DEPARTMENT OF ECONOMICS

Approaches to Measuring Schooling Value

BY

SOFIA N. ANDREOU

DISSERTATION

Submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy in Economics in the Department of Economics of the University of Cyprus

April 2011

ΣΕΛΙΔΑ ΕΓΚΥΡΟΤΗΤΑΣ

Υποψήφια Διδάκτορας : Σοφία Ν. Ανδρέου

Τίτλος Διατριβής: Approaches to Measuring Schooling Value

Η παρούσα Διδακτορική Διατριβή εκπονήθηκε στο πλαίσιο των σπουδών για απόκτηση Διδακτορικού Διπλώματος στο Τμήμα Οικονομικών και εγκρίθηκε στις 21 Μαρτίου 2011 από τα μέλη της Εξεταστικής Επιτροπής.

Εξεταστική Επιτροπή

Ερευνητικός Σύμβουλος :

- Πάνος Πασιαρδής, Καθηγητής, Πανεπιστήμιο Κύπρου

Άλλα Μέλη :

- Θεοφάνης Μαμουνέας, Καθηγητής, Πανεπιστήμιο Κύπρου (Πρόεδρος)
- Σωφρόνης Κληρίδης, Αναπληρωτής Καθηγητής, Πανεπιστήμιο Κύπρου
- Ian Walker, Professor, University of Lancaster, UK
- David Brasington, Associate Professor, University of Cincinnati, US

Abstract-In Greek

Η διδακτορική αυτή διατριβή επικεντρώνεται σε τρεις διαφορετικές έρευνες.

Η πρώτη έρευνα μελετά πώς ο νέος δείκτης ποιότητας των σχολείων, γνωστός ως «Contextual Value Added (CVA) Indicator», επηρεάζει τις αξίες των σπιτιών σε περιοχές όπου υπάρχουν σχολεία πρωτοβάθμιας και δευτεροβάθμιας εκπαίδευσης στην Αγγλία. Η εμπειρική ανάλυση στηρίζεται σε αγγλικά στοιχεία τα οποία προέρχονται από τρεις νέες και ανεξάρτητες πηγές οι οποίες δεν έχουν χρησιμοποιηθεί μέχρι τώρα για τη μελέτη του θέματος αυτού. Η ανάλυση δείχνει ότι ο παράγοντας «score» του δείκτη αυτού έχει θετική και στατιστικά σημαντική επίδραση στις τιμές των σπιτιών και στα δύο επίπεδα εκπαίδευσης (πρωτοβάθμια και δευτεροβάθμια), ενώ ο «non-score» παράγοντας του δείκτη έχει σημαντική αλλά αρνητική επίδραση μόνο στην περίπτωση των σχολείων δευτεροβάθμιας εκπαίδευσης. Παρόλα αυτά, η επίδραση του CVA και των επιμέρους στοιχείων του στις τιμές των σπιτιών διαφέρει ανάλογα με το επίπεδο στο οποίο πραγματοποιείται η ανάλυση (level of spatial aggregation), υποθέτοντας περισσότερη θετική επίδραση μεταξύ από ότι εντός των Local Authorities. Η δεύτερη έρευνα προτείνει μία νέα μέθοδο η οποία στηρίζεται στην ανάλυση ζήτησης του καταναλωτή, με στόχο να εξετάσει την εμπειρική σχέση μεταξύ των αξιών των σπιτιών και την πρόσβαση των νοικοκυριών σε ψηλό επίπεδο δημόσιας εκπαίδευσης για τα παιδιά τους. Η προτεινόμενη μέθοδος επιτρέπει τα συνδυασμένα έξοδα για στέγη και εκπαίδευση από νοικοκυριά τα οποία έχουν παιδιά που φοιτούν σε δημόσια σχολεία, να κατανεμηθούν στα επιμέρους μέρη έτσι ώστε η επίδραση της ποιότητας των σχολείων στις τιμές των σπιτιών να εκτιμηθεί από στοιχεία Έρευνας Οικογενειακού Προϋπολογισμού. Η αδυναμία εφαρμογής της «hedonic» ανάλυσης αντιμετωπίζεται με τη χρήση αυτής της μεθόδου όπου δεν χρειάζεται καμία πληροφορία για την περιοχή στην οποία διαμένει το νοικοκυριό, πληροφορία η οποία είναι εμπιστευτική σε έρευνες όπως αυτή του Οικογενειακού Προϋπολογισμού. Η εμπειρική ανάλυση στηρίζεται σε αγγλικά στοιχεία για την περίοδο 1994-1997 και δείχνει ότι τα μη παρατηρούμενα έξοδα που αφορούν την εκπαίδευση έχουν θετική και στατιστικά σημαντική επίδραση στα από κοινού έξοδα για στέγη και εκπαίδευση, με τη σχετική τιμή της εκπαίδευσης-στέγης να κυμαίνεται μεταξύ 0.4-2. Η τρίτη και τελευταία έρευνα έχει ως στόχο την κατασκευή ενός χρηματικού δείκτη ικανοποίησης του οφέλους που απορρέει από την κρατική εκπαίδευση όταν τα νοικοκυριά μπορούν να συμπληρώσουν την ελάχιστη εκπαίδευση που παρέχεται δωρεάν από το κράτος με πρόσθετη εκπαίδευση που αγοράζεται μέσω της απόκτησης στέγης σε περιοχές υψηλής ποιότητας δημόσιων σχολείων. Προτείνει τρόπους για να αντιμετωπιστούν οι δυσκολίες μοντελοποίησης της συμπεριφοράς των νοικοκυριών που προκύπτουν από την κοινή κατανάλωση στέγης και εκπαίδευσης προκειμένου να επιτευχθεί ένα πλήρες και ολοκληρωμένο σύστημα ζήτησης όπου ο χρηματικός δείκτης ικανοποίησης μπορεί να υπολογιστεί από εύκολα διαθέσιμα στοιχεία (και συγκρίσιμα με άλλες χώρες), όπως αυτά από τις Έρευνες Οικογενειακών Προϋπολογισμών. Η εμπειρική ανάλυση, βασισμένη σε αγγλικά στοιχεία για την περίοδο 2001-2007, δείχνει ότι οι οικογένειες με παιδιά στην πρωτοβάθμια εκπαίδευση απολαμβάνουν ένα μεγάλο όφελος από την κρατική παροχή εκπαίδευσης, εντούτοις, το ίδιο δεν φαίνεται να ισχύει για τις οικογένειες με παιδιά που φοιτούν στην δευτεροβάθμια εκπαίδευση. Τέλος, η διατριβή αυτή μπορεί να χρησιμοποιηθεί ως βάση τόσο σε εμπειρικό όσο και σε θεωρητικό επίπεδο για να εξηγήσει/ερμηνεύσει αποτελέσματα σύγχρονων εκπαιδευτικών πολιτικών.

Abstract- In English

This PhD thesis consists of three different parts.

The first part investigates how the newly introduced Contextual Value Added (CVA) indicator of school quality affects house prices in the catchment area of primary and secondary schools in England. The empirical analysis, based on data drawn from three independent and previously unexplored 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 the non-score component of this school quality indicator has a significant negative association with house prices, but only in the analysis of secondary school data. Furthermore, the effect of CVA and its score and non-score components on house prices also varies with the level of spatial aggregation at which empirical investigation is pursued, assuming a more ‘positive’ role between rather than within Local Authorities (LAs).

The second part of the thesis proposes an approach based on the notion of separability and uses a two-stage budgeting consumer demand analysis for the investigation of the empirical relationship between the composite education-and-housing expenditure and the education component in order to estimate the relative price of the education to housing. No information about the location of households is needed; thereby the inability to apply hedonic methods in countries like the UK, where this information is considered confidential is circumvented. The empirical analysis draws on UK FES data over the period 1994-1997 and shows that the education component has a positive and significant effect on the composite education-and housing commodity, with a magnitude of a relative price of two components (education in relation to housing) to range between 0.4 to 2.

The third part of the thesis attempts to construct a money metric of the benefit derived from state schooling when households can supplement the 'minimum' education provided free of charge with additional education purchased through acquiring accommodation in the catchment area of a high quality state school. It suggests ways to circumvent difficulties in modelling household behaviour arising from joint housing-education consumption in order to reach a complete demand system, where the proposed money metric can be estimated from data readily available in household expenditure surveys. The empirical analysis based on UK EFS data over the period 2001-2007, shows that households with children in private education enjoy a large benefit from free state schooling; however, the same is not true for households with children in secondary education.

The present thesis can serve as an empirical but also as a theoretical tool for exploring issues pertaining to current educational policies.

Acknowledgements

My sincere grateful to Professor Panos Pashardes for his continued support and his invaluable guidance throughout my PhD. I gained a great deal of knowledge from him, indeed, more than it is reflected in this thesis.

Further appreciation goes to Nicoletta Pashourtidou for her help and advice throughout my PhD, especially for her guidance on statistical issues related to the thesis. Special thanks go to Adamos Andreou, Marios Charalambous and Georgia Loizide for their help on data collection process for the one of the three datasets used in this thesis. In addition, I would like to thank Costas Hadjiyiannis, and Thanasis Stengos for their valuable comments and advice all these years.

I would also like to express my gratitude to my dearest officemate Neophyta, who was always there for me, giving me her love and support. She provided an attentive audience whenever I felt that I needed somebody to talk to.

Special thanks to all my colleagues for their support and for making these years more enjoyable and meaningful in the Economics Department: Ifigeneia Georgiou, Christiana Ierodiakonou, Elena Xeni, Elena Ketteni, Kikis Agisilaou, Adamos Adamou, Ioanna Stylianou and Constantinos Kourougiannis. Last but not least, Anastasia and Litsia are acknowledged for all their assistance and encouragement all these years.

I am also thankful to my family. My parents, brothers, and extended family offered me unconditional love and support in every step of the way, and for this I am forever grateful.

Finally, I would like to thank Sotiris for his unyielding love, patience, and encouragement throughout the completion of this thesis. There is no way I could have completed this long journey without him.

Dedication

Στους γονείς μου, Νίκο και Ανδρούλλα

Table of Contents

List of Tables

List of Figures

Chapter 1

General Introduction

1.1 Introduction

Education is generally seen as a basic requirement for winning a well-paid job. According to the theory of human capital introduced by Becker (1964), education is an investment that produces knowledge acquisition and increased productivity (see Walker, 2003; Chevalier et al, 2004), which in turn leads to higher income. Schooling exerts a major effect upon earnings or occupational status and is usually determined by the social class background of the individual. More equal education could achieve significantly greater equality of economic opportunity and distribution of income.

The family milieu during early childhood is one of the most important factors for later achievement, such as educational attainment, occupational options, careers and earnings. Towards a more comprehensive economic prospective the attainments of children depend on three primary factors: (i) choices made by society (government) determining the opportunities available to both children and their parents; (ii) choices made by the parents, which are related with the family background and resources and (iii) choices that children make given the investments and opportunities available to them.

In recent years the value of expenditures on schooling and further education has become a source for many economists to explain the relation between education and income. College graduates seem to earn more than high school graduates and high school graduates seem to earn more than high school drop-outs. The rate of return to schooling is an important factor in determining educational attainment and participation and, ultimately, wages and income. There is a vast literature relating to the returns to education which has been the focus of considerable debate in the economics literature. The impact of education on earnings reveals a wide range of estimates and an equally wide range of empirical approaches that have been adopted to estimate the returns. Some of the most important studies includes Harmon and Walker, 1995; Harmon, Oosterbeek and Walker, 2000; Blundell, Dearden and Sianesi, 2001; Trostel, Walker and Woolley, 2002).

On the other hand, the attainments of children are influenced not only by the amount of family resources allocated to them, the nature and the timing of the distribution of these resources but can be also affected by their parents' choices such as the number of their siblings, the number of location moves, the type of neighbourhood in which they grow up, the education background of parents and, generally, family structure and circumstances. Regarding the aforementioned there is also a large number of studies that have looked in depth all the issues above (e.g. Bowles, 1972; Behrman et .al, 1986, 1989, 1997; Hanuskek, 1992; Kalmijn, 1994; Haveman and Wolfe, 1995).

Hence, there are different strands of literature that can be used as basis to examine how the education choices of an individual reflect on her/his economic and social status, and vice versa. The research presented in this thesis focuses primarily on the decision of people/parents regarding the choice for their children's education; and more specifically on relation to whether they favour private or state schooling and how the economic rationality behind their choice can be recovered and analysed from observed data. The implications arising from this research are important for government planning and policy implementation regarding education opportunities available to households. More details, including the insight and motivation of this thesis are discussed later in this chapter, after we briefly outline the main literature strands associated with the its subject matter.

1.2 Brief Literature Review

1.2.1 House Prices and School Quality

Good schooling is commonly upheld as an essential in life, but empirical evidence remains quite unclear when it comes to providing answers about what makes a school a 'good' one or about what people really value in education. In the Tiebout's theory (1956), people face a trade-off between local taxes they have to pay for the quality of public services, and sort themselves into different communities accordingly. These processes of sorting can be observed through the reaction of housing costs to local amenities. Empirical work on Tiebout choice is quite extensive especially in the case of local state school quality, dating back to Oates (1969) seminal paper in which he studied the effect of property tax rates and public school expenditures per pupil on house prices using aggregate (community-level) data on 53 cities in Northern New Jersey in 1960. His methodology was a starting point and influence for numerous studies examining the capitalization of fiscal variables into house values, many of which were quite critical of it (e.g. Pollakowski, 1973; Sonstelie and Portney, 1980[1]).

The recent focus of most applied empirical work on micro housing price models is on a proper estimation of the equilibrium implicit 'prices' that can be attributed to local public goods, neighbourhood /community attributes - including state schooling so as to assess the equilibrium change in housing expenditure in response to changes in these attributes. The relationship between housing expenditure and the quantities of its composite attributes has become known as the hedonic price function (see Rosen (1974) for the classic exposition, or Sheppard (1999) for a modern survey). Even though implicit prices do not provide a basis for useful welfare analysis of policy changes, they can inform policy and provide a useful general guide to the values of locally provided public services.

The recent US literature on the empirical valuation of local school quality is quite extensive. The vast majority of research in the field (e.g. Brasington, 2000, 2002; Haurin and Brasington, 1996, 2006, 2009; Black, 1999; Barrow, 2002; Barrow and Rouse, 2004; Downes and Zabel, 2002; Bogart and Cromwell, 1997, 2000, Clapp et.al, 2008; Kane et al, 2006) has looked for evidence of the value of schools in the capitalization of their benefits into housing prices by basically applying the hedonic[2] valuation method developed by Rosen (1974). Many researchers using data from different metropolitan areas of US, found that positive influences on student achievement and other school quality measures, i.e. expenditure per pupil, are highly valued in the housing market.

Most research in this field has adopted a number of empirical strategies for the examination of the relationship between school quality and house prices, such as traditional regression-based models/hedonic models, instrumental variables approaches, spatial analysis approaches, semi-parametric methods and discontinuity designs using administrative boundaries (Black, 1999[3]). The problems associated with the use of traditional regression models in this context are discussed by Gibbons and Machin (2008).

[1] They were among the first to use disaggregated data on individual housing sales usually obtained from real estate services.

[2] In hedonic methods house prices are estimated as a function of measures reflecting the quality of the school which the house in question have access to, along with several characteristics describing other house attributes such as the number of rooms, size and type.

[3] Clapp et al (2008), criticize the boundary approach on two grounds: i)attendance zones change, ii) capitalization is weaker toward the edge of an urban area.

Not surprisingly, the vast majority of studies on housing prices and education have focused on the US, given its traditionally more decentralized system of government in which more variation in educational spending might be expected as also more accessible data. For the UK there have been relatively few studies as data on the locality of individual households are not available due to confidentiality. Gibbons and Machin (2003) provided the first empirical evidence for the UK on the effect of primary school performance on property prices using postcode sector level data on house prices, income and primary school performance for the whole of England. Other researchers (Cheshire and Sheppard, 1998, 2004; Rosenthal, 2003; Leech and Campos, 2003) have looked at the value house buyers attach to secondary schools.

Gibbons and Machin (2003) estimate (through spatial and semi-parametric approach) the magnitude of the association between primary school quality and local house prices and income during the period 1996-1999. In their study they found a 10 percent increase in performance (10 per cent more students reaching target level) increases property values by 6.9 per cent. Also, their results suggest that parents strongly value high school performance and that any raise of primary school standards has a social valuation per household equivalent to 0.5% to 0.8% of the local mean property prices.

Cheshire and Sheppard (2004) use a traditional hedonic approach including Box-Cox transformations of house prices and attempted to measure a wide range of local neighbourhood characteristics, including the socioeconomic composition of the neighbourhood and other local public goods and localized amenities in the area of Reading during 1999 and 2000. They find that the quality of both local secondary and primary schools is capitalized into house prices, although the statistical significance of the quality of secondary schools is considerably greater. However, the variation in primary school quality is found to be significantly greater than that for secondary schools. So moving from worst to best secondary school would increase average house value by 18.1 per cent, whereas the corresponding increase for primary school is as high as 33.5 per cent.

Instead of focusing on a single housing market - as Cheshire and Sheppard (2004) - the study by Rosenthal (2003) examines an application of the hedonic price method for the evaluation of school quality in English secondary schools, using a substantial amount of data for all England drawn from a number of different sources for the years 1996-1998. Particular emphasis on estimation is placed on using instrumental variable techniques as means to avoid potential problems associated with omitted (simultaneity) bias. Results on

2SLS procedures, using school inspections as instruments, show that the elasticity of purchase price of houses with respect to secondary school quality in England is around 0.05.

Leech and Campos (2003) also investigate whether house prices are higher in the catchment areas of two popular secondary schools in the local area of Coventry than other areas. As Cheshire and Sheppard (2004), they use advertised list prices as a measure of house prices and similarly to Black (1999) they focus on the price differentials within blocks between the two popular schools. They employ a different approach to minimize unobserved neighbourhood differences, focusing on the effect of LEA's admission arrangements in the face of school's popularity and not the source of popularity. They claim that a catchment area premium on house prices is consistent with different scenarios. It could be that parents make out that the school provides a high standard of education because of good teaching. Alternatively, it could be that the school is popular, because it does well in the exam league tables by serving largely educationally advantaged students from middle class backgrounds; thus parents seek to join this group of students for the educational advantages it brings. They test for such effects using a sample based on two popular comprehensive schools in one local authority at the Coventry area. They compare prices of houses for sale in their catchments areas with those of neighbouring schools using techniques to allow for differences in the house size and quality. Their results suggest that there are strong school catchment area effects on house prices, with a 20% premium due to the first popular school and a 16% premium due to the other school.

1.2.2 Measures of School Quality

An extensive literature has developed on the relationship between inputs in the production of education (student-teacher ratios, teacher education, expenditures) and measures of educational “outputs”, typically standardized tests. Nevertheless, many of these studies with probably the best known work accomplished by Hanushek (1986), have found little support that increases in educational inputs, including expenditure, have much impact on standardized tests. Researchers in education and sociology of education suggest that, although parents rank academic outcomes highly among the reasons for choosing a school, there are many other factors that play an important role, such as school composition and the child's potential wellbeing at school. In the recent US literature, the measures used to capture school quality vary from expenditure per pupil, pupil/teacher ratio, reading scores, proficiency test scores, while measures like proportion of students achieving a specific target level had used in a small number of studies available to date for the UK. Rosen and Fullerton (1977) were the first who tried to follow Oates (1969), replacing educational expenditures with one of several test scores (mean reading and maths) in the hedonic equation. They find a significant relationship between property values and educational outputs while the educational expenditure appears to be insignificant in the hedonic equation.

Following Rosen and Fullerton (1977) a number of studies have included educational test scores/ standardized scores as explanatory variables. However, many education economists (e.g Mayer, 1997) claim that this kind of school quality measures tend to be infected by non-school factors that contribute to student achievement (e.g. student characteristics, family, socioeconomic background, prior achievement, community characteristics, teaching quality or resources of school) and thus cannot serve as a proper measure of school quality. Instead, they suggest that growth over time in student achievements or "value added" is the most appropriate measure of school quality.

The term "value added" in education has usually been used to describe the additional value schools bring to the learning outcomes of their students; in other words, the contribution a school makes to the learning of students. Many researchers in the area of education have adopted the value added approach in order to measure the gain in student tests results (e.g. Hanushek and Taylor, 1990; Hanushek 1992). The proponents of value-added modelling consider its results 'fairer' and more accurate than those produced by standardized tests measuring achievement only, because socioeconomic factors may greatly affect the results of those tests. Mayer (1997) argues that the results of standardized tests[4] and assessments can be used in a variety of different applications: to measure the achievement of individual students, to produce aggregate indicators of the level and distribution of achievement for groups of students, to evaluate the efficacy of specific school policies and inputs; but not as a measure of school performance/effectiveness.

There are several different models that can be used to calculate a value-added measure. Each model has a set of assumptions that need to be made explicit to reflect the intention of the measure. The approaches usually taken in value-added models are determined by the

[4] Mayer (1997) refers generally to aggregate statistics of this kind as level indicators, because they measure some feature of the level of student achievement.

data available and the intended use of the value-added measures[5]. Until very recently, UK is the only country where there is a national value-added system. The establishment of a national curriculum and new sources of linked national data enabled consistent value added models to be established. For the US, value-added measures have been introduced in states and districts such as Tennessee, North Carolina, South Carolina, Ohio, Pennsylvania, Arkansas, Minnesota and Dallas (Texas). McCaffrey et al (2004) note that value-added approaches have not been widely adopted in other states, partly because they require high-quality longitudinal data which many states and districts do not have.

Very little has been done in the housing market literature of US to test the effect of value added measures because these are not easily available to the public. Brasington (1999), Downes and Zabel (2002), and Brasington and Haurin (2006) find little support for using value added school quality measures in the capitalization model; they basically suggest that home buyers favour more traditional measures of school quality in their housing valuations.

1.2.3 State vs Private Education Choice

Parents in both US and UK locate themselves in areas where state schools can provide high quality education for their children. Access to such education is provided by houses located in a short distance from a high quality state school, because most local education authorities base admission on whether candidates live in the schools catchment area. For this reason demand and price for houses in the catchment area of high quality state schools is relatively high, as parents are willing to pay a premium to secure a place in these schools for their children. However, demand for high quality education by households can be observed not only through their choice to locate themselves in areas where there are available high quality state schools but also by choosing to send their children in private schools. It is well accepted that parents choosing private schools tend to be those who are very supportive of education and want to send their children to high quality schools. Thus, those parents with strong preferences for private schools may choose to locate themselves in areas with low public school expenditure (see Goldhaber, 1999).

