Economy* **112**(3), 497–551.

Altonji, J. G., Huang, C.-I. & Taber, C. R. (2015), 'Estimating the Cream Skimming Effect of School Choice', *Journal of Political Economy* **123**(2), 266–324.

Angrist, J. D. & Pischke, J.-S. (2008), *Mostly Harmless Econometrics: An Empiricist's Companion*, Princeton University Press.

Barseghyan, L., Clark, D. & Coate, S. (2014), Public school choice: An economic analysis. Unpublished manuscript.

Barton, S. E. & Silverman, C. J. (1994), *Common Interest Communities: Private Governments and the Public Interest*, Institute of Governmental Studies Press, Berkeley, CA.

Benson, S. E. & DeBat, D. (2014), *Escaping Condo Jail*, CreateSpace Independent Publishing Platform.

Black, S. E. (1999), 'Do better schools matter? Parental valuation of elementary education', *The Quarterly Journal of Economics* **114**(2), 577–600.

Brasington, D. M. (2002), 'Edge versus center: Finding common ground in the capitalization debate', *Journal of Urban Economics* **52**(3), 524–541.

Brunner, E. J., Cho, S. W. & Reback, R. (2012), 'Mobility, housing markets, and schools: Estimating the effects of inter-district choice programs', *Journal of Public Economics* **96**(7-8), 604–614.

Brunner, E. & Sonstelie, J. (2003), 'Homeowners, property values, and the political economy of the school voucher', *Journal of Urban Economics* **54**(2), 239–257.

Brunner, E., Sonstelie, J. & Thayer, M. (2001), 'Capitalization and the Voucher: An Analysis of Precinct Returns from California's Proposition 174', *Journal of Urban Economics* **50**(3), 517–536.

Cannaday, R. E. (1994), 'Condominium Covenants: Cats, Yes; Dogs, No', *Journal of Urban Economics* **35**, 71–82.

Cheung, R. (2008), 'The interaction between public and private governments: An empirical analysis', *Journal of Urban Economics* **63**(3), 885–901.

Cheung, R., Cunningham, C. & Meltzer, R. (2014), 'Do homeowners associations mitigate or aggravate negative spillovers from neighboring homeowner distress?', *Journal of Housing Economics* **24**, 75–88.

Cheung, R. & Meltzer, R. (2013), 'Homeowners associations and the demand for local land use regulation', *Journal of Regional Science* **53**(3), 511–534.

Cheung, R. & Meltzer, R. (2014), 'Why and Where Do Homeowners Associations Form', *Cityscape* **16**(3), 69–92.

Chung, I. H. (2015), 'School choice, housing prices, and residential sorting: Empirical evidence from inter-and intra-district choice', *Regional Science and Urban Economics* **52**, 39–49.

Clarke, W. (2016), Optional public school interdistrict open enrollment and house prices. Unpublished manuscript.

Community Associations Institute (2016*a*), 2016 homeowner satisfaction survey, Technical report, Community Associations Institute.

Community Associations Institute (2016*b*), Large-Scale Association Survey Results, 2016, Technical Report June, Community Associations Institute.

CREDO (2012), 'Charter School Performance in Indiana'.

Do, A. Q. & Sirmans, C. F. (1994), 'Residential Property Tax Capitalization: Discount Rate Evidence from California', *National Tax Journal* **47**(2), 341–48.

Education Commission of the States (2015), 'Open Enrollment 50-State Report'.
URL: *http://ecs.force.com/mbdata/mbquest4e?rep=OE1505*

Education Commission of the States (2016), 'Charter Schools 50-State Report'.
URL: *http://ecs.force.com/mbdata/mbquestNB2?rep=CS1501*

Epple, D. N. & Romano, R. (2003), Neighborhood Schools, Choice, and the Distribution of Educational Benefits, *in* 'The Economics of School Choice', National Bureau of Economic Research, pp. 227–286.

Esquivel, A. & Alvayay, J. R. (2014), 'A Guide to Understanding Residential Subdivisions in California', *California Bureau of Real Estate* .

Ferreyra, M. M. (2007), 'Estimating the Effects of Private School Vouchers in Multidistrict Economies', *The American Economic Review* **97**(3), 789–817.

Figlio, D. & Hart, C. M. D. (2014), 'Competitive effects of means-tested school vouchers', *American Economic Journal: Applied Economics* **6**(1), 133–156.

