DEPARTMENT OF ECONOMICS
UNIVERSITY OF CYPRUS

HOUSE PRICES AND SCHOOL QUALITY
The Impact of Score and Non-score Components of Contextual Value-Added

Sofia Andreou and Panos Pashardes

Discussion Paper 2010-05

P.O. Box 20537, 1678 Nicosia, CYPRUS Tel.: ++357-22893700, Fax: ++357-22895028
Web site: http://www.econ.ucy.ac.cy

House Prices and School Quality:

The Impact of Score and Non-score Components of Contextual Value Added in UK

Sofia N. Andreou[*]
Department of Economics
University of Cyprus

Panos Pashardes
Department of Economics
University of Cyprus

New Version -May 2011

Abstract

This paper investigates how the score and non-score components of the newly introduced Contextual Value Added (CVA) indicator of school quality affect house prices in the catchment area of primary and secondary schools in England. The empirical analysis, based on data drawn from UK data sources, shows that the score component of CVA is positively associated with house prices at both primary and secondary levels of education; while its non-score component has a negative significant association with house prices in the analysis of secondary school data, but no significant association in the corresponding analysis of primary school data. Furthermore, the willingness of households to pay more for a high quality secondary education appears to depend entirely on the final score and is influenced by school comparison within (rather than between) Local Authorities. For primary schools, however, the non-score component of CVA also adds to house prices. This appears to be more so between Local Authorities, highlighting the 'public good' nature of this school quality indicator.

JEL: *R21, I29*

Keywords: *School Quality, Hedonic Regression, House Prices*

---

[*] Corresponding author: Department of Economics and Economics Research Center, University of Cyprus P.O. Box 20537, Nicosia 1678, Cyprus – E-mail: andreou.sofia@ucy.ac.cy;
Tel: +357-22 893675; Fax: +357-22 895027.

1. Introduction

This paper investigates how the score and non-score components of the newly introduced Contextual Value Added (CVA) indicator of school quality is associated with house prices in the catchment area of primary and secondary schools in England. The score component includes the pupils' final academic achievement, while the non-score component includes prior academic achievement and several pupil-specific characteristics generally thought to affect schooling performance (age, language, special needs, ethnicity, income etc). Breaking down the relationship between CVA and house prices to show separately the effect of score and non-score components is important because these two components can: (i) affect house prices in opposite direction, thereby obscuring the overall effect of CVA; and (ii) convey information about private and social preferences that have distinct policy implications. This study is the first to explore the relationship between house prices and CVA. Furthermore, it investigates this relationship at both primary and secondary education level and applies semi-parametric analysis - in addition to the usual parametric estimation. The dataset used cover regions throughout England during 2008 and is constructed by combining information drawn from three independent data sources.

The capitalization of state school quality to house prices has been the object of a large body of literature, especially in the US (e.g. Brasington, 2000, 2002; Haurin and Brasington, 2006, 2009; Black, 1999; Barrow, 2002; Barrow and Rouse, 2004; Downes and Zabel, 2002; Clapp et al, 2008; Kane et al, 2006). In the UK the issue has received less attention, with only a small number of studies available to date (Gibbons and Machin, 2003, 2006, 2008; Cheshire and Sheppard, 2004; Rosenthal, 2003; Leech and Campos, 2003); this probably because the locality of individual households is not available in the UK data due to confidentiality.

Using a hedonic approach (Rosen, 1974) most authors estimate house prices as a function of measures reflecting the quality of the school which the occupants of the house have access to, along with other house attributes such as the number of rooms, size and type. Although various measures are used to capture school quality - including expenditure per pupil (Downes and Zabel, 2002) and pupil/teacher ratio (Brasington, 1999) – reading scores, proficiency test scores and other measures emphasising final academic achievement are the most commonly employed (Gibbons and Machin, 2003, 2008; Haurin and Brasington, 2006; Black, 1999; Rosenthal, 2003). These measures are consistently found to be capitalized into house prices, indicating the willingness of consumers to pay for better quality education, a point some studies also try to justify on theoretical grounds using a Tiebout-type approach (e.g. Barrow, 2002; Hoxby, 2000).

Following Black (1999), investigators have become particularly concerned about non-school factors contaminating the relationship between house prices and school outcomes measured by test score indicators. For example, ignoring neighbourhood deprivation characteristics (such as crime, poverty and unemployment) can exaggerate the positive relationship between house prices and high school scores. Local authority policies (property taxation, provision of public goods etc) can also interfere with the same relationship - Black, 1999). This concern about over-emphasising the importance of test score indicators on house prices echoes criticism by education economists that these indicators are inappropriate measures of school quality because they reflect not only on school attributes but also on individual and family characteristics and other exogenous variables, including the socioeconomic background of the pupil. A measure capturing the distinct contribution of school to pupil's academic progress is argued to be a more appropriate reflection of school performance (Downes, 2007; Hanushek and Taylor, 1990; Mayer, 1997; Hanushek 1992; Summers and Wolfe, 1977). This argument soon found its way to house price regressions, with several authors asking whether VA or test score should be used as measures of school quality in hedonic analysis of house prices. The empirical evidence so far appears to be controversial. Gibbons et al (2008) using UK data find that simple VA and prior score indicators both have a positive and significant effect on house prices. In contrast, Downes and Zabel (2002), using data from the Chicago metropolitan area find that only final score is significant; a result supported by Brasington (1999). Furthermore, Brasington and Haurin (2006) find that while the effect of VA on house prices is positive when used on its own, it becomes negative when score is also included in the hedonic equation.

In principle, CVA indicators discriminate in favour of schools operating under conditions non-conducive to learning, such as pupils with poor socio-economic background, ethically heterogeneous classes, poor/interrupted attendance etc. As such, a CVA index can guide the so called 'pupil premium' funding program proposed by the Conservatives and Liberal Democrats with the aim of narrowing the achievement gap between rich and poor, by attaching greater weight to schools with pupils from disadvantaged backgrounds. The need for using CVA type indicators to help disadvantaged schools through funding discrimination in their favour is also evident in the US, where the grant program of

President Barack Obama's introduced in response to the 2008 economic crisis provides $100 billion for schools, while asking federal officials to focus their proposals, among others, on 'turning around low-performing schools'[1].

The CVA indicator used in the empirical analysis of this paper has been recently introduced by the Department for Children, Schools and Families in England and adjusts the final score achieved by pupils to take account of limitations imposed on their school performance by low prior achievement and other pupil-specific characteristics reflecting on disadvantaged socioeconomic background (Appendix). The fact that CVA combines final score with non-components, mainly reflecting the extent to which a school operates in a deprived socioeconomic environment, can obscure its effect on house prices. As said earlier these two components of CVA can affect house prices in opposite direction: houses in the catchment area of schools with higher final score are higher in price; whereas socioeconomic deprivation characteristics decrease the prices of houses in the affected area. Thus, by separating the effect of final score and non-score components on house prices one can highlight how a CVA indicator compromises private (household) preferences for high academic achievement (final score) with the social preferences for discriminating in favour of schools with high non-score (deprivation) indicators. For instance, a large negative effect of the non-score component on house prices can be interpreted as an indication that low school performance is largely due to the final score being eroded by a disadvantaged background, pointing to the need for greater policy intervention.

Another novelty of the empirical analysis in this paper is the use of non-parametric techniques to explore non-linearity in the relationship between school performance indicators and house prices. This is useful because school performance is measured by arbitrarily normalised indices and is customary in empirical application to (re)normalised them to measure standard deviations from the mean. Therefore, the use of square and cubic terms of these indices to explore higher order effects on house price in hedonic regression is meaningless.[2] Interestingly, this point and, in general, the presence of non-linear and non-monotonic effects of school quality indicators on house prices has not received adequate attention in the literature.

---

[1] Studies focusing on education spending in the US and its distribution across communities include Fernandez and Rogerson (1996, 1998) and Chay et al (2005).

[2] While one can avoid the sign problem by dividing by the standard deviation and not differencing from the mean, the empirical estimates will still depend on the units of measurement of the school performance.

The data used for the empirical analysis are drawn from three UK sources: (i) individual house prices collected from the electronic site "Up my Street"; (ii) school quality indicators from the primary and secondary performance tables, available from the Department for Children, Schools and Families; and (iii) deprivations indices and other neighbourhood characteristics from the Office of National Statistics of UK. The data on school quality include a CVA and final score indicators. More details about the data are given in Appendix.

The paper has the following structure. Section 2 describes the methodology followed in order to estimate the distinct (marginal) contribution of the various groups of variables entering a broadly defined CVA indicator of school quality. Section 3 briefly describes the data and presents the estimates obtained from semi-parametric and parametric empirical analysis. Section 4 concludes the paper.

2. Modelling the effect of school quality on house prices

In this section we deliberate on the components of a CVA indicator with a view to modelling their effect on house prices in a way that facilitates the interpretation and highlights the policy implications of results obtained from empirical application.