[5] While value-added information is a powerful tool for analysing school performance, it is a relative measure. It is only effective when seen in combination with other factors such as test results, teacher assessments of student progress, school self-evaluation, reviewers' judgements and the profile of the school (Saunders, 1999; Downes and Vindurampulle, 2007; Ray, 2006).

At the same time, the choice between state and private schooling is a widely researched issue in the literature, where authors are primarily concerned with factors determining the outcome of this choice (e.g. Buttin, 1998). Questions like 'why households choose private schools' or 'who chooses or can choose a private school' are examined in the recent empirical literature. Family characteristics, like income/wealth tastes for education, socioeconomic background are found to be important factors in the private/state school choice decision of households. Also, other factors like race and ethnic background are also found to be important determinants of this choice (e.g. Buttin, 1998; Lankford, Lee and Wyckoff, 1995; Lankford and Wyckoff, 1992; Long and Toma,1988). These empirical studies use microdata to analyze observed choices of state versus private schools by individual households. The results can then be used to consistently estimate the effects of policy changes, such as the introduction of vouchers (i.e. Epple and Romano, 1998)[6].

Epple and Romano (1996b) suggested that when a private schooling alternative is available, high income households prefer low public expenditure locations because private-market purchase costs them less per unit than public provision. The same is also true for low income households who prefer low public expenditure locations because they are less willing to substitute education expenditure for other goods than the higher income households; while middle income households are more willing to substitute education expenditures for other expenditure because they find public provision to be less costly than private provision. So, all households with income below the mean should prefer some positive level of public provision, but the highest level of public provision will be desired by households with incomes that come near, but are still below the mean.

All schools in a private sector vary in quality and thus they provide an opportunity for households to obtain the desired level of quality by paying different educational fees. As described above households will choose a private school if the net benefits from obtaining the desired quality exceeds the extra cost. For households demanding quality smaller than that supplied by the state schools, the cost of private schooling will always exceed the net benefits, since benefits are negative and the cost positive. Finally, the net benefits may

[6] Sonstelie (1982) and Hamilton and Macauley (1991) built a theoretical model for a household facing a choice between public and private schools, without having to measure achievements or other dimensions of school quality.

outweigh the extra costs for those demanding a quality larger than that supplied by state schools (Barzel, 1973). In any other case, state schools provide education of a given quality and charge no tuition. Hence, a family may change the quality of its state school by moving to another community. In that case, a family supplements the free of charge minimum state education with higher quality state education paid through purchasing relatively more expensive accommodation in the catchment area of a high performing state school.

1.2.4 Public Provision of Private Goods: The Case of Education

A pure public good is conventionally defined as a good which has two characteristics: (a) the marginal cost of an additional person consuming is zero; (b) the cost of excluding an individual from its benefits is infinite. In addition, a *pure publicly provided private good* is a private good (a good for which there is a substantial marginal cost of an additional person consuming it) which is provided in equal quantities to all individuals without charge (Stiglitz, 1974). Education is the most important public good in which both public and private provisions play an important role. Several reasons for providing such goods publicly can be identified; high costs of exclusion or ethical (distributional) grounds are among the most important ones.

The state education has traditionally been known for its redistributive[7] effects and as a major contributor to inequality (Stiglitz, 1974). Education is thought not only to be a very important source of economic growth, but also a great social equaliser. In many societies in the past, it was felt that inequality of outcomes was largely the result of inequality of opportunity in schooling. According to Behrman et.al (1980), education is one of the major factors affecting the degree of income inequality. Their findings indicated that educational factors (higher attainment and more equal distribution of education) play an important role in changing income inequality.

A theoretical model by Besley and Coate (1991)[8] show that the quasi-private good being provided could vary in quality. They suggest that universal provision schemes can

---

[7] Governments pursue redistributive goals through a wide variety of instruments including taxes, transfers and public expenditures

[8] Levy (2005), expanded the work of Besley and Coate (1991) assuming that governments engage in education provision in order to redistribute recourses (from the old to the young).

redistribute income from the rich to the poor (this also involves deadweight loss), even if they are financed by a head tax. In addition, they show that uniform provision by the public sector financed by proportional taxation could improve social welfare. They argue that it is inefficient to provide goods with high quality and therefore it is more efficient to provide goods with low quality. This way the poor get the amount they want and the rich get nothing.

An important reason for the provision of free of charge education (and other private goods) by the state is the guarantee of minimum consumption by economically deprived social groups. However, it is not obvious in what quantity and/or quality these goods should be provided; and how this provision should be differentiated to reflect on consumer preferences/needs. Furthermore, it is not clear how much the free provision of private goods by the state adds to the welfare of the consumer. Apart from the implementation of hedonic analysis methods and contingent valuation surveys, to the best of our knowledge, there are no consumer valuations of the public provision of private goods - such as education - in the literature based on economic theory tools. A few studies attempt to estimate how much value consumers attach to publicly provided private goods through the use of an expensive and difficulty conducted contingent valuation surveys (Brookshire and Coursey, 1987; Clinch and Murphy, 2001; Hanemann, 1994). Nevertheless, there are no studies available on consumer valuations of the publicly provided education based on the use of a complete consumer demand system derived from utility maximisation theory.

1.3 Aims of the Thesis

The main aim of this thesis is to explore alternative approaches to measuring schooling value building on different theoretical models and implementing different datasets for the case of UK.

The first issue addressed by this thesis is the investigation of the effect of a newly introduced contextual value-added indicator of school performance on house prices for primary and secondary schools in England. A value-added indicator is argued by many researchers in education to be a more appropriate measure of school performance in comparison with the final achievement indices that are widely used in the literature. This argument soon found its way to house price regressions, with several authors speculating whether value added indicators should be used as a measure of school quality in hedonic analysis of house prices. The empirical analysis in Chapter 2 has been based on data drawn from three independent and previously unexplored UK data sources which have been collected during 2008 for the purposes of this study due to the fact that locality of individual households in not publicly available for confidentiality reasons.

The second aim of this thesis, Chapter 3, is to propose a method circumventing the problem of confidentiality of the location of an individual household, by estimating the capitalization of local state school quality to house prices through consumer demand analysis theory using publicly available data, such as those in Family Expenditure Surveys. This approach is motivated by the argument that demand for high quality education can be observed not only from house prices in the catchment area of high quality state schools but also from the education expenditure of households choosing to have their children in private schools. The theoretical model used the notion of separability and two-stage budgeting consumer demand analysis.

The purpose of the last study, Chapter 4, is to model the economic valuation of freely provided state education in the context of a complete demand system satisfying the fundamental principles of economic theory. This study constructs a money metric of the benefit derived from state schooling when households can supplement the minimum education provided free of charge with additional education purchased through acquiring accommodation in the catchment area of a high quality state school. It also suggested ways to circumvent difficulties in modelling household behaviour arising from joint housing-education consumption. The empirical analysis is based on data drawn from the UK Expenditure and Food Surveys for the years 2001-2007.

The final section of this thesis, Chapter 5, provides a general discussion of the outcomes of the three studies and suggests ways that these could be implemented for research in different countries as well as their potential implications for policy making in education.

Chapter 2

House Prices and School Quality: The Impact of Score and Non-score Components of Contextual Value-Added in the UK

2.1 Introduction

This chapter examines how contextual value-added (CVA) indicators of school performance impact on house prices focusing on separating the effects of score and non-score components: the score component includes pupils' final achievements while the non-score component includes pupils' prior achievement and general pupil specific characteristics. This separation is important because these two components: (i) can affect house prices in opposite direction, thereby obscuring the overall effect of CVA; and (ii) convey different information about private and social preferences that have distinct policy implications. This study is the first to explore the effect of CVA and its components on house prices for England at different schooling levels (primary and secondary schools). We use a unique dataset that covers regions throughout England and constructed (during 2008) by combining three independent data sources, solely for the purposes of this research. The empirical investigation also performs semi-parametric in addition to the usual parametric hedonic analysis.

The capitalization of state school quality to house prices has been an 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), probably due to the fact that the locality of individual households is not available in the data due to confidentiality.

Using a hedonic approach (Rosen, 1974) 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) – the most commonly employed are reading scores, proficiency test scores and other measures emphasising final academic achievement (Gibbons and Machin, 2003, 2008; Haurin and Brasington, 2006; Black, 1999; Rosenthal, 2003). These measures are consistently found to be capitalized in housing 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, and 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 (crime, poverty, unemployment) can exaggerate the positive relationship between house prices and high school scores. Local authority policies (property taxation, provision of public goods) can also interfere with the same relationship (Black, 1999). This concern about over-emphasising the importance of test score indicators on house prices echoes education economists criticising these indicators for being inappropriate measures of school quality because they capture not only the contribution of the school but also the contribution of individual and family characteristics and other exogenous variables, including the socioeconomic background of the pupil. A value-added indicator, measuring the distinct contribution of the school to pupil's academic progress, is argued to be a more appropriate measure 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 value added 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, Machin and Silva (2009) using UK data found that simple value added and final score indicators both had a positive and significant effect on house prices. In contrast, Downes and Zabel (2002), using data from the Chicago metropolitan area found that only final score was significant; a result supported by Brasington (1999). Furthermore, Brasington and Haurin (2006) found that while the effect of value added on house prices was positive when used on its own, it becomes negative when test scores was 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. For instance, the CVA indicator used in the empirical analysis of this study 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. 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'[9].

The fact that CVA indicators combine final score with other components reflecting the extent to which a school operates in a deprived socioeconomic environment can be the reason behind the contradictory results about the effect of value-added house prices reported in the literature. This is because these two components of CVA 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. More importantly, by emphasising the difference in 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. 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.

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

The empirical investigation in this study also applies semi-parametric analysis in addition to a usual parametric estimation. The fact that school performance is measured by arbitrarily normalised indices is customary in empirical application to (re)normalised them to measure standard deviations from the mean. This renders semi-parametric analysis an essential tool for exploring higher order effects on house prices, given that including square and cubic standard deviations in the hedonic regression is meaningless. 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. Yet, as we shall see in the empirical analysis in this study, the relationship between school quality indicators and house prices may contain non-linearity that merits investigation.

The data are constructed using three independent 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 A.

The chapter 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 chapter.

2.2 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 value-added, 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, value-added can be also conditioned on prior achievement, denoted by A and defined as the final score achieved prior to the period over which value-added is measured. In this case the value-added can be written as $V = E\{(X - A)|Z,Y\}$ or, more generally, $V = E\{X|(A,Z,Y)\}$.[10]

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

$$P_{si} = a + \beta V_s + \Sigma_k \gamma_k Q_{ki} + \Sigma_m \varepsilon_m Y_{ms} + u_{si} \tag{2.1}$$

where: $Q_{ki},\ k = 1, \dots ., K$ is the vector of house-specific variables (size, type etc) and $Y_{ms},\ m = 1, \dots ., M$ the vector of neighbourhood-specific variables affecting house prices; $a, \beta, \gamma_k, all\ k = 1, \dots ., K$, and $\varepsilon_m, all\ m = 1, \dots ., M$ 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 (2.1) 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,

$$V_s = X_s + bA_s + \Sigma_j d_j Z_{js} + \Sigma_m e_m Y_{ms} \tag{2.2}$$

where $d_j, all\ j = 1, \dots, J$, and $e_m\ all\ m = 1, \dots, M$ are some known parameters. Replacing (2.2) in (2.1) we obtain the *reduced* form hedonic equation

$$P_{is} = a_i + \beta X_s + \theta A_s + \Sigma_j \delta_j \mathrm{Z_{js}} + \Sigma_m \varphi_m Y_{ms} + \Sigma_k \gamma_k Q_{ki} + u_{is} \tag{2.3}$$

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

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

As said in the introduction, CVA indicators discriminate in favour of schools operating under disadvantageous conditions (low prior achievement, poor socio-economic background, ethically heterogeneous classes, poor attendance etc), effectively awarding higher marks to schools that achieve a given final score in circumstances non-conducive to learning. In the context of equation (2.3) 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 (2.3) 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 other variables in (2.3), 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[11] in the vector Y_{ms}, is also ambiguous: assuming that these variables measure deprivation, then the parameters e_m in (2.2) will be positive (achieving a given final score in deprived neighbourhoods increases value-added); 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 value added βe_{m} - dominates.

The discussion above assumes that one knows how the contextual value was constructed, how it can be decomposed and how its individual elements can be used as variables in the house price equation. In practice, of course, the CVA indicators of school quality which are available for empirical analysis would normally be in the form of an index representing the outcome of complex quantitative and qualitative manipulations of final and prior score and/or other variables. Therefore, it would not be possible to identify the component effects of a CVA indicator, by estimating a reduced form equation like (3). Indeed, this is

[11] One could also include distance from London, however the fact that we use LA's dummy variables can be though as also reflecting distance from London.

the case with the English data used in the empirical analysis in this study. This empirical limitation necessitates modification of the theoretical analysis described above as follows.

The focus of investigation in this study is to compare the effect of final score with that of other components of the CVA indicator. We therefore defined CVA as an unknown function of final score, prior achievement and a range of pupil-specific characteristics.[12] We denote this value-added 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 V_s on X_s to obtain $V_s^* = E(V_s|X_s)$, i.e. make the CVA indicator orthogonal to the final score[13]. Then, estimating the house price equation

$$P_{si} = a + \mu X_s + \rho V_s^* + \Sigma_k \gamma_k Q_{ki} + \Sigma_m \varepsilon_m Y_{ms} + u_{si} \qquad (2.4)$$

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

In the empirical analysis that follows we use equations (2.1) and (2.4) 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.

2.3 Empirical analysis

2.3.1 Data

The postal address of households participating in official UK data (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 study, which were drawn from various sources. A detailed description is given in Appendix A3 of this chapter.

---

12 The contextual value added used in our analysis did not take into account the impact of neighbourhood characteristics on school performance.

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

The individual house price data have been collected during 2008 from the electronic site "Up my Street" and, in addition to prices, P_{si}, include number of bedrooms, number of total rooms, type of the house, postal code and, in general, the house-specific variables denoted by the vector Q_{ki}, $k = 1, \dots, K$ in equation (2.4). 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[14].

- 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%[15].
- 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 secondary education; in our sample this proportion is, on average, around 47%.

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 are constructed by adjusting the score indicator described above to take into account pupils' prior achievement. The more complex CVA indicator used in this study, 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

[14] Appendix A1 presents details on the National Curriculum and Key Stage tests. In addition, it describes the main four types of mainstream state schools in UK.

[15] There is no concern about a censored regression analysis because none of the schools has 100% achievement proportion. Only 5% of the sample is above 95%.

for secondary schools[16]. Details about the calculation methods and the range of background factors used in construction of CVA by the Department of Education are given in Appendix A2 of the chapter.

Data on deprivations indices and other neighbourhood characteristics, denoted by the vector Y_{ms}, $m = 1, \dots, M$ in (2.4), come from the Office of National Statistics of UK. All data were collected between the periods 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.[17].

2.3.2 Semi-Parametric Analysis

The CVA indicator, as seen 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 to circumvent problems of comparison and interpretation most empirical analysis in the literature is conducted by (re)normalising 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 are meaningless. In this context semi-parametric analysis becomes an essential tool for investigating non-linear effects of school quality on house prices and, thereby, finding appropriate ways to specify these effects in parametric analysis. The semi-parametric estimator used in this study was the 'nearest neighbour' one proposed by Estes and Honore (1995)[18] briefly described as follows.

We write equation (2.1) as

---

[16] 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.

[17] 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 did not appear to 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.

[18] 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 droped 2% of the sample from each end of the distribution.

$$P_{si} = a + f(V_s) + \Sigma_k \gamma_k Q_{ki} + \Sigma_m \varepsilon_m Y_{ms} + u_{si} \quad (2.5)$$

where $f(V_s)$ is an unknown function, while all variables and notation in (2.5) are as defined in (2.1). 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

$$\Delta P_{si} = \Sigma_k \gamma_k \Delta Q_{ki} + \Sigma_m \varepsilon_k \Delta Y_{ms} + u_{si} \quad (2.6)$$

ΔP_{si} measures the difference of the house price of the current observation with the house price of the observation that has V_s closest to the current. Note that, as the data were sorted by V_s 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 (2.6), we compute the part of P_{si} not explained by the right hand side variables,

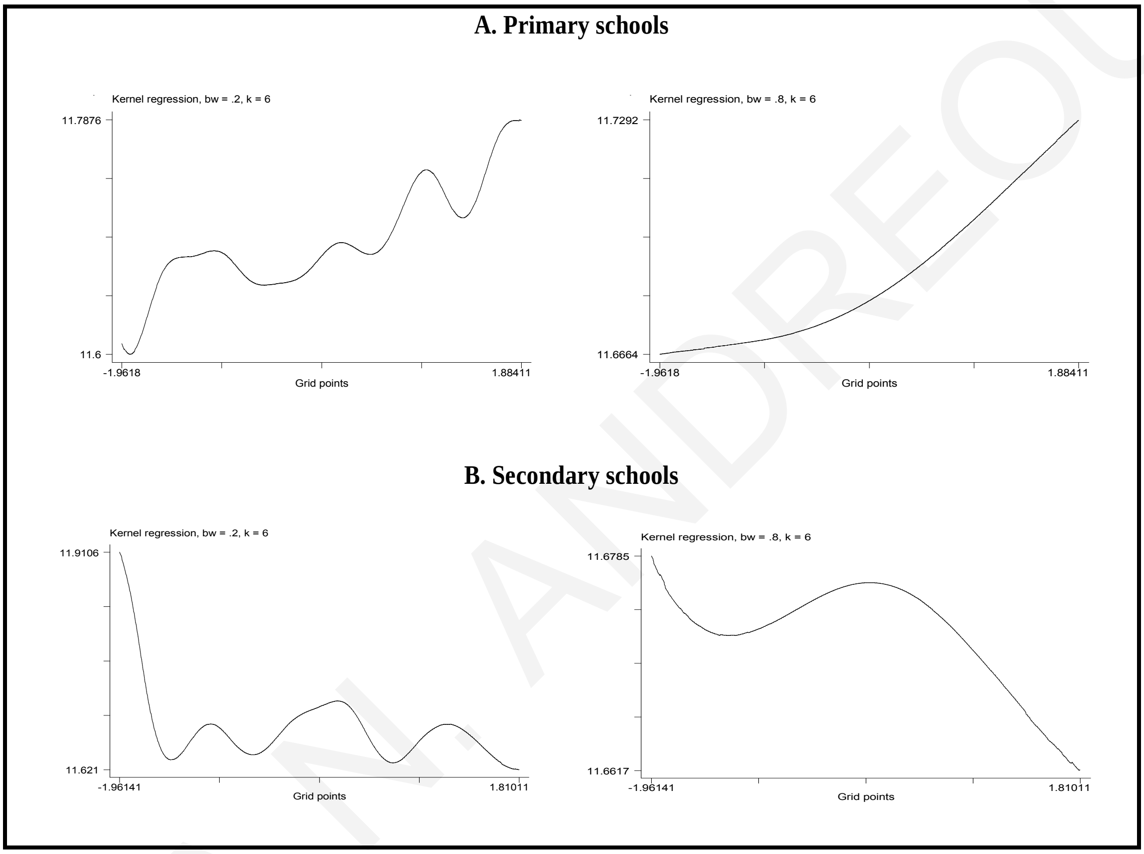
$$\hat{r}_{si} = P_{si} - \Sigma_k \hat{\gamma}_k \Delta Q_{ki} - \Sigma_m \hat{\varepsilon}_k \Delta Y_{ms} \quad (2.7)$$

and performed 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 2.1 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 bandwidths employed, 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.

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 2.1. For this we perform semi-parametric regression of $\hat{r}_{si}$ on the score (X_s) and non-score (V_s^*), following the same nearest neighbour estimator described above.[19] The Gaussian kernel weighted estimates obtained from these regressions (again, using two bandwidths, 0.2 and 0.8) are plotted in Figure 2.2.

Figure 2.1: Kernel estimates for CVA

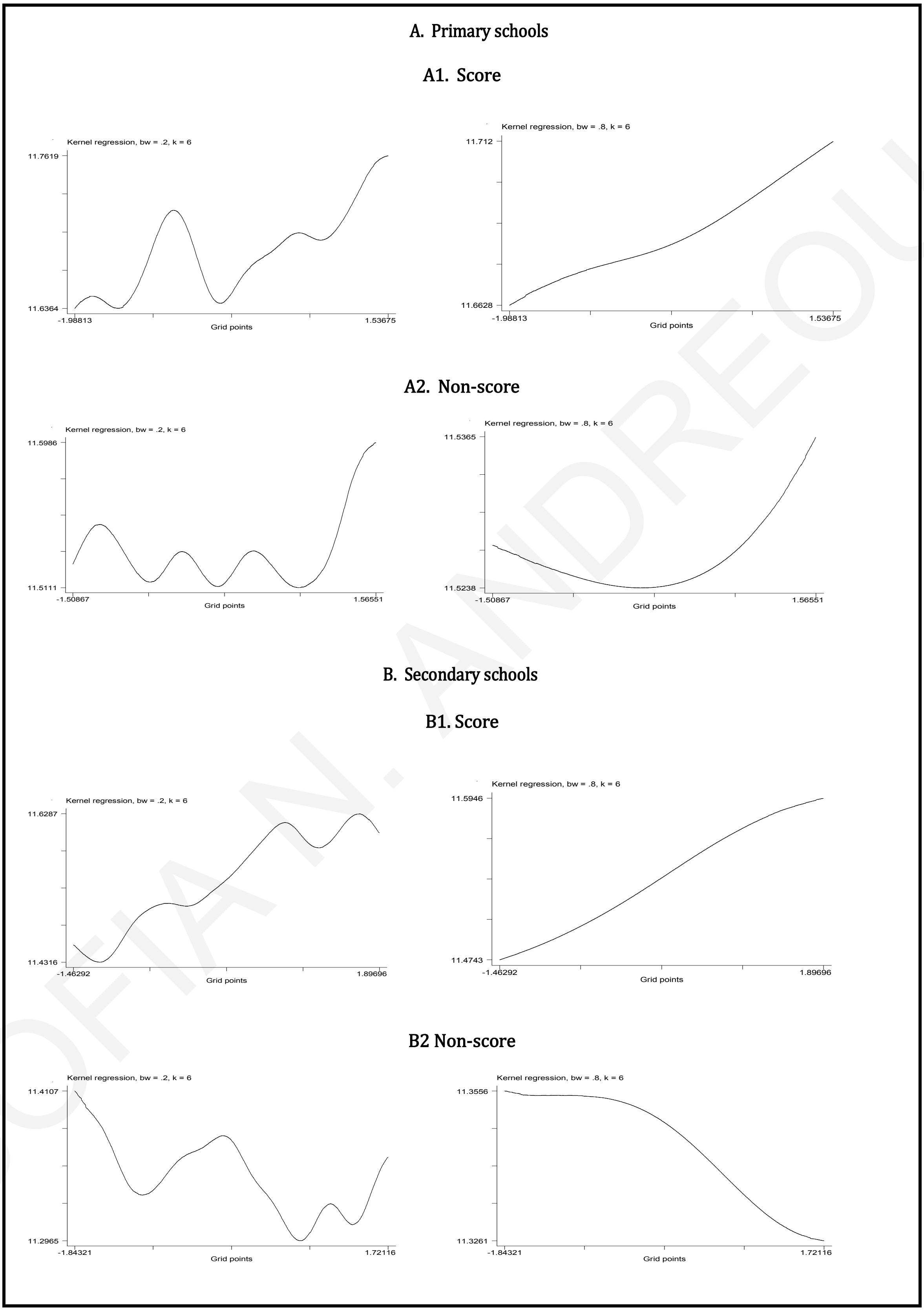


The semi-parametric effects of the score and non-score components of CVA on log house prices for primary schools are reported in Parts A1 and A2 of Figure 2.2, respectively. Here the plots show that score has a positive effect on house prices, even though the plot obtained with the smaller (0.2) bandwidth appears to be interrupted for middle values of this CVA component. The plots for the non-score component show that no clear effect on log house prices can be traced from the semi-parametric results. If such an effect exists in the data is likely to be positive and present only at large values of non-score. Put together, these results suggest that the positive effect of CVA on log house prices shown for primary schools in Part A of Figure 1 is primarily attributed to its score component.