Foldvary, F. (1994), *Public Goods and Private Communities*, Edward Elgar Publishing Company.

Glaeser, E. L. & Gyourko, J. (2005), 'Urban Decline and Durable Housing', *Journal of Political Economy* **113**(2), 345–000.

Groves, J. R. (2008), 'Finding the Missing Premium : An Explanation of Home Values within Residential Community Associations', *Land Economics* **84**(2), 188–208.

Guberman, R. (2004), 'Home is Where the Heart Is', *Legal Affairs* .
URL: *https://www.legalaffairs.org/issues/November-December-2004/feature_guberman_novdec04.msp*

Hastie, T. & Tibshirani, R. (1990), *Generalized Additive Models*, Chapman and Hall/CRC.

Hastings, J. S., Kane, T. J. & Staiger, D. O. (2009), 'Heterogeneous Preferences and the Efficacy of Public School Choice', *NBER Working Paper* **2145**(May), 1–46.

Helsley, B. R. W. & Strange, W. C. (2000), 'Social Interactions and the Institutions of Local Government', *American Economic Review* **90**(5), 1477–1490.

Helsley, R. W. & Strange, W. C. (1998), 'Private government', *Journal of Public Economics* **69**(2), 281–304.

Hilber, C. A. L. (2010), 'New housing supply and the dilution of social capital', *Journal of Urban Economics* **67**(3), 419–437.

Hilber, C. A. L. (2015), 'The Economic Implications of House Price Capitalization: A Synthesis', *Real Estate Economics* **00**(0), 1–39.

Hilber, C. A. L. & Mayer, C. (2009), 'Why do households without children support local public schools? Linking house price capitalization to school spending', *Journal of Urban Economics* **65**(1), 74–90.

Hilber, C. A. L., Rouwendal, J. & Vermeulen, W. (2014), 'Local Economic Conditions and the Nature of New Housing Supply', *SSRN Electronic Journal* .

Hopkins, E. A. (2017), 'The impact of community associations on residential property values', *Housing and Society* **00**(00), 1–11.

Hsieh, C. T. & Urquiola, M. (2006), 'The effects of generalized school choice on achievement and stratification: Evidence from Chile's voucher program', *Journal of Public Economics* **90**(8-9), 1477–1503.

Hughes, W. T. & Turnbull, G. K. (1996), 'Uncertain Neighborhood Effects and Restrictive Covenants', *Journal of Urban Economics* **39**, 160–172.

IDOE (2015), Choice Scholarship Program Annual Report: Participation and Payment Data, Technical Report February, Indiana Department of Education, Office of School Finance.

Kiernan, J. S. (2018), '2018's Property Taxes by State'.

Lopez, S. (2017), 'Github Repository'.
URL: *https://github.com/slopez15*

Lucas, W. (2013), *Neighbors at War: The Creepy Case Against Your Homeowner's Association*, Hogback Publishing, Denver, CO.

Manson, S., Schroeder, J., Van Riper, D. & Ruggles, S. (2017), 'IPUMS National Historical Geographic Information System: Version 12.0 [Database]'.

Mccabe, B. C. (2011), 'Homeowners associations as private governments: What we know, what we don't know, and why it matters', *Public Administration Review* **71**(4), 535–542.

McKenzie, E. (1994), *Privatopia: Homeowner Associations and the Rise of Residential Private Government*, Yale University Press.

Meltzer, R. (2013), 'Do Homeowners Associations Affect Citywide Segregation? Evidence From Florida Municipalities'.

Meltzer, R. & Cheung, R. (2014), 'How are homeowners associations capitalized into property values?', *Regional Science and Urban Economics* **46**, 93–102.

National Conference of State Legislatures (2016), 'School Voucher Laws: State-by-State Comparison'.
URL: *http://www.ncsl.org/research/education/voucher-law-comparison.aspx*

Nechyba, T. (2000), 'Mobility, Targeting, and Private-School Vouchers', *American Economic Review* **96**(1), 222–235.

Nechyba, T. (2003*a*), Introducing School Choice Into Multi-District Public School Systems, *in* 'The Economics of School Choice', The University of; Chicago, pp. 145–194.

Nechyba, T. (2003*b*), 'School finance, spatial income segregation, and the nature of communities', *Journal of Urban Economics* **54**(1), 61–88.