We break down the variables affecting school performance which are exogenous to the school into: (i) pupil-specific (ability and family background), denoted by Z; and (ii) neighbourhood-specific (crime, poverty, environment, ethnic heterogeneity etc), denoted by Y. In this context CVA, denoted by V, can be defined as the expected final score X achieved at given values of Z and Y, i.e. $V = E(X|Z, Y)$. When one wishes to focus on progress during a particular period of school attendance, e.g. secondary education, CVA can be also conditioned on prior achievement, denoted by A and defined as the final score achieved prior to the period over which CVA is measured. In this case the CVA can be written as $V = E\{(X - A)|Z, Y\}$ or, more generally, $V = E(X|A, Z, Y)$.[3]

The fact that CVA is a composite indicator of school performance raises the question how the various components of this indicator reflect on household perception of school quality and, thereby, on house prices. To examine this question we consider the following hedonic

[3] This definition of VA comes close to what the Department of Education, Children and Families in the UK terms 'contextual' VA, used in the empirical analysis below.

equation for the cross-section analysis of the (log) price P of house $i = 1, \ldots . S$, in school catchment area $s = 1, \ldots S$,

$$P_{si} = a + \beta V_s + \Sigma_k \gamma_k Q_{ki} + \Sigma_m \varepsilon_m Y_{ms} + u_{si}, \tag{1}$$

where: $Q_{ki}, k = 1, \ldots ., K$ is the vector of house-specific variables (size, type etc) and $Y_{ms}, m = 1, \ldots ., M$ the vector of neighbourhood-specific variables affecting house prices; $a, \beta, \gamma_k, all\ k = 1, \ldots ., K$, and $\varepsilon_m, all\ m = 1, \ldots ., M$ are parameters; and u_{si} is a randomly distributed error.

To keep matters simple we consider the effect of various components of CVA on house prices assuming that V_s in equation (1) is defined as the final score achieved by the school, linearly modified to account for prior achievement and pupil- and neighbourhood-specific factors affecting school performance,

$$V_s = X_s + bA_s + \Sigma_j d_j Z_{js} + \Sigma_m e_m Y_{ms}, \tag{2}$$

where $d_j, all\ j = 1, \ldots, J$, and $e_m\ all\ m = 1, \ldots, M$ are some known parameters.

Replacing (2) in (1) we obtain the *reduced* form hedonic equation

$$P_{is} = a_i + \beta X_s + \theta A_s + \Sigma_j \delta_j \mathrm{Z_{js}} + \Sigma_m \varphi_m Y_{ms} + \Sigma_k \gamma_k Q_{ki} + u_{is}, \tag{3}$$

where $\theta = \beta b$ shows the effect of A_s on price; $\delta_1 = \beta d_1, \delta_2 = \beta d_2, \ldots, \delta_J = \beta d_j$ the effect of variables in the vector Z_{js}; and $\varphi_1 = \beta e_1 + \varepsilon_1$, $\varphi_2 = \beta e_2 + \varepsilon_2 = \beta d_2$, ... , $\varphi_M = \beta e_M + \varepsilon_M \delta_M$ the effect of variables in the vector Y_{ms}.

As said in the introduction, CVA indicators discriminate in favour of schools operating under disadvantageous conditions (such as low prior achievement, poor socio-economic background, ethically heterogeneous classes and poor attendance), effectively awarding higher marks to schools that achieve a given final score in circumstances non-conducive to learning. In the context of equation (3) household aversion to such circumstances can be estimated and contrasted with preference for final score and other desirable components of the CVA. More specifically, the parameter β in equation (3) should be positive, indicating the willingness of households to pay for high final score, a conjecture strongly supported by empirical evidence in the literature. In contrast, the parameter $\theta = \beta b$ is likely to be negative: reaching a given final score starting with a high prior achievement represents poor school performance (low value added, i.e. $b < 0$).

The effect of variables other than X_s and A_s in (3), however, is unclear and will depend on how they are incorporated in the construction of the CVA indicator. For instance, the effects of pupil-specific characteristics will be negative or positive, depending on whether the variables in the Z_{js} vector increase or decrease with learning capacity. The effect of neighbourhood-specific variables in the vector Y_{ms}, is also ambiguous: assuming that these variables measure deprivation the parameters e_m in (2) will be positive (achieving a given final score in deprived neighbourhoods increases VA); whereas the parameters ε_m will be negative (neighbourhood deprivation decreases house prices). Therefore, the effects of neighbourhood characteristics, $\varphi_m = \beta e_{\mathrm{m}} + \varepsilon_{\mathrm{m}}$, which are obtained from estimating (3) may be positive or negative, depending on which of the two components - the direct effect on house prices ε_{m} or the indirect effect through VA βe_{m} - dominates.

The discussion above assumes that one knows how the contextual value is constructed, how it can be decomposed and how its individual elements can be used as variables in the house price equation. In practice, CVA is likely to be published in the form of an index representing the outcome of complex quantitative and qualitative manipulations of final and prior score and other variables, making it impossible to identify the component effects of CVA on house prices, by estimating a reduced form equation like (3). Indeed, this is the case with the English data used in the empirical analysis in this paper. This limitation necessitates modification of the theoretical analysis described above as follows.

Recall that the focus of investigation in this paper is to compare the effect of final score with that of other components of the CVA indicator. We therefore define CVA as an unknown function of final score, prior achievement and a range of pupil-specific characteristics.[4] We denote this by $V_s = v(X_s, A_s, Z_{1s} \dots . Z_{Js})$. To separate the effect of X_s from that of A_s and $Z_{1s} \dots . Z_{Js}$ we project X_s on V_s to obtain $V_s^* = E(V_s|X_s)$, i.e. make the CVA indicator orthogonal to the final score[5]. Then, estimating the house price equation

$$P_{si} = a + \mu X_s + \rho V_s^* + \Sigma_k \gamma_k Q_{ki} + \Sigma_m \varepsilon_m Y_{ms} + u_{si}, \tag{4}$$

the parameter μ captures the effect of final score X_s on house prices, while ρ captures the effect of V_s^*, i.e. the information contained in CVA other than final score. This additional information comes from prior achievement A_s and pupil-specific characteristics $Z_{1s} \dots . Z_{Js}$.

[4] The contextual value added used in our analysis does not take into account the impact of neighbourhood characteristics on school performance.

[5] The non-score component (V*) is defined as the residuals from regressing CVA on its score component.

In the empirical analysis that follows we use equations (1) and (4) for primary and secondary education in England to: (i) estimate the relationships between CVA and house prices; and (ii) find how this relationship is shaped by each of the two components of CVA, score and non-score, as these are defined above.

3. Empirical analysis

3.1 Data

The postal address of households participating in official UK data surveys (e.g. the Family Expenditure and General Household surveys) is unavailable to the public for confidentiality reasons. In this sub-section we describe briefly the data used in the empirical analysis in this paper which are drawn from various sources, as detailed in Appendix.

The individual house price data are collected during 2008 from the internet site "Up my Street", which advertises houses for sale in the UK. In addition to prices, P_{si}, also collected from this site are the house-specific variables denoted by the vector $Q_{ki}, k = 1, \dots, K$ in equation (4), i.e. number of bedrooms, number of total rooms, type of the house, postal code etc. The average price of houses in our sample is around 252.000 GBP and 272.000 GBP for primary and secondary school datasets, respectively.

The two main school quality indicators used in our empirical analysis, the score and CVA, X_s and V_s, come from the primary and secondary education performance tables, available from the Department for Children, Schools and Families. These tables include background information on the schools in 2007.

- For primary education the score indicates the proportion of pupils reaching Level 4 in the Key Stage 2 (KS2) standard assessment tests administered at age 11; in our sample this proportion averages to around 81%[6].

- For secondary education the score indicates the proportion of pupils aged 15 years who pass five or more General Certificate of Secondary Education (GCSE) subjects at grades A to C; in our sample this proportion is, on average, around 47%.

---

[6] Concern about a censored regression problem is mitigated by the fact that none of the schools has 100% proportion of pupils reaching Level 4. Indeed, only 5% of the sample are above 95%.

England is the only state so far, where a national CVA indicator is constructed for primary and secondary schools using annual pupil-level data collected by the Pupil Level Annual Schools Census (PLASC). Initially simple value added indicators were constructed by adjusting the score indicator described above to take into account pupils' prior achievement. The more complex CVA indicator used in this paper, is calculated - using multilevel models – for all pupils as the difference (positive or negative) between their own 'output' point score and the median achieved by others with the same or similar 'starting' (or 'input') point score, after taking account the contextual factors collecting by PLASC. In our sample the average CVA indicator is equal to 99.92 for primary schools and 1002.02 for secondary schools[7].