---

[19] 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 2.2: Kernel estimates for the score and non-score CVA components



The results obtained from semi-parametric analysis of secondary school data, shown in Part B2 of Figure 2.2, 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, but not as strong as the one between score and log house prices.

Nevertheless, the positive effect of score and the negative effect of non-score on house prices are probably the reason why the CVA appears to have no effect on house price in the case of secondary schools (Figure 2.1, Part B).

2.3.3 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/or the same across the range of its values. As explained earlier the existence of non-linear and non-monotonic effects of CVA on house prices cannot be investigated using higher order (quadratic or cubic) terms, because the CVA (and its components) is measured in standard deviations from the mean, i.e. include negative and positive values.

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 $\ell = 1 \ldots L$, in the hedonic regression

$$P_{si} = a + \sum_{\ell} \beta_{\ell} (V_s D_\ell) + \Sigma_k \gamma_k Q_{ki} + \Sigma_m \varepsilon_m Y_{ms} + u_{si} \qquad (2.1')$$

For instance, estimating equation (2.1’) for primary education, we create and include in (2.1’) two dummy variables (i.e. $\ell = 2$): 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 2.1) 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 2.1. In the case of secondary schools we create and include in (2.1’) three dummy variables (i.e. $\ell = 3$): 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. 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 were estimated from the hedonic regression

$$P_{si} = a + \mu X_s + \sum_{\ell} \rho_{\ell} \; (V_s^* D_\ell) + \Sigma_k \gamma_k Q_{ki} + \Sigma_m \varepsilon_m Y_{ms} + u_{si} \qquad (2.4')$$

For primary education, we create and include in (2.4') two dummy variables ($\ell = 2$) to allow the effect of non-score V_s^* to differ for values below and above the minimum attained (with a bandwidth 0.8) in the semi-parametric relationship in the corresponding plot in Figure 2, Part A2. Two dummy variables ($\ell = 2$) are also created and included in (2.4') for secondary education, in this case to investigate the semi-parametric finding that the relationship between non-score and log house prices was weaker for small values of non-score, as shown by the plot in Figure 2.2, Part B2.

The full regression results obtained from the estimation of (2.1') and (2.4') are reported in Appendix A4. 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 (2.1') 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.

In the column under the heading 'Model II' are the parameter estimates of (2.4'), where the non-score component of CVA is allowed to vary as suggested by the non-parametric analysis. The results here 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 2.1 reported in the columns under the headings 'Model III and 'Model IV' correspond to the estimation of (2.1') and (2.4') 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 always 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[20].

'Model IV' is also estimated by 2SLS (last column of Table 2.1) in order to circumvent the problem of potential endogeneity and measurement error of the score component of CVA. As Gibbons & Machine (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 that schools in these neighbourhoods perform better because their pupils are more receptive to education[21]. In addition, results published in the national tables (a single year measure) are argued to be noisy measures of long-run school quality.

As in Gibbons & Machine (2003), we investigate the endogeneity and/or measurement error problems by 2-SLS methods, 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 Machine (2003) indicating that errors in the measurement of the score variable may be a more serious source of bias than unaccounted endogeneity from neighbourhood quality effects on school performance[22].

[20] 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.

[21] Children in schools located in richer, higher-house-price neighbourhood are doing better academically than children elsewhere (Blanden and Machin, 2004).

[22] The validity of the instruments was supported by the high F-statistics (68.5 for primary and 75.8 for secondary schools) obtained from the first stage prediction regressions for Model IV in two datasets. See Table A4.1 in Appendix A4. Furthermore, over-identification tests (Sargan test) based on the R-square obtained from regressing the predicted errors from the 2-SLS estimation on all exogenous variables suggested that all models are just identified.

Table 2.1: 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***	0.060***
			(0.010)		(0.010)	(0.021)
Non-Score	ρ				-0.043***	-0.045***
					(0.010)	(0.010)
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.

It is worth noting here that one can consider the use of CVA and, specifically, its non-score component as an instrument towards tackling the endogeneity of final score, a problem that has received big attention in the literature (Gibbons and Machin, 2008). We investigate this in the case of the primary schools where the non-score component of CVA appears to be exogenous (Model II in Table 2.1A) by including this variable, along with the school type, the admissions age-range and the student gender, as an instrument for the final score. The results remain unchanged, i.e. the score coefficient is 0.079*** as is when the non-score component of CVA is not included in the instruments (Model IV in Table 2.1A), suggesting that the additional instrument does not add to identification.[23]

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. When separating the overall effect of the CVA indicator into score and non-score components, it becomes evident that its positive effect on house prices in primary education is entirely attributed to the score component; whereas the negative effect in secondary education is due to the large and significant negative effect of the non-score component that more than compensates the equally significant but not so large effect of the score component.

The estimates reported in Table 2.1 are conditional on Local Authorities (LA) – see Appendix A4. 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 'across' LAs we re-estimate the parsimonious versions of (2.1') and (2.4'), i.e. the versions without dummies for changing slopes, removing the LA-specific effects from the regression. The results of these estimations are reported in Table 2.2.

Commenting on the results for primary schools in Table 2.2, one can say that the between LAs effect of CVA on house prices is positive and significant, as is within LAs. However, the estimated parameter now is larger in size and significance compared to that estimated within LAs. Furthermore, the enhanced CVA effect between LAs comes from equal in size and significance contributions from both the score and non-score components; unlike the

---

[23] In the case of the secondary schools the non-score component of CVA is not exogenous and should not be included in the instruments - Model IV in Table 2.1B).

CVA effect estimated conditional on LA, which is entirely attributed to score. This means that households are willing to pay a higher price for a house in the catchment area of primary schools with both higher non-score and score component of CVA in different LAs; 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 2.2: Estimated parameters when Local Authorities are excluded from the regression (robust standard errors in brackets)

Variable	Parameter	Primary schools		Secondary schools	
CVA	β	0.039*** (0.010)		-0.007 (0.010)	
Score	μ		0.030*** (0.010)		0.030*** (0.011)
Non-Score	ρ		0.031*** (0.012)		-0.017 (0.010)
R-squared		0.747	0.747	0.747	0.749
No. of observations		1385	1385	1209	1209

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

The parameters for secondary schools reported in Table 2.2, again, change substantially compared to those obtained conditional on LA-specific effects. The changes are in the same direction as those for primary schools, in the sense that the non-score component of CVA appear to have a less negative effect on house prices between rather than within LAs. 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, while the results in Table 2.1 suggested that within LAs there is a price reduction for houses in catchment areas of secondary schools with a high CVA, across LAs no significant price differences exist between houses in catchment areas of secondary schools with a different CVA.

2.4 Discussion and Conclusions

The relationship between housing prices and education has been a topic of interest among urban economists and economists examining educational issues more generally. Evidence on the impact of educational output, as measured by student test scores on standardized tests, on house values is consistently positive and often represents a substantial component of the house value (e.g. Black, 1999; Gibbons and Machin, 2003; Haurin and Brasington, 2006; Rosenthal, 2003). These measures of educational output do appear to matter in both the USA and UK suggesting that families appear to pay attention to these measures and are in fact willing to pay for higher performance. In addition, a number of studies have included both measures of educational inputs (usually expenditure per pupil) and educational output (test scores) as explanatory variables in order to determine which measure seems to have more of an effect on property values (e.g. Haurin and Brasington, 1996; Brasington, 1999; Black,1999; Downes and Zabel, 1992). However as already mentioned earlier in this chapter, education economists (Mayer, 1997; Hanushek and Taylor, 1990; Summers and Wolfe, 1977) argue that growth over time in student achievement or 'value added' can be served as the most appropriate measure of school quality. Authors, mainly in the US, took up the challenge of investigating this argument empirically using various measures of value added as measures of school quality in hedonic analysis of house prices (Hayes and Taylor, 1996; Brasington, 1999; Haurin and Brasington, 2006; Downes and Zabel, 2002). The findings of these studies are briefly described below.

Hayes and Taylor (1996), using a small number (288 observations) of Dallas data, employes a two-stage procedure in which they attempt to distinguish how much of the difference in achievement scores, is due to a school effect and how much is due to the demographic composition of the students, the peer group effect. In order to distinguish the school and peer group effect, they follow the approach described in Hanushek and Taylor (1990) in the context of which they estimate an equation using data on individual students' test performance, where demographic characteristics of the student and his/her household and a dummy variable for the school that the student attended are explanatory variables. They find that, while spending has no effect on property values, test scores do. Further separating the test score into the school and peer effects, Hayes and Taylor (1996) suggest that the impact on property values is due to the school effect (termed value added), not the peer effect.

Brasington (1999) explicitly consider a number of measures of educational inputs and outputs in examining the impact of education on single family house sales price for the six largest metropolitan areas of Ohio in 1991. However, his measure of value-added was quite different from that of Hayes and Taylor (1996): he measures value-added as the percentage change in the number of proficient students in the district between grades.

Using a variety of formulations of value added (calculation of different relative performance in terms of grades, i.e. fourth, ninth and twelve[24], separately for reading, maths, science, and writing subjects) and performing both a traditional and a spatial analysis he finds little evidence of capitalization. Specifically the effects of value added on house price vary between positive, negative or no effect depending on the metropolitan area. However, he does find that several of the absolute test score measures, expenditures per pupil and teacher salary have a positive effect on house price.

Downes and Zabel (2002) use a sample of 1,173 house price observations in the Chicago metropolitan area to test alternative models of the impact of school quality on house prices. Their measure of value added was an eighth grade proficiency test, holding constant sixth grade proficiency test results from 2 years prior. In contrast to Hayes and Taylor (1996), they find that higher average levels of school achievement raise house values, but their measure of a school district's value added does not. They claim that even if value added is the theoretically preferred measure, what is important is the attribute of school quality that is appreciated by households.

Last, Brasington and Haurin (2006) measure the amount of capitalization of various schooling measures into 77.000 houses sold in 2000 in Ohio and find strong support for the capitalization of student proficiency tests and school expenditure onto house values. Using a variety of formulations of value added measures (like that of Brasington (1999), Hayes and Taylor (1996) and Downes and Zabel (2002)), however, they find little evidence of capitalization. Specifically, the effect of value added measure (calculated as in Brasington, 1999) is found to be negative and significant in the hedonic equation when controlling for other measures of school quality, like ninth grade proficiency tests and expenditure per pupil. In addition, the effect of value added measure (calculated as Hayes and Taylor (1996) proposed) appears in most cases positive when only the value added measure is controlled for.

[24] Pupils are between ages 14-15 and 17-18 for ninth and twelve grade levels respectively.

For the case of UK, Gibbons, Machin and Silva (2009) investigate the contribution of simple value-added (VA)[25] and prior achievements, KS1, into house prices only for primary schools in years 2003-2006[26]. They use boundary discontinuity regressions and their results show that a one-standard deviation change in either value added or prior achievement raises prices by around 3-4 %. However, the results from this study show that the effect of value added (measured by the CVA indicator which is different from the Simple Value-Added as explained earlier) is also positive and significant only in primary education. In secondary education the effect is negative and very significant. Furthermore, when final results (named "score component") are included in the regression the effect of CVA net of score (or the non-score component in our analysis) becomes negative and significant in secondary schools. This is similar to the results reported by Haurin and Brasington (2006); but the same does not hold true for the case of primary education where the non-score component appears to be insignificant.

This chapter of the thesis, focuses on how CVA affects house prices in the catchment area of primary and secondary schools in England. The emphasis in the analysis 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, between different levels of education and between neighbourhoods and spatially larger entities, like LAs.

The empirical analysis, based on data drawn from three independent UK data sources, showed that a linear hedonic regression is sufficient to capture the relationship between log house prices and the (normalised) measures of the school quality indicators under consideration. Furthermore, it shows that CVA has a significant effect on house prices, but positive in the case of primary and negative effect in the case of secondary education. This contradictory result is explained by the fact that the score component of CVA is positive and strongly significant for both primary and secondary schools, while the non-score component is insignificant for primary but very significant and negative for secondary schools. Furthermore, when LA-specific effects were ignored the non-score component of

---

[25] Look Appendix A2 for the definition of simple value-added.

[26] They used value-added scores as an output measure of school quality and Key Stage-1 scores as an input measure of school quality (not publicly available from the Department for Children, Schools and Families).

CVA assumes a more 'positive' role, becoming positive (from insignificant) for primary and insignificant (from negative) for the secondary schools.

Our findings are in line with results already published in the literature, in the sense that they confirm the strong positive effect of final score measures on house prices. Furthermore, they highlight the complexities of using broad measures of school quality, such as the CVA, in hedonic price regressions. In particular, we have found that the non-score component of CVA plays no role at primary level of education. This is probably a reflection of the limited variation of prior achievement (the most important item in the non-score part of CVA) between pupils at this education level, so that CVA may not contain much more information than the final score itself. In contrast, the non-score part of CVA plays a significant role in the analysis of secondary schools data, apparently, because in this case pupils can vary substantially in their prior achievement. Therefore, reaching a given final score starting from a high initial level subtracts from the school's image of quality and, thereby, the prices of houses in its catchment area. 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 was explained by this component; whereas for secondary schools, the corresponding figure was only 18%.

In addition to differences in the effect of CVA on house prices between low and high education levels, this effect can also vary with the level of spatial aggregation at which empirical investigation is pursued. Our finding that CVA has a positive (or not negative) effect on house prices between rather than within LAs can be interpreted as highlighting the 'public good' nature of this school quality indicator. In other words, 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) are diffused over a greater geographical area than that of the catchment area of a particular school. Thus, policies aimed at raising CVA are more of a task that can be undertaken at the LA (rather than the individual school) level and their success is visible to the household at that level. Once households are within a particular LA, however, the non-score components of CVA become either irrelevant (primary schools); or they can even have a negative effect on house prices (secondary schools).

In terms of policy implications, the analysis in this study suggested that the recently adopted practice of using CVA as a measure of school quality in England can encourage

LAs to pay more attention to raising the non-score quality characteristics of their schools. This policy appears to conform to household preferences, expressed by their willingness to pay more (not less) for houses in the catchment area of primary (secondary) schools in LAs with higher CVA average.

Appendix A

A1. The National Curriculum and Key Stage Tests

In 1988 the National Curriculum was introduced in England, setting out the subjects and programmes of study which maintained schools are obligated to cover from ages 5-16.

The National Curriculum and Key Stage tests are maintained by the QCA (Qualifications and Curriculum Authority)[27]. Independent schools (private schools) do not need to operate the National Curriculum or the Key Stage tests[28]. All stages from the age of 3 to 16 are presented below.

Foundation stage (age 3-5)

This covers children in nurseries and the reception year at primary school. The Foundation Stage Profile was introduced into schools and settings in 2002/3, with 13 summary scales covering the six learning areas (personal, social and emotional development; communication, language and literacy; mathematical development; knowledge and understanding of the world; physical development; creative development). Attainment on these scales is assessed by the teachers for each child receiving government-funded education by the end of the pupil's time in the foundation stage.

Key Stage 1 (age 5-7)

This covers Year 1 and Year 2 in primary schools, with pupils assessed at the end of Year 2 when most are 7 years old. The national curriculum specifies learning across a range of subjects such as history, art and information technology, but the three core subjects are English, mathematics and science. Pupils take tests in reading, writing and mathematics, but since 2005 these tests have only been used to inform overall teacher assessments

Key Stage 2 (age 7-11)

Twice the length of Key Stage 1, this takes pupils from Year 3 to Year 6, up to the age of 11 which is usually seen as the end of 'primary education': the following year most pupils

[27] For more information see http://www.qca.org.uk/index.html.

[28] Private Schools educate around 6-7% of pupils in England as a whole.

in maintained schools move to secondary schools. At the end of the four years pupils are assessed by teachers and take tests in English, Mathematics and Science.

Key Stage 3 (age 11-14)

This covers the first three years of secondary schooling (Year 7 to Year 9). Again there is a national curriculum across a range of subjects, with teacher assessment and tests in English, Mathematics and Science.

Key Stage 4 (age 14-16)

This covers the final period (Year 10 and 11) of compulsory schooling during which pupils are working towards a range of academic and vocational qualifications, partly assessed via coursework. Most of the assessment is at the end of Year 11. The qualifications are set by various independent awarding bodies. The main qualification is the GCSE – there are currently over forty academic subjects on offer and eight vocational subjects. However there are also a very wide range of other qualifications which can be taken by this age group. The number of subjects taken by pupils will vary.

A1.1 Mainstream state schools

The main four types of mainstream state schools are all funded by local authorities. They all follow the National Curriculum and are regularly inspected by the Office for Standards in Education (Ofsted)[29].

Community schools (CY): Community schools are run by local authorities, who employ the staff, own the land and buildings, and are responsible for admitting pupils. Community schools make strong links with their local community, offering their facilities and providing services such as childcare and adult learning classes.

Foundation schools (FD)*:* Foundation schools are managed by a governing body which employs the staff and sets the admission criteria. The land and buildings are usually owned by either the governing body or a charitable foundation. By managing themselves, foundation schools believe they can provide the best education for their pupils.

Voluntary-aided schools (VA)**:** Voluntary-aided schools are mainly funded, but not owned, by their local authority. A governing body employs the staff and sets the admission criteria. The school's buildings and land are normally owned by a charitable foundation,

[29] www.ofsted.gov.uk

often a religious organisation, and the governing body contributes to building and maintenance costs.

***Voluntary-controlled schools (VC)*:** Voluntary-controlled schools are run by the local authority, which employs the school's staff and runs the admission procedure. The school's land and buildings are normally owned by a charity, often a religious organisation, which appoints some of the members of the governing body.

A2. Value Added (VA) and Contextual Value Added (CVA) Scores in UK

Value added (VA)[30] is a measure of the progress pupils make between different stages of education. The value added 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. As educational economists claimed, CVA[31] gives a much fairer statistical measure of the effectiveness of a school and provides a solid basis for comparisons. However, it is important to remember that CVA is a measure of progress over a period of time from a given starting point and not a measure of absolute attainment. In order to calculate this measure a more complex model is needed but the principle remains the same as for the value added median line approach (VA).

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)

[30] Simple Value-Added was first published for the primary schools in 2003.

[31] CVA models have been produced separately for mainstream and special schools. This study focuses on CVA scores for mainstream schools.

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 each year group in England.[32] 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. Details on these variables are described just below.

- **Gender:** It allows for the different rates of progress made by boys and girls by adjusting predictions for females.
- **Special Educational Needs (SEN)**: Special Educational Needs' (SEN) covers a wide range of needs that are often interrelated as well as specific needs that usually relate to particular types of impairment. Children with SEN will have needs and requirements which may fall into at least one of four areas: i) communication and interaction; ii) cognition and learning; iii) behaviour, emotional and social development and iv) sensory and / or physical needs. Funding for SEN can differ in each Authority depending on the funding formula agreed with their schools.
- **Ethnicity:** PLASC collects data for 18 main ethnic groups, with a 19th code available for 'unclassified' since provision of this data is voluntary. Even with 18 ethnic groups, the codes obviously have to cover pupils with very different characteristics. For example, the Black African category covers pupils who may or may not speak English, or who may come from recent or long established immigrant communities. The source of the data usually comes from the school, parents or the pupil themselves.

[32] 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.

- **Language:** Adjustment for the effect of pupils whose first language is other than English (it does not distinguish different languages) other than English. The size of this adjustment depends on the pupil's prior attainment. This is because the effect of this factor tends to taper, with the greatest effect for pupils starting below expected levels and lesser effects for pupils already working at higher levels. As with ethnicity, there is some small changes year-on-year in this variable, even though it ought to remain constant.

- **Eligible for Free Schools Meals:** Children whose parents receive the social welfare benefit Income Support, and some related benefits, are entitled to claim free school meals (FSM). The size of this adjustment depends on the pupil's ethnic group. This is because the data demonstrates that the size of the FSM effect varies between ethnic groups. To become entitled to FSM the parents have to indicate a wish for their child to have a school meal and give a proof of benefit receipt. FSM is the only direct measure on PLASC that relates to the pupil's family income. It is a useful variable, but cannot be considered an accurate proxy for social class or income more generally because it is a simple binary flag[33].

- **Date of entry/Mobility:** PLASC records the date of entry into the school for each pupil. This information can be used to obtain measures of pupil mobility: the current method is to flag pupils who joined at non-standard times - any month other than July, August or September. It is also possible to treat pupils differently depending on how recently they joined the school.

- **Age:** Look at a pupil's age within year based on their date of birth.

- **In Care:** This variable counts pupils who are in the care of their Local Authority. These children may be living with foster parents or prospective adopters, placed in

[33] Even as a simple proxy measure for 'deprivation' FSM has disadvantages. The fact that parents have to register an interest in their pupils' having school meals may discourage some from applying. Those not applying but eligible may be an unrepresentative group, e.g. choosing not to register for cultural reasons or on the basis of dietary preferences that may correlate with attitudes to education. The benefit rules may also exclude some families who could be considered deprived, e.g. some low paid workers. Clearly a pupil's entitlement to a free meal may change when family circumstances change and this may not always get recorded, although it is possible to trace movements in and out of FSM status on PLASC. More generally, the propensity of parents to register as entitled to FSM may vary in response to the quality of meals, which may vary from school to school or between Local Authorities

children's homes or some other form of residential care, or placed at home with their parents.

