Reading Note: Do Better Schools Matter? Parental Valuation of Elementary Education

Bob Lin. Mar. 2022

The paper adopts a regression discontinuity design to study how much parents value school quality by comparing prices of houses close to school attendance district boundaries in Massachusetts. Many school reforms aim to enhance school quality, which is measured by test scores and many previous studies. However, the value of better schools is difficult to measure directly. The causality between school quality and students' future earnings is still inconclusive even though numerous researchers have worked on it. One alternative way to estimate the value of better schools is to study how much people are willing to pay to get into better schools. Past literature estimates the effect of better schools by examining how much higher the house prices are in the districts with better schools. Unfortunately, many unobserved characteristics may also differ in the districts with better schools. Hence, although previous studies have tried to control observed characteristics of houses and districts, the estimation may still be imprecise. The paper circumvents the problem by comparing the prices of houses that were geographically close but on the opposite sides of elementary school attendance district boundaries. The comparison of proximate houses reduces the possibility of potential differences in unobserved characteristics between them. Then, the main difference between the houses on the opposite sides of boundaries is the elementary schools to which the children in the houses go. As a result, the paper can estimate the effect of school quality on house prices more precisely. The paper finds that a 5% increase in test scores in elementary schools corresponds to a 2.1% increase in the price of houses, or $3,948 at the mean house price of $188,000. More importantly, such estimation is approximately half as the estimation obtained by directly comparing the average of house prices of different districts with many observed characteristics controlled, suggesting the overestimation of the value of better schools in the previous research.

The paper estimates two main regression models. The first one is the same as previous literature, regressing house prices on house characteristics, district characteristics, and test scores of the elementary schools in the district. The second one replaces the terms of district characteristics with boundary dummy variables, indicating which boundary is nearest to the house. The second model can be also viewed as a fixed effect model, with each boundary dummy variable controlling the common characteristics among the houses within a certain distance to that boundary. Model 2 also has three variations to test its robustness by using the sample of houses within 0.35, 0.20, and 0.15 miles from the nearest boundary respectively. The full sample is 22,679 single-family houses in some counties in Massachusetts that were transacted between 1993 to 1995 (so that the paper can obtain house prices by real transaction data provided by a private company). House characteristics include numbers of bedrooms and bathrooms, age of the building, and sizes of land and house. District characteristics include two parts: neighborhood characteristics and school characteristics. Neighborhood characteristics include demographic characteristics from census block data, such as ethnicity and education backgrounds. School characteristics try to capture financial inputs into schools, such as average per-pupil expenditure and pupil/teacher ratio. The school quality is measured by test scores from the fourth grade Massachusetts Educational Assessment Program. Hence, after controlling some school characteristics, test scores reflect only less observable school characteristics on school quality such as the quality of teacher and peer. The coefficients of test scores in the two motels are the main interest of this paper. For model 1, the magnitude of the test scores coefficient means that a 5% increase in the average elementary school test scores corresponds to a 4.9% increase in the house price. For model 2, the coefficients of three variations do not have a large difference and their magnitudes are roughly half as the coefficient in model 1. Also, the paper runs a variation of model 1 by restricting the sample to those houses within 0.15 miles and compares it with the variation of model 2 with the same sample. Still, the paper finds that the effect of test scores on house prices estimated in model 2 is approximately half as the estimation in model 1.

However, were the houses on the opposite sides of boundaries really similar except for school attendance? One possibility is that the boundaries of school attendance districts did reflect the actual difference in neighborhoods. The paper runs a variation of model 2 by excluding those houses that were close to boundaries which were railroad tracks, highways, and major streets, which might cause differences in characteristics across the boundaries. The paper finds that the coefficient of test scores only drops slightly after such exclusion. Another possibility is that better schools tended to be in better neighborhoods. If that is the case, the houses on the side of better schools would tend to be in better neighborhoods as they were further away from the boundaries (close to schools with better quality). Similarly, the houses on the side of worse schools would tend to be in better neighborhoods as they were close to the boundaries (close to the schools with better quality on the other side). The paper uses the houses that were 0.4 to 0.8 miles away from the boundary within the same attendance district to study the potential effect of better neighborhoods. The paper creates artificial dummies that denote if houses were closer to better schools and hence in better neighborhoods (for houses on the side of better schools, the dummy for those houses between 0.6 and 0.8 miles from the boundary is set to one; for houses on the side of worse schools, the dummy for those houses between 0.4 and 0.6 miles from the boundary is set to one). However, the paper finds that the coefficient of the dummy is small and not significant. The paper also adds controls of neighborhood characteristics to model 2 and finds little change in the coefficient of test scores. Finally, one possibility is that the school quality is correlated with unobservable differences in house quality. For example, residents who cared more about schools might also take better care of their houses. The paper regresses those observable house characteristics on test scores and the boundary dummies and finds no significant effect of test scores. The above results show that, for those houses close to the boundaries, their characteristics did not have much difference.

The paper does a great job of eliminating potential omitted variable bias. One potential problem that is worth discussing is that the paper uses real transaction data to get the prices of the houses. However, can these transacted houses represent the whole house in the study areas? If certain types of houses were more likely to be transacted, then the sample the paper uses might have bias. Also, although the paper has paid large attention to the justify the houses close to the boundaries were similar empirically, the paper can add some explanation on why such boundaries did not cause large differences in characteristics of houses between the opposite sides. After all, the schools were built in certain areas and the boundaries were set for reasons, such as population density and ethnicity, age, and economic status of residents. Those reasons may closely relate to the characteristics interest in the paper. Hence, by answering why the original reasons did not play an important role or how those factors help neighborhoods become similar, the paper can provide a more comprehensive analysis of the question.