Data on deprivation indices and other neighbourhood characteristics, denoted by the vector $Y_{ms}, m = 1, \ldots., M$ in (4), come from the Office of National Statistics of UK. All data collected cover the period June-September 2008 and include deprivation indices of income, crime, environment, housing barriers, health, and employment, and information about the density and non-domestic buildings in an area.[8]

3.2 Semi-parametric analysis

The CVA indicator, as implied from the means given above, is an arbitrarily normalised 'ordinal' measure of school performance. Indeed, this is the case for most published school quality indicators and can create problems of comparison and interpretation in empirical analysis. To avoid such problems investigators often (re)normalise school performance indicators to measure standard deviations from the mean. The measurement of a school quality indicator in standard deviations, however, limits the ability of the investigator to explore non-linearity in the relationship between this indicator and house prices in hedonic regression, insofar as higher order (quadratic or cubic) standard deviations can complicate interpretation and/or comparability of results across studies. In this context semi-parametric analysis can be a useful tool for investigating non-linear effects of school

[7] The fact that secondary CVA is V at 16 (final scores) minus V at 11 (prior scores), and primary CVA is V at 11 (final scores) minus V at 7 (prior scores) shows that there is a relationship between the two. However, this relationship cannot be exploited in this analysis because the available data measure school and not pupil performance. Details about the calculation methods and the range of background factors used in construction of CVA are given in Appendix.

[8] The houses allocated to the catchment area of a particular school are, on average, 0.2 miles away from it (range 0 to 1 miles). The results obtained from the empirical analysis do not change significantly if the smaller house-to-school maximum distances of 0.5 miles and 0.2 miles are used. This suggests that the catchment area may cover a fairly large radius around the school.

quality on house prices and, thereby, for finding appropriate ways to specify these effects in parametric analysis.

The semi-parametric estimator used in this paper is the 'nearest neighbour' one proposed by Estes and Honore (1995)[9], briefly described as follows.

We write equation (1) as

$$P_{si} = a + f(V_s) + \Sigma_k \gamma_k Q_{ki} + \Sigma_m \varepsilon_m Y_{ms} + u_{si} \tag{5}$$

where $f(V_s)$ is an unknown function, while all variables and notation in (5) are as defined in (1). Next we sort the data by V_s, and compute the differences: $\Delta P_{si} = P_{si} - P_{si-1}$; $\Delta Q_{ki} = Q_{ki} - Q_{ki-1}$, all k; and $\Delta Y_{ms} = Y_{ms} - Y_{ms-1}$, all m; where the subscript '-1' indicates the previous observation.

We then estimate the regression

$$\Delta P_{si} = \Sigma_k \gamma_k \Delta Q_{ki} + \Sigma_m \varepsilon_k \Delta Y_{ms} + u_{si}. \tag{6}$$

ΔP_{si} measures the difference in house price between the current observation and the observation that has a V_s value closest to the current. Note that, as the data are sorted by V_s (a continuous variable) no matter what the functional form of $f_x(V_s)$ is, the difference $\Delta f_x(V_s) \approx 0$ and can be ignored.

Using the parameter estimates obtained from (6), we compute the part of P_{si} not explained by the right hand side variables,

$$\hat{r}_{si} = P_{si} - \Sigma_k \hat{\gamma}_k \Delta Q_{ki} - \Sigma_m \hat{\varepsilon}_k \Delta Y_{ms}.$$

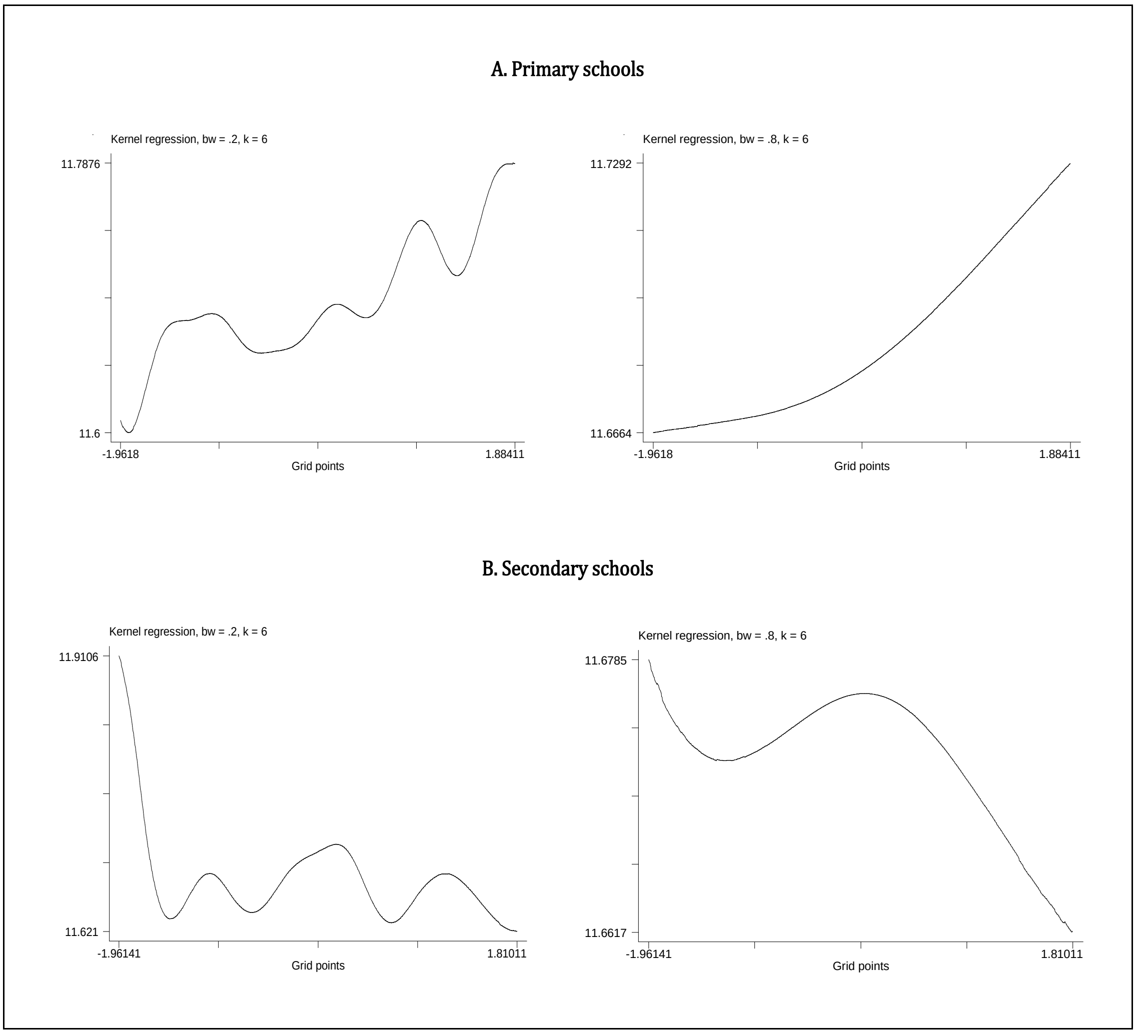
and perform separate semi-parametric regression of $\hat{r}_{si}$ on CVA using two alternative bandwidths, 0.2 and 0.8: the smaller bandwidth highlights details in the data, whereas the bigger bandwidth helps towards defining a parsimonious parametric model.

Figure 1 plots the weighted Gaussian kernel estimates of the relationship between log house prices and CVA for primary (part A) and secondary (part B) schools. In the case of primary schools it is clear that this relationship is positive for both bandwidths employed,

[9] This semi-parametric estimator is less efficient than Robinson's (1988) estimator but has computational advantages and is easier to implement. To eliminate kernel estimates based on a small number of observations we drop 2% of the sample from each end of the distribution.

0.2 and 0.8. For secondary schools, however, no (positive or negative) relationship appears to exist between log house prices and CVA. ‘Forcing’ the data to yield such a relationship with the large bandwidth (0.8) results in a complicated cubic pattern, where the effect on log house prices is negative, positive and negative for values of CVA below -0.83, between -0.83 and +0.73, and above +0.73 deviations from the mean, respectively.

Figure 1: Kernel estimates for CVA



The implications of our semi-parametric findings for modelling and estimating the effect of CVA on house prices using hedonic regressions are discussed in the next sub-section. The rest of this sub-section focuses on investigating how the score and non-score components of CVA are responsible for shaping the lines plotted in Figure 1. For this we perform semi-parametric regression of P_{si} on score (X_s) and non-score (V_s^*), following the same nearest neighbour estimator described above.[10] The Gaussian kernel weighted estimates obtained from these regressions (again, using two bandwidths, 0.2 and 0.8) are plotted in Figure 2.