Nelson, R. H. (2005), *Private Neighborhoods and the Tranformation of Local Government*, Urban Institute Press.

Nelson, R. H. (2009), 'The puzzle of local double taxation: Why do private community associations exist?', *Independent Review* **13**(3), 345–365.

Oates, W. E. (1969), 'The Effects of Property Taxes and Local Public Spending on Property Values: An Empirical Study of Tax Capitalization and the Tiebout Hypothesis', *Journal of Political Economy* **77**(6), 957–971.

Postrel, V. (2017), 'Commentary: Digging into Rand Paul's 'trivial' dispute with his neighbor'.
URL: *http://www.chicagotribune.com/news/opinion/commentary/ct-perspec-neighbors-rand-paul-1114-20171113-story.html*

Reback, R. (2005), ‘House prices and the provision of local public services: Capitalization under school choice programs’, *Journal of Urban Economics* **57**(2), 275–301.

Reback, R. (2008), ‘Demand (and supply) in an inter-district public school choice program’, *Economics of Education Review* **27**(4), 402–416.

Rogers, W. H. (2006), ‘A Market for Institutions: Assessing the Impact of Restrictive Covenants on Housing’, *Land Economics* **82**(4), 500–512.

Rogers, W. H. (2010), ‘Measuring the Price Impact of Municipal Incorporation on Homeowner Associations’, **86**(1), 91–116.

Rosen, S. (1974), ‘Hedonic Prices and Implicit Markets: Product Differentiation in Pure Competition’, *Journal of Political Economy* **82**(1), 34–55.

Saiz, A. (2010), ‘The Geographic Determinants of Housing Supply’, *The Quarterly Journal of Economics* **125**(3), 1253–1296.

Sirmans, G. S., Gatzlaff, D. H. & Macpherson, D. a. (2008), ‘The History of Property Tax Capitalization in Real Estate’, *Journal of Real Estate Literature* **16**(3), 327–344.

Speyrer, J. F. (1989), ‘The effect of land-use restrictions on market values of single-family homes in Houston’, *The Journal of Real Estate Finance and Economics* **2**(2), 117–130.

Stadelmann, D. & Billon, S. (2012), ‘Capitalisation of Fiscal Variables and Land Scarcity’, *Urban Studies* **49**(May), 1571–1594.

Tiebout, C. M. (1956), ‘A Pure Theory of Local Expenditures’, *Journal of Political Economy* **64**(5), 416–424.

Wood, S. (2017), *Generalized Additive Models: An Introduction with R*, 2 edn, Chapman and Hall/CRC.

Woodworth, J. L., Raymond, M. E., Chirbas, K., Gonzalez, M., Negassi, Y., Snow, W. & Van Donge, C. (2015), Online charter school study, Technical report, CREDO.

Ziebarth, T. & Bierlein Palmer, L. (2014), The Health of the Public Charter School Movement : A State-By-State Analysis, Technical report, The National Alliance for Public Charter Schools, Washington.

Appendix A

A.1 Clustering Algorithm Used to Attribute HOA Status to Neighborhoods

DBSCAN ("density-based spatial clustering of applications with noise") is an unsupervised clustering algorithm used to find groups of homes that likely share a common HOA status. The benefit of DBSCAN, relative to grouping by subdivision, is that it can be applied to houses where subdivisions are not recorded, where HOAs are not in a subdivision. The disadvantages of DBSCAN relative to grouping by subdivision is that the process is somewhat arbitrarily chosen, it is difficult to explain, and DBSCAN does not include homes in an HOA if, for some reason, they were built at a different time than the other homes. Altogether, it is worth grouping by both subdivision and DBSCAN cluster to [1] include more homes in the neighborhood-level measure of HOA status and [2] use the methods to validate one another.

DBSCAN works by grouping houses that are within x distance of at least n other houses into core clusters. Any additional houses that are within x distance of at least one house in the core cluster are added to the periphery of the cluster. Remaining houses are not placed in a cluster. The distance used is simple Euclidean distance plus a penalty for difference in year built. For this paper, clusters are formed with at least $n = 5$ houses in their core, located within $x = 75$ meters of each other if built in the same year or within $\frac{x}{2} = 37.5$ meters of each other if built one year apart. These parameters were determined by the author to balance [1] confidence that houses in the same cluster really do share the same HOA status and [2] inclusion of as many HOA-members houses as possible, which helps avoid focusing on a small non-representative group. DBSCAN is implemented using the scikit-learn package in Python, which provides full documentation.