- **IDACI (Income Deprivation Affecting Children Index)**: This index is provided by the Department for Communities and Local Government and it measures the proportion of children under the age of 16 in an area living in low income households. It is supplementary index to the indices of multiple deprivations[34].

A2.2 CVA Indices

Having taken into account the range of factors affecting performance, the Department base each pupil's CVA score on a comparison between their actual key stages (i.e KS-2) performance and the key stages (i.e KS-2) performance predicted for each pupil by median line approach as it explained above. Thus, an average of all pupils' CVA scores is calculated for a school. This number is presented as a number based around 100[35] and around 1000[36] to indicate the value the school has added on average for its pupils. The below table show the KS1-2 and KS2-4 CVA score indicators for primary schools and secondary schools respectively,[37] as also the percentiles which give the distribution of CVA scores and show where schools are placed nationally compared to other schools, based on the CVA measure.

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%

[34] Multiple deprivation includes income deprivation, health deprivation, crime deprivation, environment deprivation, employment deprivation and housing deprivation

[35] Primary Schools CVA Indicator

[36] Secondary Schools CVA Indicator

[37] KS2-3 CVA indicator is also available for secondary schools (see performance tables in the Department of Education, Schools and Families of UK)

Specifically, the KS1-2 CVA measure is shown as a score based around 100. For KS1-2 CVA, a measure of 101 means that, on average, the school's KS2 cohort has achieved one term's more progress than the national average. A score of 99 means that the school's pupils made a term's less progress. In addition, the KS2-4 CVA is shown as a score based around 1000. Scores above 1000 represent schools where students on average made more progress than similar students nationally, while scores below 1000 represent schools where students made less progress.

A3. Data Construction

Our data come from three major independent sources. For the house prices and characteristics we use the electronic site "Up my Street"[38]. Our primary and secondary school performance data come from the primary and secondary school performance tables, available from the Department for Children, Schools and Families[39] and the deprivations indices and other neighbourhood characteristics come from the Office of National Statistics[40].

The data collection process from the three different sources described above was as follows: The England is divided in nine regions. These nine regions are made up by one hundred fifty local authorities. Fifty local authorities were chosen from these using the following process: from each region, one third of the total number of the local authorities of the region was chosen. For example, North East region comprises of twelve local authorities. From these twelve, four were chosen on the condition that half of them had a grade mean larger than the England grade average, and the other half a lower grade mean on both primary and secondary schools[41] From each of the above local authorities, six schools were randomly selected. Three of them had a higher grade mean than the local authority grade average and three a lower grade mean. This process of collection was done

[38]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)

[39]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)

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

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

separately for primary and secondary schools[42]. Apart from grades and other characteristics, we also had information on the exact school unit code of a school.

Based on this school unit code 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 were also able to acquire information on the selling price of houses, house characteristics, distance from school and their exact unit codes. At this point it is important to mention that the distance of the house which is located from the local primary or secondary school was maximum 1 mile in order to ensure that houses were within a specific school catchment area. Thus, sometimes we were able to locate less than six houses into a specific school code due to the criterion of maximum 1 mile. In both schooling level datasets we end up with an average distance around 0.20 miles[43].

The site of Neighbourhood Statistics provides detailed statistics within specific geographic areas including deprivations indices of income, crime, environment, housing barriers, health, and employment and some other neighbourhood characteristics like population density. In order to capture neighbourhood characteristics, we used the "lower layer super output area[44]" 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.

[42] In UK there are more than four times as many primary schools than secondary. Specifically in 2006 there were 17,500 primary schools and 3,400 secondary schools. However, the focus of this study was to conduct a separate analysis for the two educational levels and have a representative sample of house prices and corresponding local state schools from the whole of England. Hence, the final sample collected is around 1400 house price observations for primary education and around 1200 for secondary education; and include 300 schools for each educational level.

[43] In the England, postcodes contain up to seven alphanumeric characters. The first two alphanumeric characters define the postcode area or the broadest postal zone (region). Examples are EC, E representing the region London, HP representing the region South East, and BD representing the region Yorkshire. Within postcodes, the next level down is the postcode districts. This is defined by a single or two-digit number following the postcode area. Examples are EC2A (a single letter further of the single or two-digit number subdivides some postcode districts in central London) E12, and BD10. Below this, we have the postcode sectors which they defined by a single number following the postcode district. Examples are EC2A 3, E12 4. Finally, we have the specific unit code of the property. This is defined by two alphanumeric characters following the postcode sectors.

[44] 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, lower layer area has about 600 households.

- 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[45].

A4. Tables

Table A4.1: Coefficients of Instruments in first stage of IV equation

	Primary Schools	Secondary Schools
Ln House Price	Model IV (2SLS)	Model IV (2SLS)
Community School[1]	-0.329***	-0.562***
	(0.060)	(0.049)
Pupils min age: 3 years old	-0.530***	-
	(0.056)	-
Pupils max age: 16 years old	-	-0.259***
	-	(0.058)
School gender: Girls[2]	-	0.583***
	-	(0.083)
School gender: Boys	-	0.013
		(0.082)
F-test of Instruments	68.59	76.56
	0.000	0.000

Notes: [1] Around 82% (61%) of the samples are community schools for primary (secondary) education;[2] 10% of the sample are schools that have only girls, 85% are mixed schools and only 5% are schools with only boys.

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

Table A4.2: Descriptive Statistics-Primary Schools

Variable Name	Mean	Std.Dev.	Min	Max
<u>House Characteristics</u>				
Bedrooms	2.838	1.021	1	7
Bathrooms	1.253	0.535	0	6
Toilet	0.361	0.540	0	5
Totalrooms	7.632	2.597	3	20
Garage	0.381	0.551	0	3
Detached	0.220	0.415	0	1
Semidetached	0.253	0.435	0	1
Flat	0.188	0.391	0	1
Terrace	0.134	0.340	0	1
Midterrace	0.100	0.300	0	1
Endterrace	0.048	0.213	0	1
Maisonette	0.017	0.131	0	1
Other	0.040	0.197	0	1
Distance	0.194	0.155	0	0.99
<u>Neighborhood Characteristics</u>				
Income_deprivation	0.167	0.121	0.010	0.630
Employment_deprivation	0.102	0.060	0.010	0.420
Health_deprivation	0.028	0.869	-2.650	2.800
Housing_deprivation	20.325	10.491	0.810	55.120
Crime_ deprivation	0.165	0.866	-2.230	9.180
Environment_deprivation	25.003	17.712	0.610	83.110
Density (persons per hectare)	49.506	34.646	0.420	242.810
Non-domestic buildings (square metres)	19.923	32.510	0.000	495.890
<u>Regions</u>				
East Midlands	0.045	0.208	0	1
East England	0.071	0.257	0	1
London	0.256	0.437	0	1
North East	0.066	0.248	0	1
North West	0.151	0.358	0	1
South East	0.118	0.323	0	1
South West	0.114	0.318	0	1
West Midlands	0.083	0.276	0	1
Yorkshire	0.095	0.294	0	1
<u>Local Authorities</u>				
Barnet	0.025	0.155	0	1
Bath	0.020	0.141	0	1
Bexley	0.025	0.155	0	1
Birmingham	0.022	0.146	0	1
Blackpool	0.021	0.143	0	1
Bradford	0.025	0.155	0	1
Brent	0.021	0.143	0	1
Buckinghamshire	0.023	0.150	0	1
Camden	0.020	0.141	0	1
Cheshire	0.020	0.141	0	1
Bristol	0.014	0.116	0	1
York	0.014	0.116	0	1
Coventry	0.024	0.153	0	1
Darlington	0.009	0.093	0	1
Derby	0.020	0.141	0	1
Enfield	0.021	0.143	0	1
Essex	0.026	0.159	0	1
Gloucestershire	0.020	0.141	0	1
Greenwich	0.023	0.150	0	1
Hampshire	0.011	0.104	0	1
Havering	0.024	0.153	0	1
Islington	0.023	0.150	0	1
Kirklees	0.023	0.150	0	1
Lambeth	0.017	0.131	0	1
Lancashire	0.021	0.143	0	1
Leeds	0.012	0.110	0	1
Manchester	0.022	0.146	0	1

Newcastle	0.013	0.113	0	1
Newham	0.018	0.133	0	1
Norfolk	0.022	0.146	0	1
Northtyneshire	0.025	0.155	0	1
Northyorkshire	0.022	0.146	0	1
Northamptoshire	0.019	0.138	0	1
Reading	0.020	0.141	0	1
Richmond	0.022	0.146	0	1
Rutland	0.006	0.076	0	1
Salford	0.019	0.136	0	1
Solihull	0.018	0.133	0	1
Somerset	0.020	0.141	0	1
Southampton	0.020	0.141	0	1
Southend	0.023	0.150	0	1
Staffordshire	0.019	0.138	0	1
Stockport	0.019	0.136	0	1
Stockton	0.019	0.138	0	1
Surrey	0.024	0.153	0	1
Sutton	0.018	0.133	0	1
Swindon	0.022	0.146	0	1
Trafford	0.019	0.138	0	1
Wigan	0.010	0.100	0	1
Wiltshire	0.018	0.133	0	1
Wokingham	0.020	0.141	0	1
n=1385				

Table A4.3: Descriptive Statistics-Secondary Education

Variable Name	Mean	Std.Dev.	Min	Max
House Characteristics				
Bedrooms	2.892	1.021	1	8
Bathrooms	1.255	0.514	1	4
Toilet	0.411	0.548	0	3
Totalrooms	7.652	2.582	3	18
Garage	0.496	0.624	0	3
Detached	0.253	0.435	0	1
Semidetached	0.318	0.466	0	1
Flat	0.179	0.383	0	1
Terrace	0.071	0.257	0	1
Midterrace	0.068	0.252	0	1
Endterrace	0.052	0.222	0	1
Maisonette	0.023	0.150	0	1
Other	0.035	0.183	0	1
Distance	0.213	0.151	0	1
Neighborhood Characteristics				
Income_deprivation	0.188	0.944	0.01	19
Employment_deprivation	0.093	0.053	0.02	0.37
Health_deprivation	-0.093	0.824	-2.52	2.53
Housing_deprivation	21.709	10.943	0.83	63.44
Crime_ deprivation	0.131	0.775	-2.22	2.51
Environment_deprivation	21.028	14.954	1.22	75.13
Density (persons per hectare)	39.150	30.985	0.1	244.76
Non-domestic buildings (square metres)	22.933	38.193	0.03	438
Regions				
East Midlands	0.046	0.210	0	1
East England	0.055	0.229	0	1
London	0.251	0.434	0	1
North East	0.073	0.260	0	1
North West	0.154	0.361	0	1
South East	0.120	0.325	0	1
South West	0.122	0.327	0	1
West Midlands	0.071	0.257	0	1
Yorkshire	0.108	0.311	0	1
Local Authorities				
Barnet	0.019	0.137	0	1
Bath	0.026	0.158	0	1
Bexley	0.018	0.134	0	1
Birmingham	0.007	0.086	0	1
Blackpool	0.029	0.168	0	1
Bradford	0.025	0.156	0	1
Brent	0.017	0.131	0	1
Buckinghamshire	0.025	0.156	0	1
Camden	0.021	0.142	0	1
Cheshire	0.017	0.131	0	1
Bristol	0.017	0.131	0	1
York	0.023	0.150	0	1
Coventry	0.022	0.148	0	1
Darlington	0.006	0.076	0	1
Derby	0.023	0.150	0	1
Enfield	0.025	0.156	0	1
Essex	0.012	0.111	0	1

Gloucestershire	0.021	0.142	0	1
Greenwich	0.023	0.150	0	1
Hampshire	0.022	0.145	0	1
Havering	0.029	0.168	0	1
Islington	0.021	0.142	0	1
Kirklees	0.022	0.145	0	1
Lambeth	0.020	0.140	0	1
Lancashire	0.019	0.137	0	1
Leeds	0.024	0.153	0	1
Manchester	0.013	0.114	0	1
Newcastle	0.020	0.140	0	1
Newham	0.023	0.150	0	1
Norfolk	0.020	0.140	0	1
Northtyneshire	0.023	0.150	0	1
Northyorkshire	0.015	0.121	0	1
Northamptoshire	0.018	0.134	0	1
Reading	0.012	0.107	0	1
Richmond	0.012	0.111	0	1
Rutland	0.005	0.070	0	1
Salford	0.011	0.103	0	1
Solihull	0.017	0.131	0	1
Somerset	0.017	0.128	0	1
Southampton	0.024	0.153	0	1
Southend	0.023	0.150	0	1
Staffordshire	0.024	0.153	0	1
Stockport	0.018	0.134	0	1
Stockton	0.024	0.153	0	1
Surrey	0.015	0.121	0	1
Sutton	0.022	0.148	0	1
Swindon	0.022	0.148	0	1
Trafford	0.020	0.140	0	1
Wigan	0.026	0.161	0	1
Wiltshire	0.019	0.137	0	1
Wokingham	0.023	0.150	0	1

n=1209

Table A4.4: Estimated parameters for all variables in the hedonic analysis-Part A

	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%.

Table A4.5: Estimated parameters for all variables in the hedonic analysis-Part B

	Primary Schools		Secondary Schools	
Ln House Prices	Model III	Model IV	Model III	Model IV
CVA	0.022**		-0.027***	
	(0.009)		(0.009)	
Score		0.035***		0.039***
		(0.008)		(0.010)
Non-Score		-0.001		-0.043***
		(0.011)		(0.010)
<u>House Characteristics</u>				
Bedrooms	0.095***	0.094***	0.097***	0.099***
	(0.015)	(0.015)	(0.017)	(0.017)
Bathrooms	0.094***	0.092***	0.035	0.035
	(0.021)	(0.021)	(0.025)	(0.024)
Toilet	-0.012	-0.015	-0.021	-0.023
	(0.017)	(0.017)	(0.021)	(0.021)
Totalrooms	0.056***	0.057***	0.080***	0.078***
	(0.008)	(0.008)	(0.009)	(0.009)
Garage	0.038***	0.039***	0.061***	0.059***
	(0.014)	(0.014)	(0.015)	(0.015)
Semidetached	-0.097***	-0.094***	-0.138***	-0.133***
	(0.021)	(0.020)	(0.021)	(0.021)
Flat	-0.273***	-0.269***	-0.153***	-0.163***
	(0.031)	(0.031)	(0.037)	(0.037)
Terrace	-0.168***	-0.169***	-0.187***	-0.184***
	(0.027)	(0.026)	(0.034)	(0.033)
Midterrace	-0.217***	-0.216***	-0.264***	-0.266***
	(0.028)	(0.028)	(0.031)	(0.031)
Endterrace	-0.205***	-0.204***	-0.211***	-0.213***
	(0.031)	(0.031)	(0.031)	(0.031)
Maisonette	-0.263***	-0.271***	-0.285***	-0.293***
	(0.043)	(0.044)	(0.049)	(0.049)
Other	-0.078*	-0.081**	-0.006	-0.010
	(0.040)	(0.040)	(0.057)	(0.056)
<u>Neighborhood Characteristics</u>				
Income_deprivation	-0.047**	-0.040**	-0.008*	-0.010**
	(0.019)	(0.019)	(0.004)	(0.004)
Housing_deprivation	0.007	0.009	-0.030**	-0.026*
	(0.012)	(0.012)	(0.014)	(0.014)
Crime_ deprivation	-0.007	-0.003	-0.042***	-0.041***
	(0.009)	(0.009)	(0.011)	(0.011)
Environment_deprivation	-0.005	-0.004	0.048***	0.045***
	(0.013)	(0.013)	(0.014)	(0.014)
Health_deprivation	-0.044*	-0.034	-0.100***	-0.082***
	(0.023)	(0.024)	(0.024)	(0.023)
Employment_deprivation	-0.037*	-0.041*	-0.040**	-0.037**
	(0.021)	(0.021)	(0.018)	(0.017)
Density (persons per hectare)	-0.034***	-0.036***	-0.039***	-0.034***
	(0.010)	(0.010)	(0.011)	(0.011)
Non-domestic buildings (square metres)	-0.002	-0.002	0.002	0.006
	(0.009)	(0.009)	(0.007)	(0.007)
<u>Regions</u>				
East Midlands	-0.902***	-0.911***	-0.590***	-0.602***
	(0.057)	(0.056)	(0.120)	(0.116)
East England	-0.490***	-0.479***	-0.480***	-0.482***
	(0.059)	(0.059)	(0.126)	(0.123)

North East	-0.831***	-0.824***	-0.817***	-0.820***
	(0.060)	(0.061)	(0.087)	(0.083)
North West	-0.645***	-0.649***	-0.559***	-0.588***
	(0.069)	(0.069)	(0.083)	(0.077)
South East	-0.494***	-0.487***	-0.361***	-0.370***
	(0.078)	(0.078)	(0.094)	(0.090)
South West	-0.503***	-0.497***	-0.475***	-0.464***
	(0.092)	(0.092)	(0.079)	(0.074)
West Midlands	-0.649***	-0.657***	-0.473***	-0.506***
	(0.053)	(0.053)	(0.081)	(0.083)
Yorkshire	-0.603***	-0.605***	-0.730***	-0.744***
	(0.077)	(0.077)	(0.070)	(0.068)
Local Authorities				
Bath	-0.074	-0.075	-0.167*	-0.189**
	(0.093)	(0.092)	(0.085)	(0.083)
Bexley	-0.335***	-0.336***	-0.416***	-0.426***
	(0.058)	(0.058)	(0.071)	(0.067)
Blackpool	-0.328***	-0.349***	-0.190***	-0.160***
	(0.053)	(0.052)	(0.066)	(0.061)
Bradford	-0.156**	-0.165**	-0.066	-0.058
	(0.068)	(0.068)	(0.055)	(0.052)
Brent	0.013	0.025	0.044	0.018
	(0.068)	(0.068)	(0.095)	(0.096)
Buckinghamshire	-0.421***	-0.423***	-0.096	-0.079
	(0.051)	(0.052)	(0.077)	(0.075)
Camden	0.686***	0.703***	0.600***	0.566***
	(0.068)	(0.068)	(0.072)	(0.071)
Cheshire	-0.183**	-0.199**	-0.250***	-0.234***
	(0.081)	(0.080)	(0.078)	(0.072)
Coventry	-0.176***	-0.151***	-0.300***	-0.281***
	(0.043)	(0.043)	(0.066)	(0.070)
Darlington	0.203**	0.182*	0.185	0.174
	(0.102)	(0.099)	(0.168)	(0.152)
Derby	-0.445***	-0.472***	-0.347***	-0.345***
	(0.085)	(0.083)	(0.116)	(0.113)
Enfield	-0.114*	-0.093	-0.101	-0.132**
	(0.069)	(0.070)	(0.065)	(0.062)
Gloucestershire	-0.168*	-0.166*	-0.084	-0.103
	(0.097)	(0.097)	(0.092)	(0.091)
Greenwich	-0.108	-0.089	0.101	0.096
	(0.067)	(0.068)	(0.101)	(0.101)
Hampshire	-0.387***	-0.419***	-0.259***	-0.266***
	(0.057)	(0.061)	(0.083)	(0.079)
Havering	-0.383***	-0.392***	-0.351***	-0.392***
	(0.057)	(0.057)	(0.077)	(0.074)
Islington	0.733***	0.741***	0.722***	0.713***
	(0.066)	(0.068)	(0.082)	(0.080)
Kirklees	-0.262***	-0.267***	-0.144*	-0.152*
	(0.071)	(0.071)	(0.079)	(0.079)
Lambeth	0.458***	0.453***	0.424***	0.402***
	(0.070)	(0.070)	(0.081)	(0.078)
Lancashire	-0.297***	-0.330***	-0.367***	-0.357***
	(0.071)	(0.071)	(0.070)	(0.066)
Leeds	-0.259***	-0.274***	-0.158***	-0.167***
	(0.082)	(0.083)	(0.049)	(0.048)
Newham	-0.029	-0.024	0.075	0.014
	(0.057)	(0.057)	(0.072)	(0.070)
Norfolk	-0.298***	-0.296***	-0.397***	-0.409***
	(0.055)	(0.055)	(0.111)	(0.109)

Northtyneside	0.033	0.014	0.057	0.028
	(0.052)	(0.052)	(0.054)	(0.052)
Northyorkshire	-0.158**	-0.158**	-0.072	-0.056
	(0.066)	(0.066)	(0.063)	(0.063)
Northamptonshire	-0.424***	-0.437***	-0.322***	-0.345***
	(0.088)	(0.084)	(0.107)	(0.105)
Reading	-0.217***	-0.227***	-0.034	-0.021
	(0.056)	(0.055)	(0.092)	(0.087)
Richmond	0.222***	0.241***	-0.123	-0.135
	(0.075)	(0.076)	(0.085)	(0.082)
Salford	-0.109	-0.130*	0.065	0.064
	(0.072)	(0.069)	(0.090)	(0.089)
Solihull	-0.020	-0.016	-0.164**	-0.178**
	(0.071)	(0.071)	(0.074)	(0.076)
Somerset	-0.195**	-0.195**	-0.218***	-0.227***
	(0.085)	(0.085)	(0.068)	(0.065)
Southampton	-0.431***	-0.421***	-0.234***	-0.231***
	(0.045)	(0.045)	(0.077)	(0.074)
Southend	0.053	0.059	-0.064	-0.091
	(0.051)	(0.050)	(0.104)	(0.103)
Staffordshire	-0.201***	-0.197***	-0.414***	-0.409***
	(0.055)	(0.055)	(0.072)	(0.074)
Stockport	-0.154***	-0.166***	-0.056	-0.071
	(0.058)	(0.057)	(0.072)	(0.065)
Stockton	-0.035	-0.039	-0.098	-0.097
	(0.058)	(0.059)	(0.060)	(0.060)
Sutton	-0.262***	-0.233***	-0.397***	-0.430***
	(0.058)	(0.059)	(0.071)	(0.069)
Swindon	-0.319***	-0.309***	-0.331***	-0.352***
	(0.085)	(0.085)	(0.076)	(0.072)
Trafford	0.080	0.068	0.033	0.057
	(0.064)	(0.063)	(0.077)	(0.073)
Wigan	-0.316***	-0.338***	-0.262***	-0.270***
	(0.073)	(0.071)	(0.065)	(0.060)
Wiltshire	-0.102	-0.092	-0.274***	-0.307***
	(0.093)	(0.093)	(0.072)	(0.070)
Wokingham	-0.272***	-0.264***	-0.252***	-0.236***
	(0.053)	(0.052)	(0.079)	(0.076)
Constant	12.068***	12.060***	12.012***	12.039***
	(0.064)	(0.064)	(0.082)	(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%.