Part A1 of Figure 2 reports the effects of score and Part A2 and the effects of non-score component of CVA on log house prices for primary schools. The plots show that score has a positive effect on house prices, although a smaller (0.2) bandwidth shows this to be interrupted for middle values of this CVA component. The plots for the non-score component show that no clear non-score effect on log house prices can be traced, except for large values of non-score, where the effect is likely to be positive. Put together, these results suggest that the positive effect of CVA on house prices shown for primary schools in Part A of Figure 1 is primarily attributed to its score component.

The results obtained from semi-parametric analysis of secondary school data, shown in Figure 2 for the score (Part B1) and non-score (Part B2) components of CVA, indicate a clearly positive relationship between the score component of CVA and log house prices. In contrast, the relationship between the non-score component and log house prices appears to be negative and less strong than the one corresponding to score component. It should be noted here that the opposite effect of score and non-score on house prices is probably the reason why the CVA indicator appears to have a non-significant effect on house prices in the analysis of secondary school data (Figure 1, Part B).

3.3 Parametric Analysis

The semi-parametric results discussed above imply that the effect of CVA (and its score and non-score components) on log house prices may or may not be linear and monotonic. As said earlier, the existence of non-linear and non-monotonic effects of CVA on house prices can complicate interpretation and/or cross-study comparison of higher order (quadratic or cubic) terms, because the CVA and its components are measured in standard deviations from the mean. Here we circumvent this problem by incorporating non-linearity and non-monotonicity in the effects of CVA on house prices using dummy variables, D_ℓ for $\ell = 1 \dots L$, in the hedonic regression

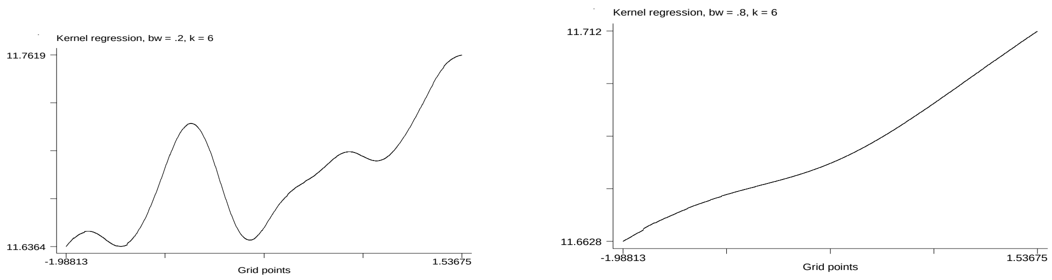
$$P_{si} = a + \sum_\ell \beta_\ell (V_s D_\ell) + \Sigma_k \gamma_k Q_{ki} + \Sigma_m \varepsilon_m Y_{ms} + u_{si}. \tag{1'}$$

[10] The non-score component (V_s^*) is defined as the residuals from regressing CVA on its score component. Given the orthogonality of X_s and V_s^* we investigate the semi-parametric relationship between $\hat{r}_{si}$ and each of the two indicators of school quality separately.

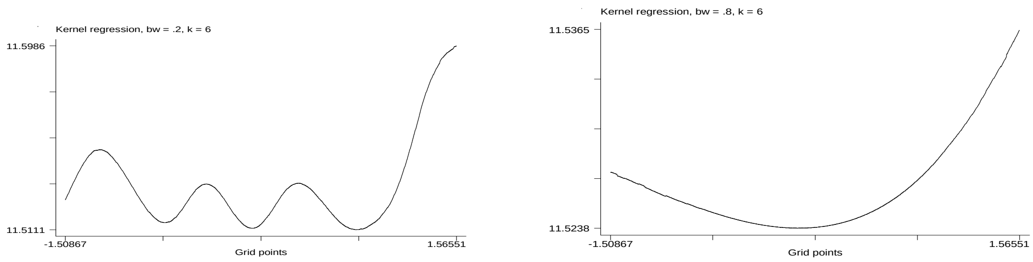
Figure 2: Kernel estimates for the score and non-score CVA components

A. Primary schools

A1. Score

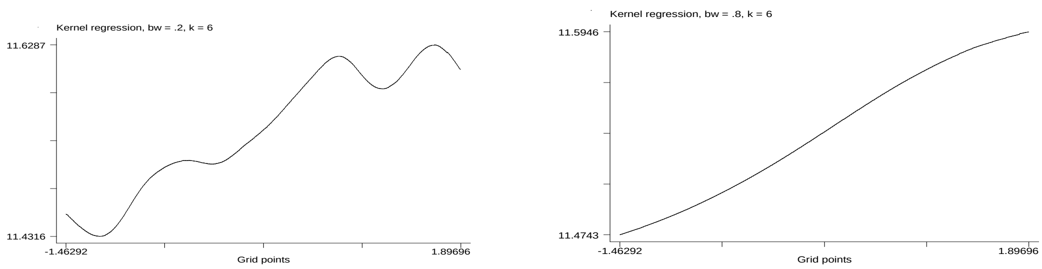


A2. Non-score

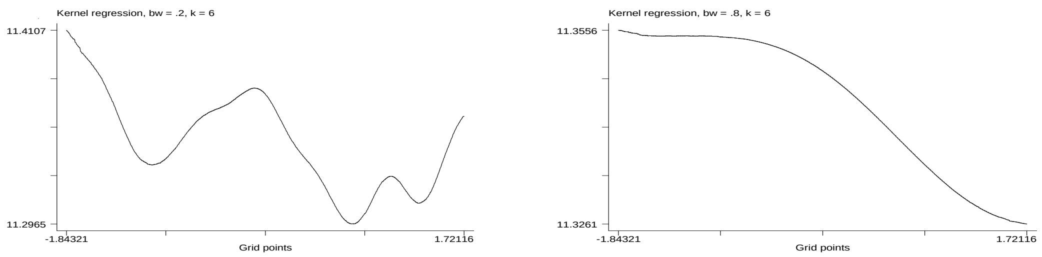


B. Secondary schools

B1. Score



B2 Non-score



For instance, we create and include in (1') two dummy variables: D1=1 if CVA<−1.75 and D1=0 otherwise; and D2=1 if D1=0 and D2=0 otherwise, where −1.75 is the value of CVA where the slope of the line in the corresponding semi-parametric plot (top right-hand side in Figure 1) change most. This allows for the effect of CVA on log house prices to differ between values below and above the threshold suggested by the semi-parametric results reported in Figure 1. In the case of secondary schools we create and include in (1') three dummy variables: D1=1 if CVA<−0.83 and D1=0 otherwise; D2=1 if −0.83≤CVA≤+0.73 and D2=0 otherwise; and D3=1 if CVA>+0.73 and D3=0. The values −0.83 and +0.73 correspond to the first and second reflection points of the line in the bottom right semi-parametric plot in Figure 1, respectively. Thus, the idea here is to investigate the possibility arising from the semi-parametric analysis that the effect of CVA on house prices is negative, positive and negative for values of CVA below −0.83, between −0.83 and +0.73, and above +0.73 deviations from the mean, respectively.

The parameters corresponding to the score and non-score components of CVA are estimated from the hedonic regression

$$P_{si} = a + \mu X_s + \sum_{\ell} \rho_{\ell} \left(V_s^{*} D_{\ell}\right) + \Sigma_k \gamma_k Q_{ki} + \Sigma_m \varepsilon_m Y_{ms} + u_{si}. \qquad (4')$$

For primary education, we create and include in (4') two dummy variables ($\ell = 2$) to allow the effect of non-score V_s^* to differ for values below and above −1.75, the minimum attained (with a bandwidth 0.8) in the semi-parametric relationship (Figure 2, Part A2). Three dummy variables ($\ell = 3$) are created and included in (4') for secondary education, in this case to investigate the semi-parametric finding that the relationship between non-score and log house prices is weaker for values of CVA below -0.83, between −0.83 and +0.73, and above +0.73 deviations from the mean (Figure 2, Part B2).

The full regression results obtained from the estimation of (1') and (4') are reported in Appendix (Table A3). Here, we focus only on results relating to the question how CVA and its score and non-score components affect house prices. The parameter estimates helping to answer this question are reported in Table 1 and, more or less, conform to expectation arising from the non-parametric analysis. In the column under the heading 'Model I' are the parameter estimates of (1') where the effect of CVA on log house prices is allowed to vary for values above and below −1.75. In the case of primary education the effect of CVA is positive and significant only for values above this threshold. For secondary education, the effect of CVA on log house price is generally negative but significant only for values above +0.73 standard deviations from the mean.