Figure A.1 shows the degree to which houses in groups formed using DBSCAN and by subdivision all have, or do not have, HOA mortgage flags. If the groups matched HOA boundaries perfectly and all house-level indicators of HOA status were correct, the histograms would show spikes at 0 percent and 100 percent with nothing in the middle. Table A.1 demonstrates that grouping by subdivision or by DBSCAN results in the same HOA designation for over 99 percent of the 14 million houses covered by both methods.

Figure A.1: Histograms of the Percentage of Homes within a Subdivision or Cluster Having a House-Level HOA Indicator

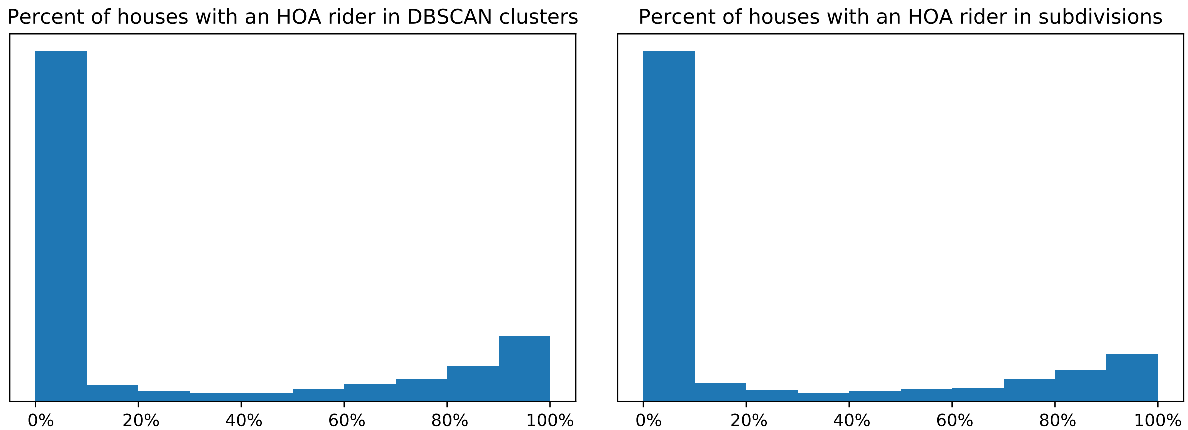


Table A.1: Agreement Between Subdivision and DBSCAN Methods for Assigning Neighborhood HOA Status

		Cluster classification	
		No HOA	HOA
Subdivision classification	No HOA	8,487,937	40,900
	HOA	17,783	5,513,593

A.2 Census Division and MSA-Level HOA Price Premiums

Table A.2: Census Division and MSA-Level Regression Results

Note: Each line presents the coefficient and standard deviation for HOA membership (plus sample size) from a GAM regression of house price on housing characteristics identical to the one found in Table 3.5, Column 4 but restricted to sales from the corresponding Census division or metropolitan statistical area. MSAs with sample size of at least 1,000 sales are reported. The New England division is not reported since ZTRAX does not include mortgage riders from those states.