Chapter 3

A Structural Demand Analysis Approach to Estimating the Capitalization of School Quality to House Prices

3.1 Introduction

Applying hedonic analysis to publicly accessible data to estimate the capitalisation of local state school quality to house prices is not possible in countries like the UK, where family expenditure surveys do not contain information about the location of households; thus, it is not possible to associate the price of their house with indicators reflecting the quality of the local state school. To circumvent this problem this chapter proposes an alternative (to hedonic) approach based on the theory of consumer demand and uses the resulting model to investigate the capitalization of local state school quality to house prices using publicly available UK Family Expenditure Surveys (FES). The parameter estimate obtained from this approach is the shadow price of education relative to housing, and largely resembles the corresponding parameter obtained from standard hedonic analysis.

The rationale behind the proposed model is the fact that the increased price observed for houses located in the catchment area of high quality state schools expresses the household's willingness to pay for an education level above the «minimum» provided free by the state. Thus families may locate in areas where local authorities spending on education are high enough to match their educational preferences; or select a location outside these areas and pay out-of-pocket for the education of their children by enrolling them to private schools.

The capitalisation of local state school quality to house prices has been an object of a large body of literature where hedonic analysis (Rosen, 1974) is used to reach an empirical valuation of the quality of local state schools, especially in the US. The measures used to capture school quality vary from expenditure per pupil (Downes and Zabel, 2002), pupil/teacher ratio (Brasington, 1999), reading scores, value added measures and proficiency test scores (Hayes and Taylor, 1996; Downes and Zabel; 2002, Haurin and

Brasington, 2006). These measures are found to be consistently capitalised into housing prices. Studies in the US reporting a positive effect of school quality on houses prices include Black (1999), Bogart and Cromwell (1997, 2000), Brasington (2000, 2002), Barrow (2002), Barrow and Rouse (2004), Downes and Zabel (2002), Brasington and Haurin (2006, 2009), Kane et. all (2006).

In the UK, the issue has received less attention, with only a small number of studies available to date, probably because data about the location of individual households are not available at a sufficiently disaggregate level due to confidentiality. Nevertheless, it is well known that parents in the UK seek to locate themselves in the catchment area of state schools providing high quality education. Gibbons and Machin (2003), the first authors attempting to estimate the capitalisation of local state school quality to house prices in the UK, find that a rise of primary school standards has an effect equivalent to between 0.5% to 0.8% on the local mean property prices. A positive effect of school quality on house prices is also reported for primary and secondary schools in Reading by Cheshire and Sheppard (2004) where - as in Black (1999) - the authors emphasise the importance of including neighbourhood characteristics in the specification to avoid estimation bias. Rosenthal (2003) evaluate the effect of school quality of English secondary schools on house prices applying instrumental variable techniques to data drawn from various sources to account for potential endogeneity bias arising from the fact that schools can achieve better educational outcomes from being located in areas where house prices are high. They found the house price elasticity of school quality to be around 0.05. Leech and Campos (2003) found house prices in the catchment area of two popular comprehensive schools in a local authority at the Coventry area to be between 16% and 20% higher than those in the catchment area of other schools.

As hinted earlier, applying hedonic analysis to estimate the capitalisation of local state school quality to house prices is not possible in countries (like the UK) where family expenditure surveys do not disclose information about the location of households enabling one to associate the price of their house with indicators reflecting the quality of the local state school. This study proposes a method circumventing this problem based on consumer demand analysis. This approach is motivated by the argument made in the context of the widely researched issue of state vs private school selection that communities with a high proportion of students enrolled in private schools can be less willing to pay taxes to finance state schools (Goldhaber, 1999). Therefore, families may locate in areas where local authority spending on education and property tax rates (hence, housing costs) are high enough to match their educational desires; or pay out-of-pocket for the education of their children by enrolling them to private schools (Brasington, 2006; Lankford, Lee and Wyckoff, 1995; Lonh and Toma, 1988). This means that paying more for a house in the catchment area of a high quality state school and purchasing education from the private sector can be considered as alternative ways to affect consumer demand for high quality education. Very recently Fack and Grenet (2010) investigated the effect of private schools on the housing price premium attach to state school performance using comprehensive data on middle schools and housing sales in Paris over the period 1997-2004. They found evidence that good private schools tend to attenuate the capitalization of state school performance in housing prices by providing a valuable outside option to parents.

While literature on state vs private school selection is mainly concerned with the relationship between private school enrollment rates and the level of public spending on education (per pupil), it highlights the fact that there is a link between the quality of local state school and the willingness of households with children in state schools to pay more for a given housing level than those with children in private schools. This study exploits this link in order to predict the notional education expenditure of households with children in state schools and use this as an indicator of school quality in order to estimate the education component of the jointly education-and-housing expenditure. The proposed method was applied to data drawn from the UK Family Expenditure Surveys (FES) for the years 1994-1997 where information about the purchase value of the property of the participating households is available. The analysis uses households with children in private or state schools: the former report their education and housing expenditures separately; while the latter report only housing expenditures that also embody education costs in the form of higher house price. The FES also includes information on the income and expenditure of households and their members, as well as a large number of demographic and other characteristics, including many housing attributes.

The theoretical model is based on the notion of separability and two-stage budgeting (Deaton and Muellbauer, 1980a). The unobserved education component as well as the jointly composite education-and-housing expenditure is approximated by Heckman-type methods. The parametric empirical analysis based on the propose theoretical model, recasts a new structural model as a household expenditure (cost) equation and estimate the relative (shadow) price of education and housing using both alternatives cost functions, i.e: the

Translog and the Quadratic Logarithmic. In addition, a usual hedonic approach is applied (based on both parametric and semi-parametric method) to the same data using the available information for the purchase value of the property of the participating households. In that case the notional education expenditure approximated by Heckman type methods is used as an indicator of willingness to pay for high school quality.

The structure of this chapter is as follows; the next section presents the theoretical model of the jointly education-and-housing expenditure and proposes a method for the estimation of the education component via both Translog and Quadratic Logarithmic cost functions; section 3 discuss how the extrapolation of the composite education-and-housing expenditure and education component are calculated; section 4 provides the empirical analysis and finally section 5 concludes the chapter.

3.2 Theoretical Model

We assume households consume n commodities from which they derive utility according to the function

$$U = u(q_0, q_1, \dots \dots, q_n) \tag{3.1}$$

where $q_0, q_1, q_2, .., q_n$ are the quantities of the n commodities and $\partial^2 U/q_i > 0$,

$\partial^2 U/q_i^2 \leq 0$, $i = 1, .. n$. While all commodities in the consumption vector q = $(q_0, q_1, q_2, .., q_n)$ can be considered to be composite (here taken to mean that they consist of several items, e.g. food consists of meat, vegetables, milk etc), in the analysis below we consider (without loss of generality) the case of one such commodity, say, q_0. We suppose that this commodity consists of 2 items q_{01}, q_{02}, education and housing respectively. Furthermore, writing

$$U = u[q_0(q_{01}, q_{02}), q_1, q_2, \dots, q_n] \tag{3.2}$$

We assume separability of the items in the sub-function $q_0(q_{01}, q_{02})$, i.e. demand for the items which make up the composite commodity in are not directly affected by changes in the relative prices of $\{q_1, q_2, \dots, q_n\}$; there can be an indirect effect only - through a change in the amount of consumption allocated to q_0, if $q_0(q_{01}, q_{02})$ is non-homothetic. It should be noted here that there are various forms of separability, typically employed in demand analysis to allow for a particular two-stage budgeting scheme: at the first stage the allocation of aggregate expenditure is assumed to take place among commodities defined as broad categories of goods, as in (3.2); at the second stage, the expenditure allocated to each commodity category is allocated to the goods which make up this category[46].

By duality, maximisation of (3.2) subject to the budget constraint $X = r_0 q_0 + \sum_{i=1}^{n} p_i\, q_i$ (where r_0 and p_i are the prices of q_0 and q_i all $i = 1, .. n$, respectively) is equivalent to minimising the cost function

$$C(p, U) = c[c_0(p_{01}, p_{02}, U), p_1, \dots, p_n, U] \tag{3.3}$$

where $c_0(.)$ is a homogeneous and increasing function, and can be thought as the price (index) of items q_{01} and q_{02}, which depends on utility[47]

The Hicksian demand for the j item in q_0 is given by

$$q_{0j} = h_j(p, U) = \left(\frac{\partial c(p, U)}{\partial p_{0j}}\right) = \left(\frac{\partial C}{\partial c_0}\right)\left(\frac{\partial c_0}{\partial p_{0j}}\right) \quad j = 1{,}2 \tag{3.4}$$

where $\partial C / \partial c_0 = q_0$. Furthermore, replacing $\partial c_0 / \partial p_{0j}$ with $(\partial ln\, c_0 / \partial ln\, p_{0j})(c_0 / p_{0j})$ in (3.4) we can see that

$$\frac{\partial lnc_0}{\partial lnp_{0j}} = \omega_{0j} \tag{3.5}$$

where ω_{oj} is the (Hicksian) share of item q_{oj} and is equal to $q_{oj} p_{oj} / q_o\, r_0$. This means that to obtain the components of a composite commodity (as value shares), one has to assume an explicit function for the price index $ln\, c_0(.)$ and apply Shephard's lemma, i.e. calculate the derivatives of this function with respect to log prices. In this study, we assume that $ln\, c_0(.)$ can have both the Translog and the Quadratic Logarithmic form. While the Quadratic Logarithmic is more complicate and has higher order utility terms, it is the most general functional form that allows recovery of the cost .The demands generated by this

46 The different concepts of separability in consumer demand analysis (additive, weak, implicit etc) are discussed in depth in Blackorby et al (1979) and Blackorby and Shorrocks (1996).

47 The dependence of $p_0(.)$ on U implies that consumer demand for q_{0j} , j=1,2, is non-homothetic. Also, the fact that $p_0(.)$ depends on utility defined on the aggregate consumption vector $q_0, q_1, q_2, .., q_n$ (and not on the sub-vector (q_{01}, q_{02}) implies that this function is implicitly (and not weakly) separable.

class are shown to be rank-3 which, as proved by Gorman (1981), is the maximum possible rank for any demand system that is linear in functions of income.

3.2.1 Translog

Here we assume that $ln\ c_0(.)$ has the Translog form

$$\ln c_0(p_0, U) = a_0 + \sum_{j=1}^{2} a_j\ lnp_{0j} + \frac{1}{2}\sum_{j=1}^{2}\sum_{k=1}^{2} \gamma_{j\kappa}\ lnp_{0j}\ lnp_{0k}$$

$$+\sum_{j=1}^{2} \beta_j lnp_{0j}\ U \qquad (3.6)$$

where the parameters a_j, γ_{jk} and β_j for all $j, k = 1,2$ obey the following restrictions: $\sum_{j=1}^{2} a_j = 1$ and $\sum_{j=1}^{2} \gamma_{jk} = \sum_{j=1}^{2} \beta_j = 0$ for adding up, $\sum_{k=1}^{2} \gamma_{jk} = 0$ for homogeneity and $\gamma_{jk} = \gamma_{kj}$ for symmetry. The parameter α_0, the arbitrary point of second order approximation of the Translog function, is the lnc_0 at base prices and subsistence expenditure (zero utility). Then, the (Hicksian) share of item q_{0j} can be written as

$$\omega_{0j} = a_j + \sum_{k=1}^{2} \gamma_{jk} lnp_{0k} + \beta_j U \qquad (3.7)$$

and substituting back in (3.6) we obtain

$$ln\ c_0 = a_0 + \sum_{j=1}^{2} \omega_{0j} lnp_{0j} - \frac{1}{2}\sum_{k=1}^{2}\sum_{j=1}^{2} \gamma_{j\kappa}\ lnp_{0\kappa}\ lnp_{oj} \qquad (3.8)$$

Equation (3.8) provides a theoretical basis for the conduct of hedonic analysis, in the sense that in cross-section data lnp_{0j} is fixed while ω_{0j} (and, consequently $ln\ c_0$) varies across households. Therefore, one can treat lnp_{0j} as parameters and ω_{0j} as the corresponding explanatory variables to estimate the structure equation

$$lnc_0^h = \varepsilon_0 + \sum_{j=1}^{2} \varepsilon_j\ \omega_{0j}^h + e^h \qquad (3.9)$$

where the superscript h denotes the household and e^h is a random error. Regarding the parameters, the intercept $\varepsilon_0 = a_0 - \frac{1}{2}\sum_{k=1}^{2} \gamma_{jk}\ lnp_{0k} lnp_{0j}$ shows the price of the `subsistence' quantity of the composite commodity (adjusted for substitution effects); and ε_j the log shadow price of the j^{th} component of this commodity.

In practice, the cost shares ω_{0j}^h are typically not observed for most composite goods. What is observed instead is whether a particular component is present and if so in what quantity/quality. For example, while some items like the length, bed configuration and meals may be itemised in the cost of a hotel stay, the same is unlikely to be case for most other items like cable tv, gym, sauna, swimming pool and distance from the city centre. Therefore, most often the cost shares ω_{0j}^h are approximated in empirical application by their determinants, usually variables indicating the quantity/quality of component items. In the resulting reduced form equation, however, the parameter estimates have a different interpretation, i.e. they no longer represent the shadow prices of the components of the composite good but the shadow prices of whatever quantity/quality indicators are used as explanatory variables for the composite commodity cost.

As mentioned in the introduction of this chapter, various school performance measures have been routinely used as an indicator of quality by investigators attempting to find how much this indicator, along with other variables (e.g. characteristics of the house and the neighbourhood) affect house prices in the school's catchement area. Our approach emphasizes on remaining closer to the structure model (3.9) defined by the fundamentals of consumer theory described above and also to consider demand for housing and education by households with children attending state schools to be jointly affected through the purchase of a house.

Using the terminology in the analysis above, the quantity demanded of this composite education-and-housing commodity can be written as $q_0(q_{01}, q_{02})$ and, under Translog and Quadratic Logarithmic form preferences just describe below, the relative price (unit value) of the two components at a given point in time can be estimated from household survey data using

$$ln\, c_0^h = \varepsilon^* + \varepsilon_1^* \omega_{0j}^h + e^h \qquad (3.10)$$

where $\varepsilon^* = a_0 - \frac{1}{2}\sum_{k=1}^{2} \gamma_{jk} \ln p_{ok} \ln p_{oj}$ and $\varepsilon_1^* = \ln(p_{01}/p_{02})$. Thus labelling education as the first and housing as the second component of this cost, the explanatory variable ω_{0j}^h in

(3.10) is the share of education in the composite `education-and-housing' expenditure; and the parameter ε_1^* the (log of the) price of education relative to housing.

3.2.2 Quadratic Logarithmic

In the following we assume that the cost function relating to the composite commodity q_0 takes the Quadratic Logarithmic form (Lewbel 1990)

$$lnc_0(p_0, U) = a(p_0) + \frac{b(p_0)U}{1 - \lambda(p_0)U} \quad (3.11)$$

where $a(p_0) = a_0 + \sum_{j=1}^{2} a_j lnp_{0j} + \frac{1}{2}\sum_{k=1}^{2}\sum_{j=1}^{2} \gamma_{jk} lnp_{0k} lnp_{0j}; b(p_o) = \prod_{j=1}^{2} p_{0j}{}^{\beta_j}$ and $\lambda(p_0) = \sum_{j=1}^{2} \lambda_j lnp_{0j}$. Moreover, the parameters $a_j, \gamma_{jk}, \beta_{jk}$ and λ_j for all j, $k = 1,2$ obey the following restrictions: (i) $\sum_{j=1}^{2} a_j = 1$, $\sum_{j=1}^{2} \gamma_{jk} = 0$, and $\sum_{j=1}^{2} \beta_j = \sum_{j=1}^{2} \lambda_j = 0$ for adding up, (ii) $\sum_{k=1}^{2} \gamma_{jk} = 0$ for homogeneity and (iii) $\gamma_{ik} = \gamma_{kj}$ for symmetry. Then, the Hicksian share of item q_{0j} in the cost of the composite commodity q_0 is given by

$$\omega_{0j} = a_j + \sum_{k=1}^{2} \gamma_{jk} lnp_{0k} + \beta_j V + \frac{\lambda_j}{b(p_0)} V^2 \quad (3.12)$$

where $V = \frac{b(p_0)U}{1-\lambda(p_0)U}$. Multiply both sides of (3.12) with $\sum_{j=1}^{2} lnp_{oj}$, and using the specific functional forms of $a(p_0)$, $b(p_o)$ and $\lambda(p_0)$ as defined above, the price index for the composite commodity can be written as

$$\sum_{j=1}^{m} \omega_{0j} lnp_{0j} = \left[.5 \sum_{j=1}^{2} \sum_{k=1}^{2} \gamma_{jk} lnp_{0k} lnp_{0j} - a_0\right] + [a(p_0) + V] \quad (3.13)$$

$$+ [lnb(p_0) - 1]V + \frac{\lambda(p_0)}{b(p_0)} V^2$$

where the RHS is obtained by adding and subtracting the terms $(V + \alpha_0)$. Noting that $lnc_0(p_0, U) = a(p_0) + V$, (3.13) can be solved with respect to the cost function

$$lnc_0(p_0, U) = \varepsilon_0 + \sum_{j=1}^{2} \omega_{0j} lnp_{0j} + B(p_0)V + \Lambda(p_0)\, V^2 \quad (3.14)$$

where $\varepsilon_0 = a_0 - \frac{1}{2}\sum_{j=1}^{2}\sum_{k=1}^{2}\gamma_{jk} lnp_{0k}\, lnp_{0j;}\; B(p_0) = 1 - ln\, b(p_0)\; and\; \Lambda(p_0) = -\lambda(p_0)/b(p_0)$. With cross section data where lnp_{0j} is fixed but ω_{0j} and $ln\, c_0$ vary across households, lnp_{0j} can be treated as parameter and (3.14) can provide the basis for the estimation of the log shadow price of the j^{th} composite commodity. Additionally, with cross section data and fixed prices the monotonic (non-decreasing) transformation of utiliy, V is observationally equivalent to U and can be approximated by (log) total household expenditure, since utility U is generated by the quantities of all goods $q_0, q_1, q_2 \dots \dots .., q_{n,}$determined at a previous budget allocation stage. Then (3.14) is written in the form of an estimable model,

$$lnc_0^h = \varepsilon_0 + \sum_{j=1}^{2} \varepsilon_j \omega_{0j}^h + \delta_1 lnx^h + \delta_2 (lnx^h)^2 + e^h \qquad (3.15)$$

where the superscript h denotes the household , x^h is the total household expenditure and e^h is a random error. The intercept in (3.15) can also vary across households through the dependence of the minimum cost parameter α_0 on household-specific variables. Such variables can include, in general, all factors affecting the minimum utility (cost) determined at an earlier stage, for example the quantities of other commodities, demographic variables, region, area characteristics and variables relating to the presence of durable goods in the household. Thus a more general estimable structural model is given by

$$lnc_0^h = \varepsilon_0^h + \sum_{j=1}^{2} \varepsilon_j \omega_{0j}^h + \delta_1 lnx^h + \delta_2 (lnx^h)^2 + \eta^h \qquad (3.16)$$

where $\varepsilon_0^h = \zeta_0 + \sum_{i=1}^{r} \zeta_i z_i^h$ and η^h is the error term. Equations (3.15) and (3.16) provide unconditional and conditional estimates respectively of the shadow prices of the components of the composite commodity. In the second instance the estimate of ε_j represents the price a household is willing to pay for the j^{th} component of the composite commodity after all other household decisions, apart from the allocation of expenditure between the components of the composite commodity under examination (e.g. housing and education), have been made. In other words (3.16) gives an estimate of the constrained shadow price after all other household commitments have been fulfilled.

Comparing (3.16) with a reduced form hedonic-type model, we observe that the composite commodity price index (cost) in (3.16) can depend not only on good-specific characteristics, which is the case in hedonic price indices, but also on household-specific characteristics and the cost shares of the composite commodity items, ω_{0j}^{h}. The model (3.16) can provide estimates for both income (through the inclusion of ω_{0j}^{h} and substitution (through the inclusion characteristics) effects. In hedonic regressions where ω_{0j}^{h} are absent, any income effects are absorbed by characteristics that determine the quality of the composite commodity components, hence income and substitution effects cannot be distinguished (If ω_{0j}^{h} were not included in the model then the estimated coefficients of the characteristics included represent a mixture of fixed and marginal cost of the composite commodity). Equation (3.16) can also include the quantities of other goods as conditioning variables, as derived from a two-stage budgeting model; such quantities are not traditionally used in hedonic models.

In the next section we consider how the composite education-and-housing $(ln\, c_{0}^{h})$ commodity is calculated using the FES data, as also how the unobserved education component (w_{0j}^{h}) for households with children in state schools is interpolated from that of households with children in private education by applying Heckman-type methods.

3.3 Extrapolating the Composite Commodity and Education Component

We apply a Heckman estimation approach to approximate the housing component of the composite commodity by the calculated imputed rent, which for the case of households with children in state schools also includes the education component. Imputed rent refers to the rent owners would have had to pay themselves to live in the property they own, and for the case of households in rent-free accommodation to the rent they would normally have had to pay. Following the theoretical model proposed above and having in mind that housing expenditure of households with children in state schools also embody the education expenditure, we approximate $ln\, c_{0}^{h}$ using the calculated imputed rent in the following way.

An equation determining the sample selection has been used to model whether a household rents a house (furnished or unfurnished) as a function of characteristics of the house and the household. A model for household expenditure on actual rent is also specified as a function of a subset of the characteristics used in the selection equation, and a term correcting for the bias due to sample selection; i.e. the error terms from the actual rent expenditure equation and the selection model are allowed to be correlated since the expenditure on rent is observed only for households that do not own their house and, therefore, likely to be at the lower end of expenditure distribution. Variables on housing characteristics (total rooms, heating, region e.tc), household characteristics (number of adults, number of children, e.tc) and also expenditure on council, water and sewerage tax are included in both structural and selection equation. Income sources of the hrp and age of hrp were also included into the selection equation for identification purposes. Income sources can be used as instruments in the sense that they are important factors when people take the decision to rent or buy a house[48]. So, if people have a permanent source of income this will increase their probability to buy a house.

After the estimation of the two models we construct predictions about the rent expenditure for all households, by extrapolating the rent expenditure for households that own their house from the estimated equation obtained for households that rent a house, and by multiplying it with the probability of a household renting a house[49].The extrapolated values form the imputed rent for owner-occupier households.