Table 1: The effect of CVA and its components on log house prices
(Robust standard errors in brackets)

A. Primary schools

Variable	Parameter	Model I	Model II	Model III	Model IV	Model IV (2SLS)
CVA	β			0.022**		
				(0.009)		
CVA×D1	β_1	0.068				
		(0.060)				
CVA×D2	β_2	0.019**				
		(0.009)				
Score	μ		0.034***		0.035***	0.079***
			(0.008)		(0.008)	(0.027)
Non-Score	ρ				-0.001	-0.001
					(0.011)	(0.011)
Non-Score×D1	ρ_1		-0.024			
			(0.018)			
Non-Score×D2	ρ_2		0.008			
			(0.013)			
R-squared		0.851	0.852	0.851	0.852	0.849
No. of observations		1385	1385	1385	1385	1385

B. Secondary schools

Variable	Parameter	Model I	Model II	Model III	Model IV	Model IV (2SLS)
CVA	β			-0.027***		
				(0.009)		
CVA×D1	β_1	-0.016				
		(0.017)				
CVA×D2	β_2	-0.026				
		(0.021)				
CVA×D3	β_3	-0.035**				
		(0.015)				
Score	μ		0.038***		0.039***	0.068***
			(0.010)		(0.010)	(0.024)
Non-Score	ρ				-0.043***	-0.046***
					(0.010)	(0.011)
Non-Score×D1	ρ_1		-0.031			
			(0.020)			
Non-Score×D2	ρ_2		-0.048***			
			(0.012)			
R-squared		0.837	0.84	0.837	0.84	0.838
No. of observations		1209	1209	1209	1209	1209

Notes: *** and ** indicate significance at 0.01 and 0.05 level, respectively.

In the column under the heading 'Model II' are the parameter estimates of (4'), where the effects of score and non-score components of CVA are estimated separately. The results demonstrate the positive and significant effect of score on log house prices at both primary and secondary levels of education. In contrast, the effect of the non-score component is negative and significant only for secondary education and for values above -0.82 standard deviations from the mean.

The parameters in Table 1 reported in the columns under the headings 'Model III' and 'Model IV' correspond to the estimation of (1') and (4'), assuming that $\beta_\ell = \beta$ and $\rho_\ell = \rho$ for all $\ell = 1 \dots L$, respectively, i.e. the effect of CVA and its components on log house prices is assumed to be linear. As expected this assumption is accepted for both equations and levels of education, given that when the effect (slope) of each variable is differentiated by dummies, only one of these effects is statistically significant.[11]

'Model IV' is also estimated by 2SLS (last column of Table 1) in order to get round the problem of potential endogeneity and measurement error of the score component of CVA. As Gibbons and Machin (2003) suggest school performance is likely to be related to house prices through factors other than school quality sorting: prosperous parents may purchase houses in neighbourhoods with better amenities, so schools in these neighbourhoods perform better because their pupils are more receptive to education.[12] In addition, results published in the national tables (a single year measure) can be noisy measures of long-run school quality.

As in Gibbons and Machin (2003), we investigate the endogeneity and/or measurement error problems by instrumenting school quality indicators with variables that are available in the school performance tables. More specifically the instruments used for the CVA are the school type, the admissions age-range and the student gender (available only for secondary schools). The 2-SLS results suggest that the effect of score component on house prices is higher than that obtained from simple regression. This is the case for both primary and secondary schools. The same upward 2-SLS 'correction' of the score effect on house prices estimated by simple regression is reported by Gibbons and Machin (2003) indicating that errors in the measurement of the score variable may be a more serious

[11] The F-values are: 0.66 for $\beta_1 = \beta_2 = \beta$ in primary and 0.31 for $\beta_1 = \beta_2 = \beta_3 = \beta$ in secondary education; and 0.11 for $\rho_1 = \rho_2 = \rho$ in primary and .50 for $\rho_1 = \rho_2 = \rho$ in secondary education.

[12] At the same time, however, one has to be cautious

source of bias than unaccounted endogeneity from neighbourhood quality effects on school performance.[13]

It is worth noting here that one can also consider the use of CVA and, specifically, its non-score component as a potential instrument for tackling the endogeneity of final score. To investigate this possibility in the case of the primary schools (where the non-score component of CVA appears to be exogenous - Model II in Table 2.1A) we include this variable as an instrument for the final score along with the school type, the admissions age-range and the student gender. The results remain unchanged, i.e. the score coefficient is 0.079*** as is without using the non-score as an the instrument for the score component of CVA (Model IV in Table 2.1A), suggesting that the additional instrument does not add to identification.[14]

The conclusion emerging from the parametric and non-parametric analysis so far is that CVA has a significant positive effect on house price in the case of primary and a strong significant negative effect in the case of secondary education. By separating the overall effect of the CVA indicator into its score and non-score components, it becomes evident that positive CVA effect on house prices in the analysis of primary school data is entirely attributed to the score component; whereas the negative effect in the analysis of secondary school data is due to the large and significant negative effect of the non-score component, which more than compensates the equally significant but not so large positive effect of the score component.

The estimates reported in Table 1 are conditional on Local Authorities (LA) – see Appendix A3. As such they reflect the effect of CVA and its score and non-score components on house prices 'within' LAs and are interpreted as indicators of willingness to pay for school quality by households already located in a particular LA. In order to also investigate how differences in school quality affect house prices 'between' LAs we re-estimate (1') and (4') using the LA means. The results of these estimations are reported in Table 2.

[13] One can argue here that the IV estimates can be associated with a Local Average Treatment Effect (LATE), i.e. reflect on observations in the sample to which the instrument is relevant, for instance changes in CVA can be specific to the school type. Nevertheless, the validity of the instruments used in the paper is supported by the high F-statistics (68.5 for primary and 75.8 for secondary schools) and the fact that over-identification tests, based on the R-square obtained from regressing the predicted errors from the 2-SLS estimation on all exogenous variables, suggest that all models are just identified.

[14] The same exogeneity test cannot be used in the case of the secondary schools because the non-score component of CVA is not exogenous (Model IV in Table 2.1B).

For primary schools, the results in Table 2 suggest that the between LAs effect of CVA on house prices is positive, as is within LAs (Table 1). However, the estimated parameter now is larger in size and significance compared to that estimated within LAs. Furthermore, the enhancement of the CVA effect between LAs appears to come from equal in size and significance contributions from both the score and non-score components; unlike the LA-conditional CVA effect, which is attributed entirely to score. This means that households are willing to pay a higher price for houses in LAs where primary schools have, on average, higher non-score and score components of CVA; whereas, once they are in a given LA, their willingness to pay a higher price for a house is limited only to primary schools with a high score.

Table 2: Estimated parameters 'between' Local Authorities
(robust standard errors in brackets)

Variable	Parameter	Primary schools		Secondary schools	
CVA	β	0.315*** (0.090)		-0.011 (0.064)	
Score	μ		0.232* (0.128)		0.035 (0.072)
Non-Score	ρ		0.295*** (0.094)		-0.015 (0.065)
R-squared		0.920	0.921	0.907	0.908
No. of observations		51	51	51	51

Notes: *** and ** indicate significance at 0.01 and 0.05 level, respectively.

The parameters for secondary schools reported in Table 2, again, change substantially compared to those obtained conditional on LA-specific effects. In this case, however, the CVA and its score and non-score components do not have a significant effect on house prices. For instance, while within LAs house prices are reduced, on average, by 2,7% for an increase in the standard deviation of CVA by one unit, house prices across LAs do not appear to be affected. Also, the results in Table 1 suggest that within LAs house prices are lower in catchment areas of secondary schools with a high CVA; while Table 2 suggests that between LAs house prices are not significantly associated with CVA differences between these schools. On the basis of these results one can argue that the willingness of households to pay more for a high quality secondary education appears to depend entirely on the final score and is influenced by school comparison within (rather than between) LAs.

4. Discussion

The impact of educational output, as measured by student test scores on standardised tests, on house values is consistently positive (Black, 1999; Gibbons and Machin, 2003; Haurin and Brasington, 2006; Rosenthal, 2003), suggesting that families are willing to pay for higher educational output. The same house price effect, however, is not so clear when measures of school quality other than educational output are used in the hedonic regression, e.g. educational inputs, usually expenditure per pupil (Haurin and Brasington, 1996; Brasington, 1999; Black,1999; Downes and Zabel, 1992); and value added (VA), as measured by the growth of student achievement over time (Hayes and Taylor, 1996; Brasington, 1999; Haurin and Brasington, 2006; Downes and Zabel, 2002).

This paper focuses on the relationship between house prices and a new measure of school quality recently introduced in England, termed Contextual Value Added (CVA). Unlike VA the CVA rewards growth of student achievement that takes place under disadvantageous conditions, such as poor socio-economic background, ethically heterogeneous classes, poor/interrupted attendance etc. The emphasis in the paper is on: (i) finding an appropriate empirical specification to capture the effects of this school indicator and its score (school factor) and non-score (non-school factor) components on house prices; and (ii) highlighting the potential heterogeneity in these effects at different levels of education and between neighbourhoods and spatially larger entities, like Local Authorities (LAs).