Place Name	Coefficient	Std deviation	Sample size
CENSUS DIVISIONS			
Middle Atlantic (NJ, NY, PA)	0.0118	0.0011	675,715
East North Central (IL, IN, MI, OH, WI)	0.0391	0.0011	857,906
West North Central (IA, KS, MN, MO, NE, ND, SD)	0.0465	0.0021	354,844
South Atlantic (DE, DC, FL, GA, MD, NC, SC, VA, WV)	0.0546	0.0006	3,216,936
East South Central (AL, KY, MI, TN)	0.0561	0.0020	437,499
West South Central (AR, LA, OK, TX)	0.0500	0.0029	269,095
Mountain (AZ, CO, ID, MT, NV, NM, UT, WY)	0.0393	0.0006	1,813,442
Pacific (AK, CA, HI, OR, WA)	0.0325	0.0005	2,709,604
MSAs			
Cape Coral-Fort Myers, FL	0.2132	0.0047	113,806
Hilton Head Island-Bluffton-Beaufort, SC	0.2053	0.0129	16,436
Show Low, AZ	0.1831	0.0155	4,843
Huntsville, AL	0.1818	0.0276	17,764
Naples-Immokalee-Marco Island, FL	0.1806	0.0069	36,096
Port St. Lucie, FL	0.1779	0.0045	65,954
Greenville-Anderson-Mauldin, SC	0.1779	0.0077	24,099
Ocala, FL	0.1598	0.0065	39,725
Wilmington, NC	0.1566	0.0165	8,392
Boone, NC	0.1562	0.0263	2,882
Lynchburg, VA	0.1487	0.0273	2,051
Little Rock-North Little Rock-Conway, AR	0.1427	0.0109	27,188
Sanford, NC	0.1406	0.0272	2,860
Asheville, NC	0.1392	0.0097	18,510
Kahului-Wailuku-Lahaina, HI	0.1381	0.0372	4,166
Mobile, AL	0.1370	0.0114	12,276
Key West, FL	0.1339	0.0183	7,809
Truckee-Grass Valley, CA	0.1292	0.0110	9,933
Bremerton-Silverdale, WA	0.1283	0.0201	4,981
Pahrump, NV	0.1282	0.0299	3,810
Newport, OR	0.1251	0.0193	4,530
Durham-Chapel Hill, NC	0.1242	0.0055	35,766
Oklahoma City, OK	0.1217	0.0052	97,174
Canton-Massillon, OH	0.1216	0.0115	17,080
Lake Havasu City-Kingman, AZ	0.1188	0.0084	25,656
Ames, IA	0.1173	0.0264	3,201
Wenatchee, WA	0.1139	0.0287	3,168
Winston-Salem, NC	0.1139	0.0070	37,827

Continued on next page

– continued from previous page

Place Name	Coefficient	Std deviation	Sample size
Wichita, KS	0.1131	0.0201	1,989
Payson, AZ	0.1095	0.0181	3,111
Kennewick-Richland, WA	0.1094	0.0094	19,936
Pueblo, CO	0.1087	0.0125	13,094
Salisbury, MD-DE	0.1024	0.0121	8,552
Hickory-Lenoir-Morganton, NC	0.1014	0.0158	13,383
Urban Honolulu, HI	0.1002	0.0152	21,080
Roseburg, OR	0.0998	0.0175	6,208
Dunn, NC	0.0993	0.0177	7,386
Morehead City, NC	0.0992	0.0161	6,139
Akron, OH	0.0985	0.0061	35,913
Palm Bay-Melbourne-Titusville, FL	0.0950	0.0046	67,726
Greensboro-High Point, NC	0.0933	0.0058	47,545
Spartanburg, SC	0.0919	0.0210	3,573
Lufkin, TX	0.0906	0.0321	2,008
Rocky Mount, NC	0.0896	0.0237	3,857
Homosassa Springs, FL	0.0894	0.0102	17,140
Flagstaff, AZ	0.0890	0.0092	7,901
Charleston-North Charleston, SC	0.0886	0.0055	37,009
Augusta-Richmond County, GA-SC	0.0884	0.0133	9,927
Tampa-St. Petersburg-Clearwater, FL	0.0878	0.0020	287,488
Tallahassee, FL	0.0878	0.0069	27,553
Tulsa, OK	0.0874	0.0061	47,954
Buffalo-Cheektowaga-Niagara Falls, NY	0.0873	0.0099	23,746
York-Hanover, PA	0.0853	0.0072	28,807
Lincoln, NE	0.0851	0.0071	25,755
North Port-Sarasota-Bradenton, FL	0.0843	0.0045	86,263
Kill Devil Hills, NC	0.0839	0.0201	5,134
New Bern, NC	0.0825	0.0173	6,452
Bend-Redmond, OR	0.0811	0.0050	29,635
Cincinnati, OH-KY-IN	0.0803	0.0032	108,652
Kapaa, HI	0.0798	0.0537	2,099
Pensacola-Ferry Pass-Brent, FL	0.0789	0.0044	47,230
Houston-The Woodlands-Sugar Land, TX	0.0784	0.0216	3,999
Birmingham-Hoover, AL	0.0771	0.0080	27,167
Fort Collins, CO	0.0741	0.0042	44,852
Knoxville, TN	0.0740	0.0055	47,351
Raleigh, NC	0.0730	0.0029	168,938
Dover, DE	0.0725	0.0228	14,612
Montgomery, AL	0.0723	0.0127	14,856
Lubbock, TX	0.0719	0.0321	2,625
Charlotte-Concord-Gastonia, NC-SC	0.0704	0.0026	223,357
St. Louis, MO-IL	0.0697	0.0030	86,708
Kansas City, MO-KS	0.0695	0.0204	1,467
Tucson, AZ	0.0694	0.0024	102,041
Fresno, CA	0.0693	0.0033	73,993
Greenville, NC	0.0691	0.0135	9,492
Columbus, OH	0.0690	0.0033	88,884
Orlando-Kissimmee-Sanford, FL	0.0686	0.0022	224,913
Cleveland-Elyria, OH	0.0683	0.0040	68,868
Durango, CO	0.0677	0.0187	3,044