In addition, we approximate the education component using again the Heckman sample selection model by applying a similar approach. An equation determining the sample selection is used to model whether the children in a household attend private schools as a function of characteristics of the household head and other members. A model for household expenditure on education is also specified as function of a subset of the characteristics which are used in the selection equation and a term correcting for the bias due to sample selection; i.e. the error terms from the education expenditure equation and the selection model are allowed to be correlated, since expenditure on education is observed only for the households with children in private schools, which are likely to be at the top end of expenditure distribution. After the estimation of the two models predictions about the education expenditure for all households is constructed by extrapolating the education expenditure for households with children in state schools from the estimated equation obtained for households with children in private schools, and multiplying it by the

[48] Wages, self-employment and annuities are all insignificant if we included them into the structural equation (rent expenditure equation).

[49] Although Family Expenditure Surveys (EFS) includes information on actual rent paid by tenants, this information does not apply in the case of house owners.

probability that the children in the household attend private school. The extrapolated values for the education expenditure form the notional (unobserved) education expenditure for households with children in state schools.

The rest of the section describes of how the extrapolation of education component from a standard selection Heckman model introduced by Heckman (1976) is conducted. A regression model is specified for the log of education expenditure (y_h)

$$y_h = \beta_1' x_{1h} + u_{1h} \tag{3.17}$$

where x_{1h} is a vector of characteristics (e.g. total expenditure, region) that affect education expenditure. The dependent variable y_h is observed for a household h if the household has children attending private schools. Children are sent to private schools when the cost of private schooling is lower than the minimum cost the household is willing to incur via the purchase of a house in a particular location served by a state school of the same quality. We denote this difference by d_h^* and we defined a variable d_h, which take the value 1 if $d_h^* > 0$ and the value 0 if $d_h^* \leq 0$. The variable d_h indicate whether a household has children in private (d_h=1) or in state (d_h =0) school. Hence, we specify a selection equation

$$d_h^* = \beta_2' x_{2h} + u_{2h} \tag{3.18}$$

such that y_h is observed and d_h=1 if d_h^* > 0 and y_h is not observed and d_h= 0 if $d_h^* \leq$ 0.Thus equation (3.18) constitutes a binary choice model. The vector of characteristics x_{2h} includes all characteristics in x_{1h} as well as some additional characteristics that affect a household's choice between private and state education. The error terms in (3.17) and (3.18) are jointly normally distributed with mean zero and variances Var $(u_{1h}) = \sigma_1$, Var $(u_{2h}) = 1$, and correlation, Corr(u_{2h},u_{2h}) = ρ.

Equation (3.17) and (3.18) form a standard selection model (Gronau 1974, Heckman 1976), which can be estimated using either maximum likelihood (see e.g. Verbeek 2002) or the two-step procedure proposed by Heckman (1979). Estimation results obtained by maximum likelihood are shown in Tables B2.1 and B2.2 in the Appendix of the chapter.

After the estimation of the parameters in (3.17) and (3.18) we extrapolate (predict) log education expenditure for all households in the sample and therefore for those with children in state schools, by computing the unconditional expectation of y_h, assuming that the expected education expenditure is zero for households with children in state schools. The probability that the children in a household attend private schools is determined by the selection equation for all households and the predicted values of the probability are denoted by Pr (d_h=1). Thus the extrapolated log education expenditure for the entire sample, denoted by y_h^* , is given by

$$y_h^* = \Pr(\widehat{d_h = 1}) \times E(y_h | \widehat{d_h} = 1) \quad (3.19)$$

where $E\ (y_h | \widehat{d_h} = 1)$ are the fitted values from regression equation (3.17), estimated with sample selection methods, that take into account the sample selection bias[50]. Despite the fact that the predicted values $E\ (y_h | \widehat{d_h} = 1)$ can only be directly compared to observed education expenditure, these values can be computed for the whole sample. Thus, (3.19) provides the extrapolated (predicted) values of log education expenditure for the whole sample. The extrapolated log education expenditure for households with children in state schools, which we also term as notional education expenditure, are used in the analysis that follows. The same procedure has been implemented for the calculation of a composite commodity. The estimation results obtained by maximum likelihood are shown in Table B2.3 in the Appendix of the chapter.

3.4 Empirical Analysis

3.4.1 Data

The data used in the empirical analysis are drawn from the UK Family Expenditure Survey (FES) conducted between 1994 and 1997, the only period for which the FES provides data concerning the purchase value of the property of the participating households. The FES data also include information on income and expenditure of households and their members (including expenditure for rent), as well as a large number of individual and household characteristics. A sample of households with properties acquired after 1983 is chosen and all property purchase values in the data are expressed in prices corresponding to the year of the survey. The households consist of two adults and one to four children up to 15 years of age, with at least one child attending primary or secondary school. The sample chosen is

---

[50] The extrapolated education expenditure y_h^* is the estimated unconditional expectation of y_h for each observation.

such that children of schooling age in the same household either all attend private or all attend state schools (i.e. households with children in both private and state schools are excluded from the sample)[51].

The data used for the calculation of the composite commodity in the empirical estimation of a Heckman model are also drawn from the UK FES carried out between 1994-1997 and include all households with and without children whose head is under retirement age. Variables on housing characteristics (total rooms, heating, region e.tc), household characteristics (number of adults, number of children, etc) and also expenditure on council and water tax are included in both structural and selection equations. Income sources of head as also the age of head, occupation of head, etc, are included into the selection equation for identification purposes. After all, the imputed rent expenditure which corresponds to the expenditure of the composite education-and housing commodity are matched with the sample of 2 adults and one to four children as described above. Table 3.1 presents the average actual and imputed rent expenditure expressed in GBP.

Table 3.1: Average Rent Expenditure (GBP)

	Actual Rent	Imputed Rent
Weekly expenditure	79.55	97.55
Households number[a]	1695	19191

Note: [a] 8.83% of the sample are households that rent a house

Table 3.2 shows the sample means of some key variables employed in the analysis, for the two groups of households, i.e. those with children in state and those with children in private schools. Households with children attending private schools earn and spend more, on average, per week than households with children attending state schools. The house purchase value stated in the data (expressed in prices corresponding to the year of the survey) by households with children in private schools is, on average, about twice as high as that recorded for households with children in state schools. Finally the average

[51] The reason we restricted the sample to properties purchased after 1983 was the availability of the Halifax House Price Index, which was used to transform the stated purchase prices into values corresponding to the year of survey. Let E_x be the purchase price of a property bought in year x which is provided in the survey conducted in year y. Then the transformed purchase value of the property (corresponding to the composite expenditure on housing and education) used in the analysis, denoted by E, is given by $E = E_x \frac{I_y}{I_x}$, where I_y and I_x is the Halifax House Price Index in year y and x respectively.

calculated composite commodity expenditure is higher for households with children in private schools.

Table 3.2: Sample means (GBP)

	Households with children in state schools N=2697	Households with children in private schools N=176
Composite commodity expenditure (weekly)	109.41	131.48
Property purchase value stated[a]	64,065.00	111,707.80
Weekly income	562.57	841.76
Weekly total expenditure[b]	379.01	517.57

Notes: [a] Property purchase values in the data are expressed in prices corresponding to the year of the survey; [b] Total expenditure includes expenditure on nondurable goods, durable goods and composite commodity expenditure "housing-and-education".

3.4.2 Results

Table 3.3 reports the results obtained from OLS estimation analysis of structural models for both Translog and Quadratic Logarithmic cost functions as described in Eq. (3.10) and (3.15). In the columns under the heading `Model I' the results of the unconditional structural models are presented (in terms of other quantities used in the upper stage) while in the columns under the heading `Model II' the parameters' estimates of the structural models when the quantities of other categories of total expenditure is conditioned are also reported. The parameter estimate of ω_{01}^{h} corresponds to the shadow price of education relative to housing which it appears to be positive and significant in all models.

The magnitude of ω_{01}^{h} is higher in the Translog models because total expenditure was not included in the regressions. However an important result is that the education component of households with children in private schools is actually not significant in any model apart from Model II of Translog form, in which it appears to be negative. Furthermore, households with children in private schools seem to have higher composite commodity expenditure in the Translog models, apparently because they are usually richer than households with children in state schools (i.e. in addition to high fees paid to private schools they also spend more on housing). This is also confirmed in the Quadratic Logarithmic models where this effect appears to be insignificant because total expenditure is included in the empirical estimation. Finally, a significant variation in the cost of the composite commodity between the different survey years is observed.

Following Browning and Meghir (1991), the separability of the composite commodity from the quantities of other goods can be tested empirically using Eq. (3.15) by defining $\delta_i = \delta_{i0} + \sum_{s=1}^{n} \delta_{is}\bar{q}_s \; i = 1,2$, where $\bar{q}_s$, $s = 1,2,\dots,n$, are the quantities of other goods which are assumed to be determined, and testing whether δ_{is} are zero for all i and s. Rejection of the null hypothesis indicates that the composite commodity cannot be separated from other commodity groups. In vector $\bar{q}$ the broad categories of food, alcohol and tobacco, clothing, fuel and light, household goods and services and leisure goods and services are included. Here, the null hypothesis is rejected (F-test=28.11) thus we conclude that the composite commodity cannot be separated from all other categories of goods. Consequently, both Models "I" that hold for the separability assumption cannot be accepted.

Table 3.3- Structural Form Models: OLS estimation results

	Translog				Quadratic			
Dep.Variable (log): Composite Commodity	Model I		Model II		Model I		Model II	
	Coef.[a]	s.e[b]	Coef.[a]	s.e[b]	Coef.[a]	s.e[b]	Coef.[a]	s.e[b]
w_{01}^{h}	2.015***	0.079	1.517***	0.074	1.278***	0.064	1.271***	0.062
(w_{01}^{h}) x (hholds in priv sch.)	-0.262	0.218	-0.468**	0.199	-0.151	0.183	-0.111	0.183
Hholds with children in private schools	0.132***	0.033	0.086***	0.030	0.044	0.028	0.039	0.027
Ln total expenditure	-	-	-	-	1.204***	0.210	1.045***	0.241
Ln total expenditure sq	-	-	-	-	-0.071***	0.018	-0.043**	0.021
Quantities[c]								
Qfd	-	-	0.149***	0.015	-	-	-0.070***	0.016
Qalctob	-	-	-0.057***	0.020	-	-	-0.203***	0.018
Qcf	-	-	0.056***	0.010	-	-	-0.076***	0.011
Qfl	-	-	0.251***	0.059	-	-	-0.005	0.051
Qls	-	-	0.136***	0.029	-	-	0.002	0.025
Qlg	-	-	0.037**	0.015	-	-	-0.084***	0.015
Qtr	-	-	0.068***	0.012	-	-	-0.097***	0.014
Qhs	-	-	0.081***	0.010	-	-	-0.043***	0.009
Qhg	-	-	0.075*	0.041	-	-	-0.075**	0.037
Survey Year (1994):								
1995	0.050***	0.010	0.030***	0.010	0.021**	0.008	0.021***	0.008
1996	0.096***	0.011	0.111***	0.011	0.127***	0.009	0.116***	0.009
1997	0.098***	0.010	0.082***	0.010	0.033***	0.009	0.024***	0.009
Constant	4.412***	0.011	4.249***	0.015	-0.110	0.626	0.042	0.706
N	2873	-	2873	-	2873	-	2873	-
R2	0.288	-	0.389	-	0.520	-	0.563	-

Notes: [a]The symbols *, ** and *** denote statistical significance at 10%, 5% and 1%. [b]The reported standard errors are robust to heteroscedasticity; [c] these quantities constitute the following subcategories of nondurable goods:food, alcohol and tobacco, clothing and footwear, fuel and light, leisure goods, leisure services, transport, household goods, household services.

The first two columns of Table 3.4 report the results obtained from OLS estimation of the Quadratic Logarithmic structural model, as described in Eq. (3.16). In addition, the last column reports the results from the hedonic model. In the first two columns, the education component (ω_{01}^{h}) appears to be of lower magnitude than in the corresponding columns of Table 3.3 because of extra conditioning on household characteristics and quantities. As it has been already mentioned in the previous section, the parameters estimates of the reduced form models do not represent the shadow prices of whatever quantity/quality indicators are used as explanatory variables for the composite commodity. Thus, in the hedonic model ω_{01}^{h} is likely to be approximated by variables indicating the quantity/quality of component items.

At this point, the (log) notional education expenditure is used as a measure of willingness to pay for higher quality education. The hedonic model gives a positive and significant elasticity of the composite commodity expenditure with respect to notional education expenditure (approximated by purchase property values), of magnitude 0.218.

Thus, if a household doubles its notional education expenditure the house price increases by 21.8%. The next subsection analyses a homogeneous sample of 2 adults-2 children with children only in state schools and examines this relationship in more depth using both parametric and non-parametric analysis. Looking at the effects of the remaining characteristics, we observe that they are very similar in magnitude and significance in the two Quadratic Logarithmic models. As expected, total household expenditure affects positively the dependent variables.

Households in detached houses, followed by those in semi-detached houses incur a higher expenditure on the composite commodity and purchase property values than households in other house types. The number of children has positive and linear effect in the two QUAIDS models, while the squared term of this variable was negative and significant in a hedonic model. In all three models the number of bedrooms affects positive and significant the dependent variables.

As far as the region of residence of a household is concerned, Greater London followed by South East are associated with the highest expenditure on the composite commodity education-and-housing as also in house prices, while households in Northern Ireland appeared to have the lowest effect. Again, a significant variation appears in the cost of the composite commodity and purchase property values between the different survey years.

Table 3.4- Conditional Structural and Hedonic Models: OLS estimation results

	Structural Form (Quadratic Logarithmic)				Hedonic	
	Model III		Model IV		Model I	
Dep.Variable (log):	Composite Commodity				House Prices	
	Coef.[a]	s.e[b]	Coef.[a]	s.e[b]	Coef.[a]	s.e[b]
w_{01}^{h}	0.397***	0.053	0.469***	0.053	-	-
(w_{01}^{h}) x (hholds in priv sch.)	0.078	0.093	0.084	0.094	-	-
y_h^*	-	-	-	-	0.218***	0.026
(y_h^*) x (hholds in priv sch.)	-	-	-	-	0.022	0.038
Hholds with children in private schools	0.018	0.014	0.019	0.014	0.178*	0.107
Ln total expenditure	0.414***	0.135	0.419***	0.140	0.228***	0.027
Ln total expenditure sq	-0.011	0.011	-0.007	0.012	-	-
Quantities[c]						
qfd	-	-	-0.040***	0.008	-	-
qalctob	-	-	-0.073***	0.010	-	-
qcf	-	-	-0.015***	0.005	-	-
qfl	-	-	0.092**	0.036	-	-
qls	-	-	-0.007	0.013	-	-
qlg	-	-	-0.025***	0.007	-	-
qtr	-	-	-0.028***	0.007	-	-
qhs	-	-	-0.019***	0.004	-	-
qhg	-	-	-0.051***	0.018	-	-
General Characteristics						
Age of head	0.005	0.003	0.005	0.003	0.041***	0.011
Age of head squared	-0.000	0.000	-0.000	0.000	-0.000***	0.000
Number of children	0.028***	0.011	0.025**	0.011	0.068	0.044
Number of children squared	0.001	0.002	0.001	0.002	-0.024**	0.009
Number of vehicles	-0.014***	0.003	-0.016***	0.003	0.019	0.013
Number of bedrooms	0.065***	0.004	0.060***	0.004	0.105***	0.017
House Type (other)[d]:						
Detached	0.089***	0.012	0.081***	0.012	0.408***	0.063
Semi-detached	0.043***	0.012	0.040***	0.011	0.136**	0.061
Terraced	-0.014	0.012	-0.015	0.011	-0.075	0.061
Region (South west)[d]:						
Yorkshire and Humberside	-0.085***	0.008	-0.081***	0.008	-0.181***	0.032
North East	-0.157***	0.009	-0.152***	0.009	-0.188***	0.042
Greater London	0.303***	0.008	0.298***	0.008	0.380***	0.033
North West	0.078***	0.008	0.083***	0.008	-0.109***	0.031
East Midlands	-0.162***	0.009	-0.152***	0.009	0.059	0.042
West Midlands	-0.118***	0.008	-0.112***	0.008	0.024	0.034
East Anglia	-0.096***	0.012	-0.086***	0.012	0.203***	0.054
South East	0.115***	0.007	0.117***	0.007	0.268***	0.025
Wales	-0.123***	0.011	-0.114***	0.011	-0.022	0.053
Scotland	-0.032***	0.008	-0.029***	0.008	-0.149***	0.037
Northern Ireland	-0.355***	0.019	-0.339***	0.018	-0.233***	0.064
Survey Year (1994)[d]:						
1995	0.013***	0.005	0.015***	0.005	0.009	0.022
1996	0.097***	0.005	0.093***	0.006	0.163***	0.021
1997	0.039***	0.005	0.035***	0.005	0.122***	0.021
Constant	2.170***	0.407	2.045***	0.418	7.627***	0.240
N	2873	-	2873	-	2873	-
R2	0.860	-	0.865	-	0.615	-

Notes: [a]The symbols *, ** and *** denote statistical significance at 10%, 5% and 1%. [b]The reported standard errors are robust to heteroscedasticity. [c] these quantities constitute the following subcategories of nondurable goods: food, alcohol and tobacco, clothing and footwear, fuel and light, leisure goods, leisure services, transport, household goods, household services; [d]The variable in the brackets is excluded from the regression and is used as the benchmark for comparison.

3.4.3 A hedonic Analysis Approach

In this subsection the relationship between the purchase property values and the notional education expenditure is examined using a homogeneous sample of 2 adults-2 children in state schools. We suggest that the notional education expenditure can also be viewed as an indicator of willingness to pay for quality state education and can be used to assess the contribution of the quality of state schools to housing values, in the sense that higher notional education expenditure by households with children in state schools indicates higher quality of education. In order to investigate this relationship both semi-parametric and parametric analysis are applied.

3.4.3.1 Semi-Parametric Analysis

First the relationship between the purchase property values ($ln\, P_0^h$) and the notional education expenditure (y_h^*) using semi-parametric analysis is explored.

The reduced form model is recast into an empirical model of the form

$$ln\, P_0^h = \Psi' Z_h + f(y_h^*) + e_h \tag{3.20}$$

where Z_h is a vector of general household and house characteristics and total household expenditure, $f(y_h^*)$ is unknown function of the notional education expenditure to be estimated using a semi-parametric regression method and e_h is a random error term.

Equation (3.20) is fitted only to data for households with children in state schools. We follow a nearest neighbour estimation approach proposed by Estes and Honore (1995) and employed in Lyssiotou et al. (1999). First, the data are sorted by the continuous variable y_h^* and the differences $\Delta log P_0^h = log P_0^h - log P_0^{h-1}$ and $\Delta Z_h = Z_h - Z_{h-1}$ were computed. Second, the regression $\Delta log P_0^h = \widehat{\Psi}' \Delta Z_h + \eta_h$ is estimated to get a consistent estimator of Ψ, denoted by $\widehat{\Psi}$. Then the residuals $\widehat{e_h} = log P_0^h - \widehat{\Psi}' Z_h$ are computed. Finally we estimate the non-parametric regression of the form

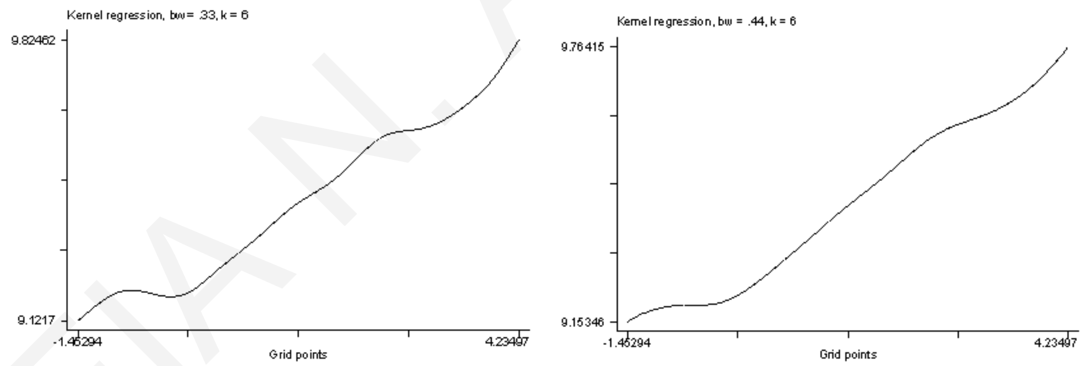
$$\widehat{e_h} = f(y_h^*) + r_h \tag{3.21}$$

to explore the relation between the composite expenditure education-housing and the notional education expenditure, net of the influence of other characteristics. As (3.21) is a non-parametric model no assumptions are made about the distribution, homoscedasticity or serial correlation of the error term r_h. Using (3.21) a kernel weighted regression estimator has been computed for f (y_h^*) (e.g. Greene, 2003).

Figure 3.1 plots the kernel weighted regression estimate for f (y_h^*), computed using the Gaussian kernel, for different bandwidth sizes.

Even though the larger bandwidth produces a smoother estimate, some form of non-linearity in the relationship between education expenditure and the composite commodity expenditure at very low and very high levels of the education expenditure is observed. This finding is exploited in the specification of the parametric models as describe above. In particular we investigate the contribution of the notional education expenditure (y_h^*) to the expenditure of the composite commodity education-and-housing (as reflected in the purchase value of the property), using alternative parametric specifications motivated by the semi-parametric analysis. As previously, this analysis is conducted for households with children in state schools, whose education expenditure is not directly observed in the data

Figure 3.1: Kernel estimates for $f(y*)$



bw =0.33 bw=0.44

3.4.3.2 Parametric Analysis

The general form of the parametric model is given by

$$log\, P_0^h = \Phi' Z_h + F(y_h^*) + \varepsilon_h \qquad (3.22)$$

where Z_h is a vector of general household and house characteristics and total household income, ε_h is a random error term and $F(y_h^*)$ is a function of the notional education expenditure, which in the case of where $log\, c_0^h$ is approximated by purchase property values defined in three alternative ways as follows:

$$F(y_h^*) = \varphi_1\, y_h^* \quad (3.23)$$

$$F(y_h^*) = \varphi_1\, y_h^* + \varphi_2\, (y_h^*)^2 \quad (3.24)$$

$$F(y_h^*) = \varphi_1\, y_h^* + \varphi_2\, (y_h^*)^2 + \varphi_3\, (y_h^*)^2 \quad (3.25)$$

Equation (3.22) is estimated for the specifications resulting from (3.23) - (3.25) by OLS and the results are shown in Table 3.5. Model I gives a positive and significant elasticity of the composite commodity expenditure with respect to notional education expenditure, equal to 0.119. Thus, if a household doubles its notional education expenditure (i.e. its willingness to pay for a higher quality state education) the composite 'education-and-housing' cost - defined here by the value of its property - increases by 11,9%. Model II tends to favour a linear relationship between the purchase property value and the notional education expenditure, since the quadratic term does not seem to affect the dependent variable[52]. The linear effect for Model II is of about the same magnitude as the effect obtained from Model I. Model III on the other hand, appears to capture a non-linear relationhsip between notional education expenditure and purchase property values.