To our knowledge Hayes and Taylor (1996) are the first to consider VA in investigating the relationship between house prices and school quality. Using Dallas data they try to separate the impact of achievement score on house prices into school (value added) and demographic composition (peer group) effects and find only the former to be significant. Brasington (1999) considers the impact of VA on house prices, along with several other measures of educational inputs, for the six largest metropolitan areas of Ohio. Measuring VA as the percentage change in the number of proficient students and using a variety of relative performance indices in hedonic and a spatial analysis he too finds no evidence of VA impact on house prices. More specifically, the results show the effect of VA to vary from positive to negative (and zero) between metropolitan areas; in contrast test score measures, expenditure per pupil and teacher salary are found to have a positive effect on house prices. These results are reinforced by Brasington and Haurin (2006) who find no evidence of capitalization of alternative VA measures into the prices of 77000 houses sold in 2000 in Ohio. In fact, after defining VA as in Brasington (1999) they obtain a negative effect on house prices when controlling for score and expenditure per pupil.

Gibbons et al (2009) investigate the effect of VA and prior achievement on house prices for UK primary schools during the period 2003-2006 using boundary discontinuity regressions. They find that a one-standard deviation change in either VA or prior achievement raises prices by around 3-4%. It is important to stress here there is no contradiction between these results and those reported in the US studies discussed above, where VA does not appear to affect house prices. This is because Gibbons et al (2009) condition the VA effect on prior achievement, whereas the VA effect in the US studies is conditioned on final score. Therefore (i) the prior achievement parameter in Gibbons et al (2009) is equivalent to the final score parameter in the US studies; and (ii) the VA parameter in Gibbons et al (2009) is equivalent to the sum of final score and VA parameters in the US studies.[15]

The empirical findings in our analysis complement those obtained in the literature, insofar as we use a CVA measure recently introduced in England to investigate the effect of school quality on house prices. The relationship between school performance and house prices is more difficult to analyse when a CVA, rather than a VA, measure is used because the CVA also accounts for various non-school factors affecting educational outcomes. For instance, we find the effect of CVA on house prices to be positive for primary and negative for secondary schools. The heterogeneity of the results at different education levels is explained when the CVA effect is broken into its score and non-score components: the former is positive and strongly significant for both primary and secondary schools; while the latter is significant and negative for secondary and insignificant for primary schools. Viewing this result in the context of the studies discussed above, one can argue that our results confirm what most other investigators report in the literature, i.e. other indicators of school quality do not have a significant positive effect on house prices once controlled for the final score.

An explanation why the analysis of secondary school data shows the non-score component of CVA to have a significantly negative effect on house price may be that pupils at this level of education can vary substantially in their prior achievement, the most important item in

[15] To see this write the house price equation $P = \alpha + \beta V + \gamma X$, where $V = X - A$ is value added, X final score and A prior achievement. Replacing X with $V + A$ gives $P = \alpha + (\beta + \gamma)V + \gamma A$.

the non-score part of CVA; whereas the opposite is true for primary schools, because there can be little variation in prior achievement among pupils at the beginning of their education.[16] Furthermore, the negative sign of the effect of the non-score component of CVA on house prices in the analysis of secondary school data can be attributed to the fact that reaching a given final score starting from a high (low) initial level subtracts from (adds to) the school's image of quality.

In addition to differences between the results obtained from primary and secondary school data, we have found heterogeneity in the effect of CVA depending on whether this effect relates to house price variation within or between LAs. For example, we find that for primary schools CVA has a positive effect on house prices between LAs, indicating the willingness of households to pay a higher price for houses in LAs where primary schools have, on average, a higher CVA, either due to the score or the non-score component. This can be interpreted as highlighting the 'public good' nature of this school quality indicator and suggests that the school quality aspects reflected in a high CVA (achieving a high final score against the odds of low prior achievement, a large proportion of pupils whose mother tongue is not English and other circumstances non-conducive to learning) may be diffused over a greater geographical area than that of the catchment area of a particular school. However, the same cannot be said about secondary schools for which the empirical analysis suggests that higher house prices in the catchment area of schools at this level of education is entirely a reflection of the score component of the school's CVA within the LA; while the non-score component appears to be either irrelevant (between LA's) or affect house prices negatively (within LAs).

A limitation of the empirical analysis in this paper is the inability to analyse separately the relationship between each individual component of CVA and house prices, due to the fact that the Department of Education, Children and Families of the UK does not publish this information. While the separation of the score from the non-score components attempted in the paper helps illustrate the potential heterogeneity of the effects of CVA on house prices, itemising the non-score components (to prior achievement, proportion of pupils in disadvantaged groups etc) would enable one to reach more informed conclusions, for instance, about how households value private and social aspects of school quality.

[16] To test this conjecture we have regressed CVA on its score component: the results showed that 41% of the variation in CVA among primary schools is explained by this component; whereas for secondary schools, the corresponding figure was only 18%.

5. Conclusion

The analysis in this paper highlights the usefulness and complexities of trying to estimate the impact of Contextual Value Added - a more comprehensive measure of school quality - on house prices using hedonic regressions. In general, the fact that this measure plays down the importance of final score, evidently a very desirable aspect of school quality from the point of view of the private household, is shown to weaken (or reverse) the positive link between school performance and house prices. Yet, households with children in primary education appear to pay a higher price for housing in Local Authorities with schools showing high non-score quality characteristics. In terms of policy, our findings suggest that a Local Authority can pay attention to improving the non-score aspects of the CVA of its primary schools, knowing that this can make it an attractive location among households looking for high quality primary education. However, the willingness of households to pay more for a house in the catchment area of a school offering high quality secondary education appears to depend entirely on the final score.

Acknowledgments

We would like to thank Steve Gibbons, David Brasington and Ian Walker for valuable comments and suggestions. We would also like to thank Adamos Andreou, Marios Charalambous and Georgia Loizide for their help on data collection process. We are, of course, responsible for all errors and omissions.

References

Barrow L. and C. Rouse, 2004. Using Market Valuation to assess Public School Spending, Journal of Public Economics 88, 1747-1769.

Barrow L., 2002. School Choice through Relocation: Evidence from the Washington DC Area, Journal of Public Economics 86,155-189.

Black S., 1999. Do Better Schools Matter? Parental Valuation of Elementary Education, The Quarterly Journal of Economics, 114, 2, 577-599.

Bogart W. and B. Cromwell, 1997. How Much More is a Good School District Worth?, National Tax Journal, 50, 2, 215-232.

Bogart W. and B .Cromwell, 2000. How Much is a Neighbourhood School Worth? Journal of Urban Economics, 47, 280-305.

Brasington D., 1999. Which Measure of school quality does the Housing Market value? Journal of Real Estate Research, 18, 395-413.

Brasington D., 2000. Demand and Supply of Public Quality in Metropolitan Areas: The Role of Private Schools, Journal of Regional Science, 40, 3, 583-605.

Brasington D., 2002. The Demand for Local Public Goods: The Case of Public School Quality", "Public Finance Review, 30, 163-187.

Brasington D. And D. Haurin, 2009. Parents, Peers, or School Inputs: Which Components of School Outcomes are Capitalized in to House Value?, Regional Science and Urbal Economics, 39, 523-529.

Chay K, P. McEwan and M. Urquiola, 2005. The Central Role of Noise in Evaluating Interventions That Use Test Scores to Rank Schools, The American Economic Review, 95, 4, 1237-1258.

Cheshire P. and S. Sheppard, 2004. Capitalizing the Value of Free Schools: The Impact of Supply Characteristics and Uncertainty, Economic Journal, 114, 499, 397-424.

Clapp J. , A. Nanda and S. Ross, 2008. Which School Attributes Matters? The Influence of School District Performance and Demographic Composition on Property Values, Journal of Urban Economics, 63, 451-466.

Downes A. and J. Zabel, 2002. The Impact of School Characteristics on House Prices: Chicago 1987-1991, Journal of Urban Economics, 52, 1-25.

Downes D. and O. Vindurampulle, 2007. Value-Added measures for School Improvement, Office for Education Policy and Innovation/Department of Education and Early Childhood Development Melbourne, No. 13.

Estes E. and B. Honore, 1995. Partially Linear Regression using one Nearest Neighbor, Princeton University, Department of Economics manuscript.

Fernandez R. and R. Rogerson, 1996. Income Distribution, Communities and the Quality of Public Education, The Quarterly Journal of Economics, 111, 1, 135-164.

Fernandez R. and R. Rogerson, 1998. Public Education and Income Distribution: A Dynamic Quantitative Evaluation of Education Finance Reform, American Economic Review, 88, 4, 813-833.

Gibbons S. and S. Machin, 2003. Valuing English Primary Schools, Journal of Urban Economics, 53, 197-219.

Gibbons S. and S. Machin, 2006. Paying for Primary Schools: Admissions Constraints, School Popularity or Congestion, The Economic Journal, 116, 77-92.

