DO BETTER SCHOOLS MATTER?

PARENTAL VALUATION OF ELEMENTARY EDUCATION

Reading Note 3, Yuzhe Wang

Researchers are interested in finding the relationship between school quality and school value. However, it is difficult to obtain conclusive results by establishing a direct link between such school quality and value as test scores and student earnings. To solve the problem, earlier studies start to estimate the school value by calculating the house prices in areas with schools. However, these studies face the problem of omitted variable bias because of the unobserved house or neighborhood characteristics. So, their results overestimate the value of better schools.

To deal with the problem of omitted variable bias and isolate the relationship between test scores (school quality) and house price (school value), the author limits the sample to the houses in the same cities and near the attendance district boundaries and adds boundary dummies in the model. More specifically, the full sample consists of single-family residences in Massachusetts within districts with more than one elementary school that overlap grades and districts without the system of intradistrict school choices. In the model set, compared to the traditional model by regressing the house prices on test scores, house characteristics, neighborhood, and school district characteristics, the author replaces the neighborhood and school district characteristics with the set of boundary dummies that decide whether houses are on either side of an attendance district boundary. Because the neighborhood and school district characteristics are similar to the houses near the boundaries, the boundary dummies can explain the unobserved characteristics shared by these houses. Based on the selected sample data and the new model, the omitted variable problems can be solved.

The results of the traditional model and the newly constructed model confirm the expectation that the results of the earlier studies overestimate the value of better schools. More specifically, when the author controls some characteristics of houses, neighborhoods, and school districts, a 5 percent increase in the average elementary school test score increases the house price by 4.9 percent. However, when the author includes the boundary fixed effects (the set of boundary dummies) in the model and restricts the sample to houses that are closer to the attendance district boundary, the effects of the test score on the house price are much lower. When the sample of houses is limited within 0.35 miles from the boundary, a 5 percent increase in the average elementary school test score increases the house price by 2.3 percent. Furthermore, when the sample is located within 0.2 miles and 0.15 miles from the boundary, the results are similar to the one within 0.35 miles. To prove that the weaker effect of school quality is not caused by the sample selection, the author uses the subsample of houses at 0.15 miles from the boundary to run the traditional model. The results show no difference between the traditional models with full sample and the subsample, indicating that the sample size doesn't lead to the change of the effect of test score. In addition, the author also confirms the expectation that houses closer to the boundary are more similar in house and neighborhood characteristics by showing the decreasing difference in means of the characteristics among houses within a smaller distance from the boundary. Conclusively, the results demonstrate that when the neighborhood characteristics can be fully controlled, the increased value of the better school quality is around half the effect of simple hedonic house price regression.

To test the sensitivity of the results, the author checks three worries by running several specification tests. 1) The first concern is that the houses on either side of the attendance district boundaries are not really the same neighborhoods and they can be quite different instead, which may lead to the bias of estimation. To check this concern, the author runs the regression by excluding the boundaries that are major streets, railroads, and highways. The slight decrease in the coefficient on test scores shows that this concern is not a problem. 2) The second concern is that better schools are in better neighborhoods, which may lead to the fact that it is not the difference of schools but the progression of neighborhoods that affects the house price. The

author tests the concern by running a true regression and a control regression on a "hi" dummy which indicates if the house is on the better side of the boundary (higher average test scores). The true regression uses the sample of houses that are within 0.2 miles of the boundary. The control regression uses the pseudo-control group of houses that are between 0.4 and 0.8 miles of the boundary. If it is the progression of neighborhoods affects the results, there should be no difference in results between the true and control models. However, the significance of the result in the true regression and insignificance of the result in the control regression shows that this concern is not a problem. In addition, when the author adds some neighborhoods characteristics in the model, the result still doesn't change much, which further indicates that the results are not pick up the progression in neighborhoods. 3) The third concern is that the results may contain some unobservable differences in house quality. The author tests the concern by regressing some observable house characteristics (lot size and internal square footage) on the test scores and boundary dummies. The results show no significant effects indicating that the house characteristics do not vary on the boundaries. Another check is by regressing the house price on the interactions of the test score and a dummy indicating whether it is a one- or two-bedroom house and the interaction of the test score and a dummy indicating whether it is a three- or more bedroom house. The results show a significant effect of the interaction of the test score and three- or more bedroom dummy and insignificant effect of the interaction of the test score and one- and two-bedroom dummy, indicating the results are still caused by the difference in schools and not just to the quality of the houses. Therefore, these two checks show that the third concern is still not a problem.

In conclusion, the author solves omitted variable bias by limiting the sample of houses near the attendance boundary and adding the set of boundary dummies in the regression model. The further checks of three concerns also indicate the robustness of the results.