The quadratic and cubic terms of education expenditure are both individually and jointly highly significant[53]. Finally, looking at the effects of the remaining characteristics, we observe that they are very similar in magnitude and significance in the three models under consideration. As expected, total household expenditure (budget), affects positively the dependent variable but with an estimated effect less than unity. This suggests that the cost of the 'education-and-housing composite commodity is not sensitive to changes in the level of households budget (income).

As expected house characteristics (such as the number of bedrooms, being detached or semi-detached etc) also influence positively the dependent variable. As far as the region of residence is concerned, households in Greater London, followed by South East, have the

---

[52] In Model II the F-test for the joint significance of y_h^* and $(y_h^*)^2$ gives a p-value equal to 0.016 (F-statistic=4.11).

[53] In Model IV the F-test for the joint significance of $(y_h^*)^2$ and $(y_h^*)^3$gives a p-value equal to 0.047(F-statistic=5.38). The F-test for the joint significance of y_h^* , $(y_h^*)^2$ and $(y_h^*)^3$ gives a p-value equal to 0.0096 (F-statistic=3.82).

highest while those in Northern Ireland the lowest expenditure on the composite education-and-housing commodity. Finally, we observe a significant variation in the cost of the composite commodity between the different survey years, with the cost peaking in 1996. In particular composite education- and-housing expenditure in 1996 was about 10% higher than in 1994, while in 1995 was about 5% lower than in 1994. This finding stems largely from the evolution of property prices during the period 1994-1997.

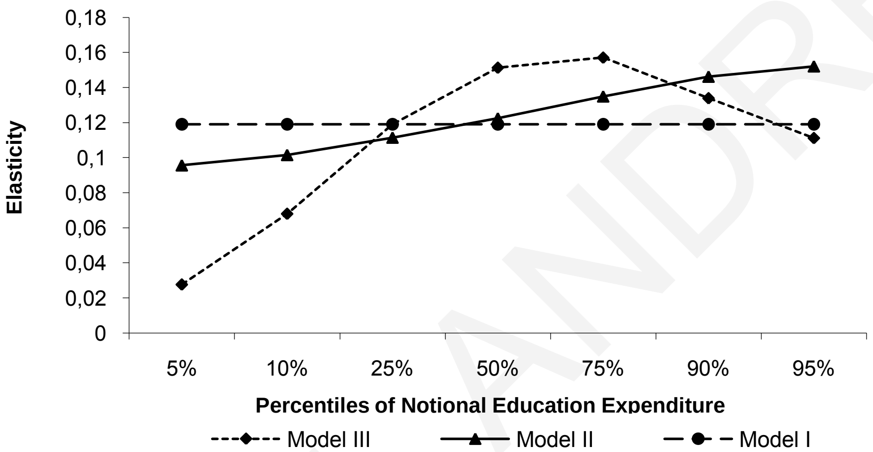
Table 3.5: OLS estimation results for hedonic models
(Sample of 2 adults and 2 children)

	Model I		Model II		Model III	
Dependent Variable (log)	House Prices		House Prices		House Prices	
	Coef.[a]	s.e[b]	Coef.[a]	s.e[b]	Coef.[a]	s.e[b]
Notional Education Exp						
y_h^*	0.119***	0.043	0.103**	0.044	0.077*	0.046
$(y_h^*)^2$	-	-	0.008	0.007	0.047**	0.022
$(y_h^*)^3$	-	-	-	-	-0.009*	0.005
Log Total Expenditure	0.191***	0.055	0.186***	0.055	0.185***	0.055
House Characteristics:						
Number of rooms	0.032	0.021	0.028	0.022	0.027	0.022
Number of bedrooms	0.071***	0.026	0.072***	0.026	0.075***	0.026
House Type (other)[c]:						
Detached	0.405***	0.088	0.404***	0.088	0.393***	0.088
Semi-detached	0.147*	0.084	0.148*	0.084	0.139	0.085
Terraced	-0.050	0.084	-0.051	0.084	-0.060	0.084
Otherl Characteristics						
Age of head	0.036**	0.015	0.036**	0.015	0.036**	0.015
Age of head_sq	-0.000**	0.000	-0.000**	0.000	-0.000**	0.000
Number of vehicles	0.041**	0.018	0.040**	0.018	0.038**	0.018
Region (South West)[c]:						
Yorkshire and Humb.	-0.133**	0.065	-0.116*	0.068	-0.111	0.068
North East	-0.187***	0.058	-0.173***	0.060	-0.170***	0.060
Greater London	0.492***	0.055	0.508***	0.058	0.508***	0.058
North West	-0.089	0.059	-0.073	0.062	-0.071	0.062
East Midlands	0.051	0.106	0.068	0.110	0.072	0.109
West Midlands	-0.031	0.054	-0.015	0.057	-0.012	0.057
East Anglia	0.223*	0.130	0.238*	0.132	0.233*	0.131
South East	0.382***	0.076	0.401***	0.080	0.407***	0.080
Wales	-0.082	0.116	-0.065	0.119	-0.063	0.119
Scotland	-0.179***	0.047	-0.170***	0.048	-0.169***	0.048
Northern Ireland	-0.288*	0.165	-0.285*	0.165	-0.328**	0.167
Survey Year (1994)[c]:						
1995	-0.066**	0.026	-0.066**	0.026	-0.067***	0.026
1996	0.071**	0.029	0.070**	0.029	0.069**	0.029
1997	0.070***	0.027	0.071***	0.027	0.071***	0.027
Constant	8.185***	0.437	8.217***	0.441	8.230***	0.441
N	1577	-	1577	-	1577	-
R2	0.601	-	0.601	-	0.602	-

Notes: [a]The symbols *, ** and *** denote statistical significance at 10%, 5% and 1%. [b]The reported standard errors are robust to heteroscedasticity; [c]The variable in the brackets is excluded from the regression and is used as the benchmark for comparison.

Figure 3.2 plots the rates of change of the composite education-and-housing expenditure[54] associated with changes in the notional education expenditure (with all other factors held constant), i.e. the elasticity of the notional education expenditure with respect to education-and-housing expenditure, for different percentiles of the distribution of the notional education expenditure. The lower 5% and upper 95% percentiles correspond to a weekly notional education expenditure of 0.63 GBP and 21.52 GBP, respectively. In Model 2 the elasticity declines as we move to higher levels of education expenditure.

Figure 3.2: Elasticity for differnet percentiles of y*



In Model 3 elasticity increases with the notional education expenditure up to the 50-75% percentile and decreases for higher levels of education expenditure. Hence, households willing to spend too much or too little on education have expenditure on education-and-housing less responsive to changes in the notional education expenditure than households in the middle of the notional education expenditure distribution. Given the positive relationship between notional education expenditure and the household budget, this is in line with results reported by Romano and Epple (1996b) that middle income households prefer higher public expenditure on education compared to households with low or high income.

54 Recall that in the context of our analysis the value of property reflects the expenditure on the composite commodity, termed education-and-housing.

3.5 Conclusion

The consumer demand based approach proposed in this chapter of the thesis can be used as an alternative to the hedonic approach to estimate the capitalisation of local state school quality to house prices in countries, like the UK, where family expenditure surveys do not contain information about the location of households. The idea here is to exploit the fact that the increased price of houses in the catchment area of high quality state schools reflects the household's willingness to pay in order to secure an education level for its children above the «minimum» level provided free by the state.

Households with children in private schools report their education and housing expenditures separately; while those with children in state schools report only housing expenditures that also includes education costs in the form of higher house price paid for location in the catchment area of a good quality state school. For the latter households the unobserved education component as also the composite education-and-housing expenditure is estimated by Heckman-type methods. The theoretical model is based on the notion of separability and two-stage budgeting; and using Translog and Quadratic Logarithmic functional forms a consumer demand system is derived allowing for the estimation of the shadow price of education relative to housing. The model is applied to data drawn from the UK Family Expenditure Surveys (FES) for the years 1994-1997 where information about the purchase value of the property of the participating households is available. The results demonstrate the presence of a positive and significant education component in the composite house-and-education expenditure with the relative price of education to housing to vary between 0.4 to 2, largely depending on its conditioning on other characteristics of the household and the quantities of the nondurable goods included in the household budget (total expenditure).

In addition, a usual hedonic analysis is performed based on both parametric and semi-parametric methods, where the notional education expenditure estimated by Heckman type methods is included among the explanatory variables (as an indicator of willingness to pay for high school quality). The results suggest a non-linear relationship, with middle-income households being more willing to pay for the education of their children through housing (i.e. having a higher elasticity of the notional education expenditure with respect to education-and-housing).

Appendix B

B1. Subgroups of Non-durable Goods

The nine subgroups commodities include in the category of nondurable goods were the following:

a. Food and Catering
b. Alcohol and Tobacco
c. Fuel and Light
d. Clothing and Footwear
e. Transport
f. Household Goods
g. Household Services
h. Leisure Goods
i. Leisure Services

For the creation of the above groups, in most of the cases the distinction of RPI[55] is followed (e.g. food and catering, clothing and footwear, fuel and light), while in the remaining cases all durable sub-categories of expenditures were excluded and a new group created (named "Durables") which is also included in the total expenditure of a household.

The sub-categories for the broad groups mentioned above are:

- Food and Catering (as RPI distinction)
- Alcohol and Tobacco (as RPI distinction)
- Fuel and Light (as RPI distinction)
- Clothing and Footwear (as RPI distinction)

[55] The Retail Prices Index (RPI) is the most familiar general purpose domestic measure of inflation in the United Kingdom. It is available continuously since June 1947. The RPI measures the average change in prices by calculating the change in price of a fixed basket of goods and services representing the items bought by all UK households (http://www.statistics.gov.uk).

- Transport (maintenance of motor vehicles, petrol and oil, vehicle tax and insurance, fares and other travel costs such as rail fares, bus and coach fares, and other travel costs)
- Households Goods (household consumables, pet care)
- Household Services (postage, telephones, telemessages, domestic services, fees and subscription[56])
- Leisure Goods (CDs and tapes, toys, photography and sports goods, books and newspaper, gardening product)
- Leisure Service (television licence and rentals, entertainment and other recreation)
- Durables (furniture, furnishings, electrical appliances, other household equipment and purchase of motor vehicles, audio visual equipment)

[56] Excluded education

B2. Tables

Table B2.1: Estimation results of the sample selection model[a]: Composite Commodity

	Rent Expenditure		Selection Equation[d]	
	Coef.[b]	s.e	Coef.[b]	s.e
Constant	2.473***	(0.106)	0.994***	(0.191)
Log total household Expenditure	0.251***	(0.021)	0.214***	(0.035)
Region (South West)[c]:				
Yorkshire and Humberside	-0.108***	(0.041)	-0.251***	(0.070)
North East	-0.208***	(0.049)	-0.321***	(0.079)
Greater London	0.290***	(0.035)	-0.161**	(0.063)
North West	0.043	(0.038)	-0.174***	(0.064)
East Midlands	-0.225***	(0.040)	-0.101	(0.071)
West Midlands	-0.158***	(0.041)	-0.267***	(0.069)
East Anglia	-0.168***	(0.048)	-0.000	(0.084)
South East	0.072**	(0.032)	-0.052	(0.056)
Wales	-0.201***	(0.048)	-0.137*	(0.081)
Scotland	-0.055	(0.053)	-0.636***	(0.088)
Northern Ireland	-0.336***	(0.089)	-0.940***	(0.134)
Other Characteristics:				
Number of rooms	0.049***	(0.008)	-0.076***	(0.014)
Number of vehicles	-0.014	(0.014)	-0.136***	(0.024)
Number of workers	-0.050**	(0.023)	-0.092**	(0.043)
Number of economically active persons	0.023	(0.022)	0.010	(0.041)
Professional (head)	0.102***	(0.035)	0.152**	(0.060)
Number of adults	0.096***	(0.015)	0.086***	(0.031)
Number of children	0.029***	(0.011)	-0.168***	(0.016)
Council tax	0.001**	(0.000)	-0.006***	(0.000)
Council water tax	-0.000	(0.000)	-0.001	(0.001)
Heating type (other)[c]:				
Electricity	0.161***	(0.029)	-0.257***	(0.053)
Gas	0.142***	(0.022)	-0.417***	(0.036)
Oil	0.112*	(0.062)	-0.192**	(0.093)
House Type (other)[c]:				
Detached	0.073*	(0.042)	-0.316***	(0.065)
Semi-detached	0.044	(0.032)	-0.464***	(0.049)
Terraced	-0.002	(0.026)	-0.356***	(0.043)
Source of Income (other)[c]:				
Investment	0.046	(0.118)	-0.854***	(0.198)
Social security benefits	0.305***	(0.030)	-0.856***	(0.087)
Wages	-	-	-0.773***	(0.089)
Self-employment	-	-	-0.568***	(0.101)
Annuities	-	-	-1.204***	(0.187)
Age of head	-	-	-0.029***	(0.001)
Survey Year (1994)[c]:				
1995	-0.010	(0.025)	0.085**	(0.042)
1996	0.009	(0.025)	0.065	(0.042)
1997	0.014	(0.024)	0.169***	(0.042)

Notes: [a] The number of observations is 1695 for the rent expenditure regression (number of households that pay rent) and 19191 for the selection equation. The estimated standard error of the rent expenditure equation is 0.358. The estimated correlation between the errors of the rent expenditure and selection equations is -0.319 (s.e.=0.085) and the LR test for the independence of the two equations ($\rho=0$) gives a p-value equal to 0.237 (chi-squared statistic=1.40); [b]The symbols *, ** and *** denote statistical significance at 10%, 5% and 1%; [c] The variable in the brackets is excluded from the regression and is used as the benchmark for comparison; [d] For identification reasons we include extra variables in selection equation after the examination test of endogeneity (the predicted errors obtained from selection equation were statistically correlated with rent expenditure, t=-2.67). The joint chi-squared for the extra variables is equal to 537.59 (p=0.000).

Table B2.2: Estimation results of the sample selection model[a]: Education Component

(1-4 children)

	Education Expenditure		Selection Equation[d]	
	Coef.[b]	s.e	Coef.[b]	s.e
Constant	-1.752	(2.014)	-5.339***	(0.829)
Log total household Expenditure	0.477*	(0.263)	0.632***	(0.129)
Region (South West)[c]:				
Yorkshire and Humberside	-0.498	(0.540)	-0.168	(0.246)
North East	-0.520	(0.542)	0.285	(0.245)
Greater London	0.130	(0.475)	0.687***	(0.210)
North West	-0.630	(0.562)	-0.115	(0.232)
East Midlands	-1.257**	(0.490)	0.091	(0.232)
West Midlands	-0.750	(0.563)	0.031	(0.243)
East Anglia	-1.321*	(0.736)	-0.453	(0.330)
South East	-0.389	(0.427)	0.224	(0.188)
Wales	-1.185	(0.731)	-0.588*	(0.347)
Scotland	-0.472	(0.493)	0.204	(0.221)
Northern Ireland	-1.503	(1.005)	-0.129	(0.387)
Other Characteristics:				
Number of rooms	0.306***	(0.064)	0.160***	(0.042)
Number of vehicles	0.168	(0.161)	0.147*	(0.076)
Heating type (other)[c]:				
Electricity	-0.111	(0.861)	0.297	(0.336)
Gas	-0.186	(0.665)	0.179	(0.221)
Oil	-0.290	(0.708)	0.523**	(0.258)
Number of bedrooms	-	-	0.104	(0.087)
Number of children	-	-	-0.497***	(0.071)
House Type (other)[c]:	-	-		
Detached	-	-	-0.224	(0.247)
Semi-detached	-	-	-0.307	(0.247)
Terraced	-	-	-0.202	(0.249)
Source of Income (other)[c]:	-	-		
Wages	-	-	-0.180	(0.294)
Self-employment	-	-	-0.049	(0.307)
Age of head	-	-	-0.021**	(0.008)
Professional (head)	-	-	0.225*	(0.121)
Survey Year (1994)[c]:				
1995	-0.275	(0.291)	-0.025	(0.139)
1996	-0.212	(0.291)	0.227	(0.140)
1997	-0.249	(0.278)	0.150	(0.135)

Notes: [a] The number of observations is 145 for the education expenditure regression (number of households that pay for education) and 2915 for the selection equation. The estimated standard error of the education expenditure equation is 1.213. The estimated correlation between the errors of the rent expenditure and selection equations is 0.591 (s.e.=0.161) and the LR test for the independence of the two equations (ρ=0) gives a p-value equal to 0.019 (chi-squared statistic=5.50); [b]The symbols *, ** and *** denote statistical significance at 10%, 5% and 1%; [c] The variable in the brackets is excluded from the regression and is used as the benchmark for comparison; [d] For identification reasons we include extra variables in selection equation after the examination test of endogeneity (the predicted errors obtained from selection equation were statistically correlated with rent expenditure, t=13.83). The joint chi-squared for the extra variables is equal to 57.93 (p=0.000).

Table B2.3: Estimation results of the sample selection model[a]: Education Component

(2 children)

	Education Expenditure		Selection Equation[d]	
	Coef.[b]	s.e	Coef.[b]	s.e
Constant	-6.213*	(3.609)	-7.102***	(1.230)
Log total household Expenditure	0.938**	(0.426)	0.758***	(0.181)
Region (South West)[c]:				
Yorkshire and Humberside	-1.352*	(0.757)	-0.218	(0.338)
North East	-0.883	(1.013)	-0.033	(0.392)
Greater London	-1.025	(0.740)	0.639**	(0.298)
North West	-1.177	(0.739)	0.082	(0.304)
East Midlands	-2.306***	(0.719)	-0.007	(0.332)
West Midlands	-1.005	(0.793)	0.217	(0.323)
East Anglia	-2.661***	(0.965)	-0.520	(0.445)
South East	-1.667**	(0.655)	0.195	(0.261)
Wales	-2.012**	(1.023)	-0.568	(0.490)
Scotland	-0.445	(0.711)	0.390	(0.295)
Northern Ireland	-2.937**	(1.318)	-0.046	(0.572)
Other Characteristics:				
Number of rooms	0.438***	(0.115)	0.156***	(0.056)
Number of vehicles	0.157	(0.251)	0.183*	(0.106)
Heating type (other)[c]:				
Electricity	-0.384	(1.221)	0.230	(0.462)
Gas	0.516	(0.892)	0.061	(0.300)
Oil	-0.013	(0.943)	0.408	(0.356)
Number of bedrooms	-	-	0.228**	(0.116)
House Type (other)[c]:	-	-		
Detached	-	-	-0.476	(0.331)
Semi-detached	-	-	-0.614*	(0.335)
Terraced	-	-	-0.271	(0.343)
Source of Income (other)[c]:	-	-		
Wages	-	-	-0.065	(0.550)
Self-employment	-	-	-0.047	(0.569)
Age of head	-	-	-0.029**	(0.012)
Professional (head)	-	-	0.301*	(0.157)
Survey Year (1994)[c]:				
1995	0.109	(0.427)	0.152	(0.188)
1996	0.305	(0.463)	0.264	(0.199)
1997	-0.096	(0.405)	0.111	(0.190)

Notes: [a] The number of observations is 78 for the education expenditure regression (number of households that pay for education) and 1689 for the selection equation. The estimated standard error of the education expenditure equation is 1.318. The estimated correlation between the errors of the rent expenditure and selection equations is 0.737 (s.e.=0.164) and the LR test for the independence of the two equations ($\rho=0$) gives a p-value equal to 0.0224 (chi-squared statistic=5.22); [b]The symbols *, ** and *** denote statistical significance at 10%, 5% and 1%; [c] The variable in the brackets is excluded from the regression and is used as the benchmark for comparison; [d] For identification reasons we include extra variables in selection equation after the examination test of endogeneity (the predicted errors obtained from selection equation were statistically correlated with rent expenditure, t=10.07). The joint chi-squared for the extra variables is equal to 17.85 (p=0.000).

Chapter 4

How much is State Schooling Worth to Consumers? The case of the UK

4.1 Introduction

The free of charge provision of a minimum level of education by the state is a world-wide phenomenon, based on both efficiency (positive externalities) and equity (less income inequality) criteria. However, it is not obvious how much the free provision of this private good by the state adds to the welfare of the consumer; and how this provision should be differentiated to reflect on consumer preferences/needs.

While no full-grown framework exists allowing for theoretical and/or empirical analysis of the valuation of state education by the consumer, there are several theoretical articles in the literature addressing welfare issues associated with public provision. In particular, public provision is treated as a means for mitigating market imperfections (such as excludability, imperfect information, externalities etc.) and, under certain conditions, for redistributing income and enhancing efficiency (Epple and Romano, 1996a, 1996b; Besley and Coate, 1991; Blomquist and Christiansen, 1995,1999; Fernandez and Rogerson,1996, 2003). The redistributive aspects of public provision in the presence of a private market, where consumers can pay for extra quality - as in education - are analysed by Ireland (1990). To our knowledge, previous empirical analysis of the welfare implications of publicly provided goods is limited to contingent valuation studies (mainly of environmental goods) and econometric modelling of willingness to pay, also elicited from contingent valuation surveys (see for example Brookshire and Coursey, 1987; Clinch and Murphy, 2001; Hanemann, 1994); or application of hedonic methods, e.g. to value air quality (Smith and Huang, 1995).

The choice between state and private schooling is a widely researched issue in literature, where authors are primarily concerned with factors determining the outcome of this choice (e.g. Buttin, 1998). In contrast to the choice between state and private schooling, very few studies have been concerned with expenditure on education (e.g. Buddin, 1998), especially in the form of supplementing the minimum quantity/quality of education provided free of charge by the state with additional quantity/quality purchased either directly in the form of evening classes or private tuition outside school hours (as it happens, for example in Greece and other Mediterranean countries) or indirectly in the form of a premium paid by households in order to acquire accommodation in the catchment area of a high performing state school (e.g Blank,1999; Gibbons and Machine, 2003).

The main theoretical contribution of this chapter lies in the use of an integrable complete demand system (Gorman, 1981; Lewbel, 1990) to evaluate household utility from publicly provided education (or any other private good). Demand analysis is basically concerned with the explanation of *behaviour differences* between households. Households differ in size, age composition, educational level and other characteristics, and in general, someone expect households with different characteristics to have different expenditure patterns. Generally, one can model differences in behaviour by making demand depend not only on prices and total expenditure but also on a long list of household characteristics and macroeconomic variables (Blundell et al, 1993). The effects of household composition, the number of children, number of adults, the ages of household members are most commonly modelled, but basically many other characteristics can also be included. Such an approach raises the question of whether such demand functions can form a basis for welfare comparisons between households. As Deaton and Muellbauer (1980) argue, such welfare comparisons are being made across households, thus someone can assume that two households who behave identically have identical welfare levels.