Gibbons, S and S. Machin, 2008. Valuing School Quality, Better Transport and Lower Crime: Evidence from House Prices, Oxford Review of Economic Policy, 24, 1, 99-119.

Gibbons S., S. Machin and O. Silva, 2009. Valuing School Quality using Boundary Discontinuity Regressions, Spatial Economics Research Centre, Discussion Paper 18.

Hanushek E. and L. Taylor, 1990. Alternative Assessments of the Performance of Schools, Journal of Human Resources, 25, 179-201.

Hanushek E., 1992. The trade off between Child Quantity and Quality, Journal of Political Economy, 100, 84-117

Haurin D. and D. Brasington, 1996. School Quality and Real House Prices, Journal of Housing Economics, 5, 351-368.

Haurin D. and D. Brasington, 2006. Educational Outcomes and House Values: A test of the Value Added Approach, Journal of Regional Science, 46, 2, 245-268.

Hoxby C., 2000. Does Competition among Public Schools benefit students and taxpayers? Comment, American Economic Review, 90, 1209-1238.

Kane T., S. Riegg and D. Staiger, 2006. School Quality, Neigborhoods and Housing Prices, American Law and Economics Review, .8, 2, 183-212.

Leech D. & E. Campos, 2003. Is Comprehensive Education Really Free? A Case Study of the Effects of Secondary School Admissions on House Prices, Journal of the Royal Statistical Society, 166, 1, 135-154.

Mayer R., 1997. Value-Added Indicators of School Performance: A primer, Economics of Education Review, 16, 3, 283-301.

Ray A., 2006. School Value Added Measures in England, Department for Education and Skills UK.

Robinson M., 1988. Root-N consistent Semiparametric Regression, Econornetrica, 56, 931-954.

Rosen S., 1974. Hedonic Prices and Implicit Markets: Product Differentiation in Pure Competition, Journal of Political Economy, 82, 34-55.

Rosenthal L. , 2003. The Value of Secondary School Quality, Oxford Bulletin of Economics and Statistics, 65, 3, 329-355.

Summers A. and B. Wolfe, 1977. Do Schools Make a Difference? The American Economic Review, 67, 4, 639-652.

Appendix

A1. Data sources

Our data come from three major independent sources. The house price and characteristics data have been drawn from the electronic site "Up my Street"[17]. The school quality data come from the primary and secondary school performance tables, available from the Department for Children, Schools and Families[18], and the deprivations indices and other neighbourhood characteristics from the Office of National Statistics[19].

The data collection process from the three different sources was as follows: England is divided in nine regions consisting of one hundred fifty Local Authorities (LAs). Fifty LAs were chosen, one third from each region, half with a larger and the other half with a lower mean grade than the England average.[20] From each of these LAs six schools were randomly selected, three with a higher and three with a lower grade mean than the LA average. This process was accomplished separately for primary and secondary schools.

Using the school postcode we were able to locate the six houses closest to the school that were up for sale using information from the Up my Street website. We collected information on the selling price of houses, house characteristics and distance from school. The average distance from the local primary or secondary school was around 0.20 miles and in no case more than one mile.

The site of Neighbourhood Statistics provides detailed statistics within specific geographic areas, including deprivation indices of income, crime, environment, housing barriers, health, and employment and other neighbourhood characteristics like population density. In order to capture neighbourhood characteristics, we used the "lower layer super output area[21]" for each specific postcode and collected the following indicators:

- Income: the proportion of the population living in low income families.
- Employment: involuntary exclusion from work of working age population.
- Health and Disability: rate of premature death, poor health and disability.
- Barriers to Housing and Services: barriers to GP premises, supermarkets, primary schools and post offices, divided into 'geographical barriers' and 'wider barriers'.
- Living Environment: 'Indoors' measuring the quality of housing and 'outdoors' measuring the air quality and road traffic accidents.
- Crime: the rate of recorded crime (burglary, theft, criminal damage and violence).
- Density: the number of persons per hectare[22].

[17]From this site we had access in all properties which have been bought and sold in the whole of England. (website: www.upmystreet.com/property-prices/find-a-property/l/n13+5rx.html)

[18]School performance tables include background information such as type of schools, are range of pupils and gender of pupils (website: www.dcsf.gov.uk/index.htm)

[19] Website: www.neighbourhood.statistics.gov.uk/dissemination

[20] For those regions containing a number of LAs that could not be divided by three, the number of LAs finally chosen was rounded up or down to the nearest one third of the total number of LAs.

[21] Roughly one LA is divided into 100-150 Lower Layer Super Output Areas. For example, Islington (LA) has 175,000 residents and it is divided into 117 Lower Layer Super Output Areas. Thus, on average each Lower Layer Area has around 1,500 residents. Also, on average there are 2, 5 persons per household. Hence, a lower layer area has about 600 households

[22] Resident people per hectare in the area at the time of the 2001 Census.

A2. Value Added and Contextual Value Added

Value added (VA) is a measure of the progress pupils make between different stages of education. The VA score for each pupil, as defined by the Department of Education, Children and Families of the UK, is the difference (positive or negative) between their own 'output' point score and the median - or middle - output point score achieved by others with the same or similar starting point, or 'input' point score. Thus, an individual pupil's progress is compared with the progress made by other pupils with the same or similar prior attainment. In order to calculate this measure the Department used a median line approach (Ray, 2006).

Contextual Value Added (CVA) was introduced in order to account for student, family and socioeconomic characteristics affecting the progress made by pupils. The technique used to derive a CVA score is called multi-level modelling (MLM) performed in four stages: (1) obtain a prediction of attainment based on the pupil’s prior attainment; (2) adjust this prediction to take account of the pupil’s set of characteristics; (3) for key stage 2-4 adjust further to account for school level prior attainment and (4) calculate the CVA by measuring the difference (positive or negative) between the pupils actual attainment and that predicted by the CVA model.

The data for the calculation of CVA are drawn from the Pupil Level Annual School Census (PLASC), a national dataset for some 600,000 pupils in England.[23] The PLASC was introduced in 2002 with the aim of collecting contextual data from schools’ administrative records on all pupils annually (i.e. not just at the end of each key stage). The main variables in the PLASC data used in the calculation of CVA are the gender, special educational needs, ethnicity, eligibility for free schools meals, language, date of entry/mobility, age, being in care and the income deprivation affecting children index.

The CVA measure for primary schools is normalised to 100, whereas for secondary schools the same measure is normalised to 1000: scores above (below) these norms represent schools where students made more (less) progress than similar students nationally. The table below shows

Table A2.1 The distribution of the CVA score indicator

Primary schools	Secondary schools	Percentiles nationally
101.5 and above	1041.11 and above	Top 5%
100.6 to 101.4	1013.41 to 1041.10	Next 20%
100.2 to 100.5	1006.11 to 1013.40	Next 15%
99.8 to 100.1	997.61 to 1006.10	Middle 20%
99.4 to 99.7	990.66 to 997.60	Next 15%
98.5 to 99.3	971.54 to 990.65	Next 20%
98.4 and above	971.53 and below	Bottom 5%

[23] Some external factors which are commonly thought to impact on pupil’s performance (e.g. parental education status/occupation) are not included in the calculation of CVA because no reliable national data are available.

A3. Estimated parameters for all variables in the hedonic analysis

	Primary Schools		Secondary Schools	
Ln House Price	Model I	Model II	Model I	Model II
CVAxD1 (low CVA)	0.068		-0.016	
	(0.060)		(0.017)	
CVAxD2 (middle CVA)	0.019**		-0.026	
	(0.009)		(0.021)	
CVAxD3 (high CVA)			-0.035**	
			(0.015)	
Score		0.034***		0.038***
		(0.008)		(0.010)
Non-ScorexD1(low Non-Score)		-0.024		-0.031
		(0.018)		(0.020)
Non-ScorexD2(high Non-Score)		0.008		-0.048***
		(0.013)		(0.012)
<u>House Characteristics</u>				
Bedrooms	0.095***	0.094***	0.097***	0.100***
	(0.015)	(0.015)	(0.017)	(0.017)
Bathrooms	0.094***	0.093***	0.034	0.035
	(0.021)	(0.021)	(0.025)	(0.024)
Toilet	-0.011	-0.014	-0.020	-0.022
	(0.017)	(0.017)	(0.021)	(0.021)
Totalrooms	0.056***	0.056***	0.080***	0.078***
	(0.008)	(0.008)	(0.009)	(0.009)
Garage	0.038***	0.038***	0.061***	0.059***
	(0.014)	(0.014)	(0.015)	(0.015)
Semidetached	-0.098***	-0.095***	-0.138***	-0.133***
	(0.021)	(0.020)	(0.021)	(0.021)
Flat	-0.271***	-0.270***	-0.154***	-0.163***
	(0.031)	(0.031)	(0.037)	(0.037)
Terrace	-0.168***	-0.170***	-0.187***	-0.185***
	(0.027)	(0.026)	(0.034)	(0.033)
Midterrace	-0.217***	-0.218***	-0.264***	-0.266***
	(0.028)	(0.028)	(0.031)	(0.031)
Endterrace	-0.206***	-0.203***	-0.210***	-0.211***
	(0.031)	(0.031)	(0.031)	(0.031)
Maisonette	-0.264***	-0.268***	-0.285***	-0.294***
	(0.043)	(0.044)	(0.049)	(0.049)
Other	-0.079**	-0.081**	-0.007	-0.011
	(0.040)	(0.040)	(0.057)	(0.056)
<u>Neighborhood Characteristics</u>				
Income_deprivation	-0.046**	-0.039**	-0.008**	-0.010**
	(0.019)	(0.019)	(0.004)	(0.004)
Housing_deprivation	0.007	0.008	-0.029**	-0.026*
	(0.012)	(0.012)	(0.014)	(0.014)
Crime_ deprivation	-0.007	-0.003	-0.042***	-0.040***
	(0.009)	(0.009)	(0.011)	(0.011)
Environment_deprivation	-0.004	-0.004	0.048***	0.045***