Continued on next page

– continued from previous page

Place Name	Coefficient	Std deviation	Sample size
Sebastian-Vero Beach, FL	0.0673	0.0080	19,929
Hot Springs, AR	0.0667	0.0229	5,734
Seattle-Tacoma-Bellevue, WA	0.0667	0.0019	269,475
Fayetteville-Springdale-Rogers, AR-MO	0.0666	0.0086	24,035
Columbus, GA-AL	0.0663	0.0136	5,076
Toledo, OH	0.0648	0.0097	22,860
Minneapolis-St. Paul-Bloomington, MN-WI	0.0646	0.0047	55,767
Iowa City, IA	0.0645	0.0135	11,582
Boulder, CO	0.0644	0.0047	31,806
Jacksonville, FL	0.0641	0.0027	141,650
Baltimore-Columbia-Towson, MD	0.0640	0.0044	28,703
Warner Robins, GA	0.0633	0.0137	13,056
Roanoke, VA	0.0633	0.0122	11,386
Chattanooga, TN-GA	0.0621	0.0100	27,629
Portland-Vancouver-Hillsboro, OR-WA	0.0615	0.0022	138,205
Salem, OR	0.0603	0.0047	33,838
Edwards, CO	0.0600	0.0190	5,315
Phoenix-Mesa-Scottsdale, AZ	0.0599	0.0010	637,585
Nashville-Davidson–Murfreesboro–Franklin, TN	0.0595	0.0037	110,001
Crestview-Fort Walton Beach-Destin, FL	0.0595	0.0060	37,758
Montrose, CO	0.0590	0.0168	2,196
Louisville/Jefferson County, KY-IN	0.0589	0.0039	54,468
Breckenridge, CO	0.0583	0.0181	3,501
Deltona-Daytona Beach-Ormond Beach, FL	0.0581	0.0036	66,380
Bakersfield, CA	0.0573	0.0037	86,593
Burlington, NC	0.0562	0.0145	9,504
New Orleans-Metairie, LA	0.0557	0.0093	7,988
Olympia-Tumwater, WA	0.0542	0.0196	8,856
Daphne-Fairhope-Foley, AL	0.0534	0.0093	14,450
Ottawa-Peru, IL	0.0528	0.0393	1,965
Greeley, CO	0.0508	0.0042	43,295
Atlantic City-Hammonton, NJ	0.0504	0.0094	14,647
San Jose-Sunnyvale-Santa Clara, CA	0.0492	0.0040	83,769
Erie, PA	0.0491	0.0182	8,557
Gettysburg, PA	0.0490	0.0154	3,819
Reading, PA	0.0488	0.0067	17,414
Oxnard-Thousand Oaks-Ventura, CA	0.0488	0.0035	54,220
Riverside-San Bernardino-Ontario, CA	0.0476	0.0009	486,965
Denver-Aurora-Lakewood, CO	0.0472	0.0013	289,605
Stillwater, OK	0.0461	0.0277	3,199
Philadelphia-Camden-Wilmington, PA-NJ-DE-MD	0.0454	0.0018	174,481
Eugene, OR	0.0445	0.0063	24,419
Davenport-Moline-Rock Island, IA-IL	0.0445	0.0115	12,832
Des Moines-West Des Moines, IA	0.0441	0.0065	47,586
Colorado Springs, CO	0.0436	0.0021	97,904
Madera, CA	0.0435	0.0143	9,187
Spokane-Spokane Valley, WA	0.0413	0.0077	25,636
State College, PA	0.0408	0.0101	9,785
Dallas-Fort Worth-Arlington, TX	0.0404	0.0198	4,359