	(0.013)	(0.013)	(0.015)	(0.014)
Health_deprivation	-0.044*	-0.035	-0.098***	-0.080***
	(0.024)	(0.024)	(0.024)	(0.023)
Employment_deprivation	-0.038*	-0.042**	-0.041**	-0.039**
	(0.021)	(0.021)	(0.018)	(0.017)
Density (persons per hectare)	-0.034***	-0.037***	-0.039***	-0.034***
	(0.010)	(0.010)	(0.011)	(0.011)
Non-domestic buildings (square metres)	-0.001	-0.002	0.001	0.006
	(0.009)	(0.009)	(0.007)	(0.007)
<u>Regions</u>				
East Midlands	-0.488***	-0.464***	-0.594***	-0.602***
	(0.092)	(0.091)	(0.120)	(0.116)
East England	-0.573***	-0.565***	-0.485***	-0.485***
	(0.065)	(0.065)	(0.127)	(0.123)
North East	-0.860***	-0.851***	-0.821***	-0.820***
	(0.062)	(0.064)	(0.088)	(0.082)
North West	-0.520***	-0.512***	-0.560***	-0.585***
	(0.067)	(0.067)	(0.084)	(0.078)
South East	-0.253***	-0.245***	-0.364***	-0.368***
	(0.060)	(0.060)	(0.093)	(0.090)
South West	-0.533***	-0.520***	-0.478***	-0.464***
	(0.091)	(0.091)	(0.079)	(0.074)
West Midlands	-0.680***	-0.680***	-0.473***	-0.503***
	(0.056)	(0.056)	(0.081)	(0.083)
Yorkshire	-0.610***	-0.633***	-0.731***	-0.742***
	(0.078)	(0.072)	(0.070)	(0.068)
<u>Local Authorities</u>				
Bath	-0.075	-0.080	-0.161*	-0.182**
	(0.093)	(0.091)	(0.086)	(0.085)
Bexley	-0.336***	-0.331***	-0.422***	-0.426***
	(0.058)	(0.058)	(0.072)	(0.067)
Blackpool	-0.329***	-0.343***	-0.193***	-0.163***
	(0.053)	(0.053)	(0.067)	(0.062)
Bradford	-0.181**	-0.159**	-0.066	-0.060
	(0.076)	(0.067)	(0.055)	(0.052)
Brent	0.028	0.024	0.040	0.020
	(0.062)	(0.067)	(0.096)	(0.096)
Buckinghamshire	-0.421***	-0.427***	-0.093	-0.078
	(0.051)	(0.053)	(0.079)	(0.075)
Camden	0.685***	0.703***	0.598***	0.565***
	(0.068)	(0.068)	(0.071)	(0.070)
Cheshire	-0.168**	-0.201**	-0.248***	-0.234***
	(0.082)	(0.081)	(0.079)	(0.073)
Coventry	-0.173***	-0.158***	-0.300***	-0.281***
	(0.043)	(0.043)	(0.066)	(0.070)
Darlington	0.202**	0.188*	0.187	0.177
	(0.102)	(0.099)	(0.170)	(0.153)
Derby	-0.446***	-0.469***	-0.342***	-0.344***
	(0.085)	(0.083)	(0.116)	(0.113)
Enfield	-0.114*	-0.098	-0.105	-0.134**
	(0.069)	(0.070)	(0.065)	(0.061)
Gloucestershire	-0.168*	-0.166*	-0.084	-0.103
	(0.097)	(0.097)	(0.092)	(0.091)

Greenwich	-0.109	-0.092	0.097	0.095
	(0.067)	(0.068)	(0.101)	(0.101)
Hampshire	-0.388***	-0.440***	-0.253***	-0.264***
	(0.057)	(0.065)	(0.085)	(0.080)
Havering	-0.385***	-0.396***	-0.349***	-0.386***
	(0.057)	(0.057)	(0.077)	(0.075)
Islington	0.731***	0.739***	0.718***	0.711***
	(0.066)	(0.068)	(0.082)	(0.080)
Kirklees	-0.288***	-0.258***	-0.141*	-0.152*
	(0.079)	(0.069)	(0.079)	(0.080)
Lambeth	0.455***	0.449***	0.420***	0.400***
	(0.071)	(0.070)	(0.080)	(0.078)
Lancashire	-0.301***	-0.325***	-0.370***	-0.363***
	(0.071)	(0.071)	(0.071)	(0.067)
Leeds	-0.283***	-0.269***	-0.163***	-0.170***
	(0.088)	(0.081)	(0.049)	(0.048)
Newham	-0.029	-0.026	0.074	0.016
	(0.057)	(0.057)	(0.072)	(0.070)
Norfolk	-0.298***	-0.303***	-0.393***	-0.402***
	(0.055)	(0.055)	(0.111)	(0.109)
Northtyneside	0.030	0.010	0.057	0.029
	(0.052)	(0.052)	(0.054)	(0.052)
Northyorkshire	-0.182**	-0.159**	-0.070	-0.056
	(0.072)	(0.065)	(0.063)	(0.063)
Northamptonshire	-0.426***	-0.443***	-0.322***	-0.345***
	(0.088)	(0.084)	(0.107)	(0.105)
Reading	-0.217***	-0.233***	-0.024	-0.016
	(0.056)	(0.055)	(0.094)	(0.087)
Richmond	0.222***	0.240***	-0.122	-0.130
	(0.075)	(0.076)	(0.086)	(0.082)
Salford	-0.111	-0.121*	0.065	0.060
	(0.072)	(0.070)	(0.091)	(0.089)
Solihull	-0.018	-0.016	-0.170**	-0.181**
	(0.071)	(0.072)	(0.075)	(0.076)
Somerset	-0.195**	-0.196**	-0.218***	-0.227***
	(0.086)	(0.085)	(0.069)	(0.065)
Southampton	-0.430***	-0.424***	-0.232***	-0.235***
	(0.045)	(0.045)	(0.079)	(0.074)
Southend	0.054	0.063	-0.060	-0.089
	(0.051)	(0.050)	(0.104)	(0.104)
Staffordshire	-0.201***	-0.208***	-0.415***	-0.405***
	(0.055)	(0.055)	(0.072)	(0.075)
Stockport	-0.156***	-0.159***	-0.052	-0.066
	(0.058)	(0.057)	(0.074)	(0.067)
Stockton	-0.035	-0.039	-0.096	-0.095
	(0.058)	(0.059)	(0.061)	(0.060)
Sutton	-0.264***	-0.236***	-0.396***	-0.427***
	(0.058)	(0.059)	(0.071)	(0.069)
Swindon	-0.319***	-0.316***	-0.331***	-0.354***
	(0.085)	(0.085)	(0.076)	(0.072)
Trafford	0.080	0.072	0.036	0.057
	(0.064)	(0.064)	(0.077)	(0.073)
Wigan	-0.317***	-0.327***	-0.263***	-0.271***

	(0.073)	(0.072)	(0.066)	(0.061)
Wiltshire	-0.103	-0.092	-0.263***	-0.297***
	(0.094)	(0.093)	(0.073)	(0.073)
Wokingham	-0.270***	-0.270***	-0.252***	-0.240***
	(0.053)	(0.052)	(0.079)	(0.077)
Constant	12.070***	12.059***	12.018***	12.041***
	(0.064)	(0.064)	(0.084)	(0.079)
N	1385	1385	1209	1209
R2	0.851	0.852	0.837	0.840

Notes: neighbourhood characteristics were standardized; reference Region is London and reference Local Authorities are: Barnet, Rutland, Bristol, Essex, York, Manchester, Birmingham, Surrey, and Newcastle (one from each Region); standard errors are robust to heteroscedasticity and the symbols *, ** and *** denote statistical significance at 10%,5% and 1%.