Continued on next page

– continued from previous page

Place Name	Coefficient	Std deviation	Sample size
Jacksonville, NC	0.0380	0.0064	25,680
Reno, NV	0.0369	0.0032	55,746
Lakeland-Winter Haven, FL	0.0354	0.0050	76,245
Panama City, FL	0.0332	0.0084	21,655
Virginia Beach-Norfolk-Newport News, VA-NC	0.0330	0.0036	69,989
Steamboat Springs, CO	0.0323	0.0309	1,729
Richmond, VA	0.0322	0.0058	29,915
Carson City, NV	0.0322	0.0127	4,357
Allentown-Bethlehem-Easton, PA-NJ	0.0317	0.0061	27,392
Memphis, TN-MS-AR	0.0294	0.0040	82,043
Chico, CA	0.0293	0.0093	15,246
Sacramento–Roseville–Arden-Arcade, CA	0.0285	0.0016	176,576
Chicago-Naperville-Elgin, IL-IN-WI	0.0283	0.0016	254,648
Dayton, OH	0.0281	0.0127	11,669
Napa, CA	0.0262	0.0097	7,405
Redding, CA	0.0258	0.0159	7,700
Stockton-Lodi, CA	0.0249	0.0024	69,010
Washington-Arlington-Alexandria, DC-VA-MD-WV	0.0242	0.0016	167,249
Gainesville, FL	0.0241	0.0102	13,323
Santa Rosa, CA	0.0225	0.0061	31,914
Albany, OR	0.0217	0.0081	8,550
Pittsburgh, PA	0.0209	0.0053	65,436
Fernley, NV	0.0208	0.0112	9,523
Santa Maria-Santa Barbara, CA	0.0204	0.0054	18,919
San Diego-Carlsbad, CA	0.0201	0.0014	190,429
Columbia, SC	0.0197	0.0223	5,594
Miami-Fort Lauderdale-West Palm Beach, FL	0.0184	0.0018	266,536
Charlottesville, VA	0.0183	0.0114	7,486
Gardnerville Ranchos, NV	0.0171	0.0112	6,076
Lebanon, PA	0.0163	0.0081	6,434
Los Angeles-Long Beach-Anaheim, CA	0.0153	0.0013	343,587
Corvallis, OR	0.0147	0.0095	6,434
East Stroudsburg, PA	0.0144	0.0118	8,437
San Francisco-Oakland-Hayward, CA	0.0136	0.0023	157,246
Elko, NV	0.0126	0.0319	3,353
San Luis Obispo-Paso Robles-Arroyo Grande, CA	0.0121	0.0058	20,478
Harrisburg-Carlisle, PA	0.0103	0.0059	32,645
Vallejo-Fairfield, CA	0.0102	0.0045	44,640
Las Vegas-Henderson-Paradise, NV	0.0100	0.0013	339,585
Medford, OR	0.0089	0.0066	19,714
Hagerstown-Martinsburg, MD-WV	0.0088	0.0138	4,411
Sierra Vista-Douglas, AZ	0.0080	0.0226	3,435
Omaha-Council Bluffs, NE-IA	0.0047	0.0066	32,486
Mount Vernon-Anacortes, WA	0.0042	0.0270	3,892
Trenton, NJ	-0.0017	0.0075	10,502
The Villages, FL	-0.0043	0.0030	25,664
Glenwood Springs, CO	-0.0090	0.0124	7,952
Santa Cruz-Watsonville, CA	-0.0100	0.0098	10,127

Continued on next page

– continued from previous page

Place Name	Coefficient	Std deviation	Sample size
New York-Newark-Jersey City, NY-NJ-PA	-0.0131	0.0021	191,616
Lancaster, PA	-0.0147	0.0050	32,650
Winchester, VA-WV	-0.0163	0.0090	5,743
Salinas, CA	-0.0243	0.0059	17,683
Detroit-Warren-Dearborn, MI	-0.0266	0.0054	31,868
Grants Pass, OR	-0.0273	0.0197	5,299
Bellingham, WA	-0.0751	0.0263	3,340
Albemarle, NC	-0.1371	0.0439	1,904
Ocean City, NJ	-0.1394	0.0212	6,347