The integrability of the demand system is necessary (and sufficient) for the derivation of a money metric to effect welfare comparisons across households with for example different sizes and compositions. These index numbers through which these comparisons are made are known in the literature as equivalence scales. Many empirical applications have been known in the literature for the estimation of equivalence scales based on demand analysis approaches. Some important studies include Deaton and Muellbauer (1986), Pashardes (1991, 1995) and Dickens et. all (1993) where both adult and child equivalence scales were calculated.

At an empirical level the approach proposed here innovates by being the first in the international literature to attempt an econometric estimation of a money measure of the publicly (free of charge) provided education. The money metric associated with free state schooling is estimated using widely available household data - such as data drawn from household expenditure surveys routinely available in many countries, including the UK.

Furthermore, in the case of UK, this is the first study that attempts to connect the quantity and quality of the publicly provided education with the prosperity/welfare of the consumer. The proposed method is applied to data drawn from the UK Expenditure and Food Surveys (EFS) for the years 2001-2007 which containing information about expenditure on a detailed commodity breakdown and a large number of demographic and other household characteristics which are found to be significant on empirical studies of consumer behaviour based on individual household data (Blundell et al, 1993; Lyssiotou et al, 2004).

The novelty of economic valuation of the state education could provide useful information and can help create incentives for more efficiency and greater equity in. It could also guide policies towards a more informed and constructive treatment of the role which can be played by the private education sector. In theory this sector could be as useful for the utility maximization of wealthy households as state education is for the utility maximization of the rest of the population. Besides, under certain conditions, it can be used as a vehicle for transferring income from the rich to the poor. Finally, the proposed theoretical approach could enable one to break away from the need to resort to costly contingent valuation surveys that can be difficult to conduct in ways consistent with economic theory.

The study consists of the following sections: section 2 presents a brief literature review for consumer demand systems; section 3 presents the theoretical model; section 4 presents the empirical model by explaining in detail the data used in the empirical analysis, and reports the results obtained; finally section 5 concludes the chapter.

4.2 Brief Review of Consumer Demand Analysis

4.2.1 Demand Systems

One of the most greatly developed empirical applications of economic theory is the specification, estimation, interpretation and application of consumer demand systems. A consumer demand system is a set of equations that describe how a consumer or household with observable demographic or other non-income characteristics (that affect its preferences) allocates its total expenditures to any combination consumption goods given the market prices of those goods. One of the most active area of demand system approach concerns welfare and cost effects of changes in demographic characteristics or other attributes of households (except of prices and income) that affect demand. Two very popular examples are the estimation of the cost of children and the calculation of the amount of money a household would require maintaining its same standard of living if one of its members died.

Demand systems are usually used to test behavior questions of economic significance. In addition, they provide estimates of price and income elasticities and any other effects of demographic characteristics that affect demand. These effects are often of direct policy interest. Even when someone wants to study only one good, it is useful to estimate its demand within the context of a demand system because the imposition of integrability constraints can increase the efficiency of estimation. Most of the behavior implications of utility maximization include cross-equation restrictions, such as Slutsky symmetry, thus are best applied in the context of a demand systems rather than to individual demand equation.

Many empirical studies of demand systems statistically rejected implications of utility maximization due in part to restrictive functional forms. The first estimated systems of demand equations derived explicitly from consumer theory were the original linear expenditure systems -LES- (see the pioneer study by Stone (1954)), the Rotterdam model first proposed by Theil (1965) and the Translog model first introduced by Christensen, Jorgenson and Lau (1975). All of these models have been extensively estimated and have been used to test the homogeneity and symmetry restrictions of demand theory. Deaton and Muellbauer (1980b) introduced the Almost Ideal Demand System (AIDS) which had many advantages in relation to the LES, Rotterdam and Translog models. In this model the budget shares of the various commodities are linearly related to the logarithm of real total expenditure and the logarithms of relative prices. However, there has been increasing evidence of empirical Engel curve studies that further terms in income are required for some, but not for all expenditure share equations. (see, for example, Atkinson et al (1990); Hausman et al (1995); Lewbel (1991a)).

Gorman (1981) analyzed the class of exactly aggregate demand systems, and found that the integrability imposes that the rank of a matrix of Engel curve coefficients in such systems cannot exceed three[57]. For example, a demand system having all linear Engel curves is rank two, while quadratic Engel curves can be either rank two or rank three. Parametric

[57] Three is the maximum amounts of prices into cost function for U to be increasing

and nonparametric empirical tests of demand system rank include among others Lewbel (1991a), Blundell, Pashardes and Weber (1993) and Banks, Blundell and Lewbel (1997).

Based on the theoretical findings of Gorman (1981), richer functional forms demand systems have been developed adopting quadratic (rather than linear) logarithmic cost functions[58] (Pollak and Wales, 1992; Blundell, et al, 1993; Banks, et al, 1997). The most popular of these systems is the Quadratic Logarithmic Almost Ideal Demand System (QUAIDS) proposed by Banks et al (1997). The QUAIDS allows quadratic log-expenditure to enter the budget share equations through an extension of the AIDS, thereby maintaining the theoretically sound and empirically practical properties of the latter demand system.

4.2.2 Equivalence Scales

The empirical model used in this chapter of the thesis in order estimate the value of state education to the consumer in based on the concept of equivalence scale. In general, an equivalence scale is used to compare the welfare of households with different demographic characteristics, e.g. different number of adults and children in various age groups, and is defined as the ratio of two costs

$$\mathrm{C}\left(z^h; \mathrm{p}, z_0^h, \mathrm{U}\right) \equiv \frac{\mathrm{C}\left[\mathrm{z}^\mathrm{h}, \mathrm{p}, \mathrm{U}\right]}{\mathrm{C}\left[\mathrm{z}_0^\mathrm{h}, \mathrm{p}, \mathrm{U}\right]}$$

where $C[z^h, p, U]$ is the cost required the household with characteristics z^h and $C[z_0^h, p, U]$ the cost required by household with characteristics z_0^h to reach utility U at prices p. For example if the two vectors of households characteristics differ only in the number of children, e.g. z^h contains one child whereas z_0^h does not) then the equivalence scale measures the increase in cost required by the household with one child to reach the same level of utility (at given prices) as the childless household.

The limitations of using demand analysis to estimate equivalence scales are well known in the literature. Difficulties associated with consumer demand estimates of demographic costs can arise due to the use of inadequate empirical specification and/or insufficient data.

---

[58] Early systematic attempts to introduce quadratic expenditure effects in a demand system date back to Pollak and Wales (1978) who proposed the Quadratic Expenditure System (QES), a rank-3 extension of the then popular Linear Expenditure System (LES).

Muellbauer (1974) was the first who showed that equivalence scales cannot be identified when a system of budget shares linear to log expenditure is applied to cross-section data. Some studies consider demographic effects in the context of demand equations that include quadratic expenditure in both rank-2 and rank-3 systems. As Pashardes (1995) suggested, the specification of rank-3 system nests the AIDS and the only rank-2 extension of AIDS that allows nonlinear log expenditure effects in the budget share equations. He showed that while the introduction of nonlinearity in rank-2 systems can identify the equivalence scales from cross-section data, the rank-3 system yields more precise estimates of demographic costs. Finally, when price variation is also presented in the data, rank-3 system is also favourable. Previews and recent work on theoretical and empirical investigation of equivalence scales in demand systems includes studies by Pollak and Wales (1979), Deaton and Muellbauer (1986), Lewbel (1986), Blundell and Walker (1984), Lewbel (1991b), Blundell and Lewbel (1991), Pashardes, (1991,1995), Dickens, Fry and Pashardes (1993).

The most serious limitation of equivalence scales, however, concerns the Independence of Base (IB) property that they need to possess in order to be meaningful measures of welfare comparison across households. As argued by Lewbel (1989) and Blackorby and Donaldson (1989) the equivalence scale does not satisfy IB unless the cost function is multiplicatively separable, i.e. it can be written as

$$\ln C(z^h, p, v(U, z^h,)) = \ln C_1[z^h, p] + \ln C_2[p, U]$$

For example, IB does not hold when the $\ln C_2[.]$ sub-function has the form $\ln C_2[z^h, p, U]$, implying that parameters that translate z^h to cost, also translate U to cost; therefore, one cannot measure the cost of z^h at constant utility. With non-IB the equivalence scale depends on the level of utility at which it is calculated; therefore in the context of cardinal utility analysis the resulting index is meaningless.

As shown by Blundell and Lewbel (1991) some, but not all, violations of multiplicative separability of the cost function are testable. Therefore, while the IB hypothesis can be rejected if these violations are empirically observed, the validity of the IB hypothesis can never be established on empirical grounds.[59] Although this is a serious limitation in using equivalence scales to make welfare comparisons across households, one can justify this approach on the grounds that such comparisons are made at 'subsistence' consumption level (e.g. the consumption of households at the bottom 1% of the budget (total expenditure) distribution. This is reasonable, given that the equivalence scales are mostly used to differentiate the cost of demographic characteristics of households at the bottom end of income distribution.

4.3 Theoretical Model

Consumers selecting to consume state education benefit from it. But what is the value of this benefit? Clearly, this question cannot be answered from market data (i.e. price and quantity of education observed in the market) because no market price exists for the education provided by the state. Here an estimate of the value of the publicly provided education will be obtained by observing how state schooling affects the consumption pattern of households.

The starting point of our analysis is that state education is provided free of charge only at some minimum level and consumers not satisfied with this level of education can either opt out of the state education system and send their children in private schools, in which case they have to pay fees; or supplement the free of charge minimum state education with more state education paid through purchasing relatively more expensive accommodation in the catchment area of a high performing state school. Thus they locate themselves in areas where state schools can provide high quality education for their children[60]. In the latter case education is purchased jointly with housing. In the context of demand analysis this

[59] To overcome non-IB problems in making welfare comparisons across households Blundell and Lewbel (1991) propose the use of the 'relative' equivalence scale, cost of living of households with different demographic characteristics attributed to price changes (e.g. the cost of living of households with children can increase more than that of other households when food prices increase faster than the prices of other goods).

[60] See for example papers by Black (1999) for the case of the US and Gibbons and Machine (2003) for the case of UK.

complicates estimation because one cannot observe education and housing expenditure separately[61], except for households that send their children to private schools.

To circumvent the problem one can use a separability assumption allowing household consumption decisions to be taken in two stages: at the first stage total expenditure is thought to be allocated among broad commodity groups, normally between non-durables and durables; at the second stage the budget for non-durables is allocated among commodities in this group. This two-stage budgeting framework is similar to the one used by Pashardes (1991) and Blundell, Pashardes and Weber (1993).

Thus, as long as one can consider education (along with housing, motor vehicles, household appliances etc) to be in the category of goods decided at an upper budgeting stage, demand analysis focusing on the allocation of expenditure at a lower budgeting stage need not include education. The logic for doing so is that once taken the decision to either supplement the freely provided minimum state education through paying for accommodation in the catchment area of a better state school or to opt out of the state system and purchase education from the private sector, implies a long-term spending commitment. In this sense it is no different from purchasing a durable good, given that households are locked in a given level of consumption that cannot be costlessly altered in response to changes in price or income.

The present analysis considers the two budgeting stages described above to be based on implicit (or quasi) separability, i.e. prices p enter the cost function partitioned into K groups where each group k has its own sub-cost function defined on the group price vector $p_k = p_{k1,\ldots,}p_{kn}$ and total utility U[62].

Thus, in the context of implicit separability, utility is common to both budgeting stages and provides a connection between them: higher (lower) consumption costs in the second

[61] In this study a method to calculate the jointly education-and-housing expenditure using a Heckman sample selection model is suggested. Details are discussed later in.

[62] The different concepts of separability in consumer demand analysis (additive, weak, implicit etc) are discussed in depth in Blackorby et al (1978) and Blackorby and Shorrocks (1996).The most popular concept of weak separability implies that the group sub-cost functions are defined on group sub-utility. Thus if preferences are weakly separable then commodities can be divided into groups so that preferences within groups can be described independently of the quantities in other groups (see Deaton and Muellbauer, 1980). Previews studies by Blundell and Walker (1982, 1986) and Browning and Meghir (1991) modelled the joint determination of household labour supplies and commodity demands assuming restriction on the household’s preferences of weak separability between goods and leisure.

budgeting stage are transmitted to the first budgeting stage effects through lower (higher) total utility; and vice versa. This connection between the two stages of spending decisions (not found, for example, in weak separability, where not only prices but utility too is separated between budgeting stages) is exploited in the analysis below to construct a money metric of the benefit derived from consuming the free (minimum) state education from the analysis of consumer behaviour at the second budgeting stage.

Households are considered to have the preference structure (Gorman, 1976; Deaton and Muellbauer, 1980) defined by the cost function

$$C(p,U) = \left[c_1\left(p_{11,\dots\dots,}p_{1n,}U\right), \dots\dots.., c_k\left(p_{k1,\dots\dots,}p_{km,}U\right), U\right] \tag{4.1}$$

where $c_k\ (.)$ is the sub-function reflecting the prices (unit cost) of the k^{th} commodity group, $q_k = q_{k1\ ,\dots\dots,}q_{kn}$, which is increasing in U and linearly homogeneous in price. The subscripts n and m indicate the number of goods in the groups. As described above, we assume that in the first stage there are three broad categories of goods, i.e. non-durables goods, durable goods and the composite education-and-housing commodity.

Consumer demand for the i^{th} good in the k^{th} group is obtained by applying Shepherd's lemma to (1),

$$q_{ki} = \frac{\partial C\ (.)}{\partial p_{ki}} = \frac{\partial C\ (.)}{\partial c_k\ (.)}\frac{\partial c_k\ (.)}{\partial p_{ki}} \tag{4.2}$$

where $\partial c_k\ (.)/\partial p_{ki}$ is the Hicksian consumer demand for the i^{th} good in the k^{th} group conditional on the (Hicksian) demand for the k^{th} commodity group given by $\partial C\ (.)/\partial c_k\ (.)$.

Writing (4.2) in the form

$$q_{ki}\ = \quad \frac{\partial lnc\ (.)}{\partial lnc_k\ (.)}\ \frac{C(.)}{c_k\ (.)}\ \frac{\partial lnc_k\ (.)}{\partial lnp_{ki}}\ \ \frac{c_k\ (.)}{p_{ki}} \tag{4.3}$$

yields the Hicksian consumer demand for the i^{th} good in the k^{th} as share in total expenditure $X = \sum_k x_k$,

$$\omega_{ki}^* \equiv \frac{q_{ki}\ \ p_{ki}}{X} = \frac{\partial lnC\ (.)}{\partial lnc_k\ (.)}\ \frac{\partial lnc_k\ (.)}{\partial lnp_{ki}} \tag{4.4}$$

where $\partial lnc_k\ (.)/\partial lnp_{ki}$ is the share of the i^{th}good in the k^{th} group expenditure, $\omega_{ki} \equiv q_{ki}\ p_{ki}/x_k$; and $\partial lnC\ (.)/\partial lnc_k\ (.)$ the share of the k^{th} group in total expenditure, $\omega_k \equiv q_k\ (.)\ p_{ki}/X$.

The analysis is conducted at the lower stage and concerns one of the three broad categories of the first stage budgeting; that of the "non-durables" commodity group[63]. The unit cost of the k^{th} commodity group is defined in the context of an integrable rank-3 demand system, the Quadratic Logarithmic (QL)[64] cost function (Lewbel 1990; Banks et al, 1997) and for household h over goods $i = 1, \ldots\ldots, n$ in period t is written

$$lnC_k(p_{kt,}U) = a(p_{kt}) + \frac{b_h(p_{kt})\ U}{1 - l_h(p_{kt})U} \tag{4.5}$$

where $a_h(p_{kt})$, $b_h(p_{kt})$ and $l_h(p_{kt})$ are some household-specific price indices. In addition, $a_h(p_{kt})$ is homogeneous of degree one in prices, whereas $b_h(p_{kt})$ and $l(p_{kt})$ are homogeneous of degree zero. This cost function yields Hicksian shares[65]

$$\omega_{hi} = a_{hi}(p) + b_{hi}(p)\left[\frac{U}{1 - l_h(p)U}\right] + \frac{l_{hi}}{b_h(p)}\left[\frac{U}{1 - l_h(p)U}\right]^2 \tag{4.6}$$

where $a_{hi}(p) = \partial lna_h/\partial lnp_i$, $b_{hi}(p) = \partial b_h/\partial lnp_i$, and $l_{hi} = \partial l_h/\partial lnp_i$

Note that U is the household utility level obtained from the total expenditure as defined at the top stage of budgeting. Therefore to obtain the Marshallian demands for nondurable goods we substitute U in Eq.(4.5) for the indirect utility function $V(p, Y^*)$, where Y^* is the top stage budget of household from, which is approximated by a popular Stone Index (also known as the Cobb-Douglas index), $\sum_i \omega_{hi}\ lnP_i$, where the weights (ω_{hi}) are the household's expenditure shares of the goods within the group.[66]

The Marsiallian budget share equations at the lower budgeting stage are then obtained by

[63] Keep in mind that, in addition to the number of children, the amount households spend on the joint composite commodity is pre-determined at the lower stage. Without this assumption one cannot estimate child costs and, therefore, the cost difference between attending private or state school.

[64] QL is the most general functional form that allows recovery of the cost and welfare effects of changes in consumer behaviour.

[65] The time t and the subscript k that can be attached to the variables in the next of this section were omitted for notational simplicity.

[66] For empirical purposes all prices were expressed in prices of January 2001

$$\omega_{hi} = a_{hi}(p) + b_{hi}(p)[lnY_h^* - lna_h(p)] + \frac{l_{hi}}{b_h(p)}[lnY_h^* - lna_h(p)]^2 \quad (4.7)$$

where $Y_h^* = Y_h - \sum_i \omega_{hi} \ lnP_i$ is the deflated budget of household h (in period t) from the first stage budgeting.

This budget share equation can be estimated and their parameters can be readily interpretable. The analysis that follows focuses on the effect of free schooling on consumer behaviour at the second budgeting stage by extending (4.7) to include child costs and other household characteristics, as described in the discussion about equivalence scales in the previous section. The empirical validity of the assumptions behind this simplified demand system is examined in the next section.

4.4 Empirical Analysis

4.4.1 Data

The empirical analysis is based on two categories of nondurable goods: "Food and Catering" and "Others". The group "Others" includes the subcategories of fuel, clothing, transport and communication, households goods and services, personal and leisure goods and services. Details on each subcategory are given in Appendix C1.The data are drawn from the 2001-2007 Expenditure and Food Survey (EFS) of UK. The Expenditure and Food Survey (EFS) took over from the Family Expenditure Survey (FES) and the National Food Survey (NFS) and started in 2001.Each EFS includes information on the income and household expenditure on a large number of goods, including education. It also includes a large number of personal/household characteristics.

The sample used in the empirical analysis consists of two-adult (non-retired) households[67] without children or with children up to 16 years old attending private or state or both at pre-primary, primary or secondary education level. This sample selection resulted in

[67] This was motivated by the need to limit heterogeneity among households to demographic characteristics of interest, i.e. the number of children of schooling age. Extending the sample to include other household categories, for example households with more than two adults, households with hrp over 65 or households with children also in higher education level, would introduce further heterogeneity and require the inclusion of additional parameters in the demand system. Thus, it is important to point out that the empirical results in this study may not hold for types of households substantially different from those in the selection considered.

14619 observations for the seven years (2001- 2007). Specifically, 61.15% (n=8939) are households that have no children at all, while the rest 38.85% (n=5680) are households with children between ages 0-16. Around 4% of the latter households have children only in private schools (n=213) and around 2% of the sample are households with children in both private and state schools (n=123). The summary statistics of the main variables used in estimations are shown in Table 4.1. A detailed description of a large number of household characteristics is given in Appendix C2 (Table C2.1).

Table 4.1: Main Descriptive Statistics

Variable	Mean	Std.Dev.	Min	Max
Number of children	0.732	1.033	0	4
Number of children only in state schools (n=5467)	1.896	0.762	1	4
Number of children only in private schools (n=213)	1.559	0.631	1	4
Number of children both in state and private schools (n=336)	1.896	0.783	1	4
<u>Shares:</u>				
SFC	0.268	0.107	0	0.928
SOTHER	0.732	0.107	0.072	1
Log household expenditure (Upper Stage)	6.138	0.448	4.554	7.432
Log household expenditure (Lower Stage)	5.554	0.546	2.895	7.033

Furthermore, for the calculation of the composite education-housing commodity a Heckman sample selection model is used in the following way. An equation determining the sample selection is used to model whether a household rents a house (furnished or unfurnished) as a function of characteristics of the house and of the household. A model for household expenditure on actual rent has also been specified as a function of a subset of the characteristics used in the selection equation and a term correcting for the sample selection bias, the latter arising from the likely correlation between the error terms from the actual rent expenditure and the sample selection models, as explained elsewhere in this thesis.

After the estimation of a model, predictions about the rent expenditure for all households are constructed by extrapolating the rent expenditure for households that own a house from the estimated equation obtained for households that rent a house, and by multiplying it with the probability of a household renting a house. The extrapolated values for the rent expenditure form the imputed rent expenditure for owner households (as there is no actual rent expenditure for these households). Thus, this imputed rent expenditure can be viewed as the housing expenditure for all households. Finally, as already mentioned above, the housing expenditure for households with children in state schools can also include an education component.

The data used in the empirical estimation of a Heckman model include all households with and without children whose household reference person (hrp) is under retirement age. Variables on housing characteristics (total rooms, heating, region e.tc), household characteristics (number of adults, number of children, age of the household reference person, e.tc) and also expenditure on council, water and sewerage tax are included in both structural and selection equation. Income sources of the hrp were also included into the selection equation for identification purposes. The estimation results, obtained by maximum likelihood, are shown in the Appendix C2 of this chapter.

4.4.2 Valuation of Free State Schooling

The effect of the freely provided state education on consumer behaviour at the second budgeting stage is modelled, along with the effects of children and other demographic and non-demographic characteristics of the h^{th} household, denoted by z_{μ}^{h}, $\mu = 1 \ldots .$ M, by allowing the 'subsistence' parameters of the cost function to depend on these characteristics. The vector z^h can include personal and household variables found to affect the level and pattern of consumption in studies analysing individual household behaviour with pooled time-series and cross-section data (e.g. Blundell et al 1992, Pashardes 1991, 1995). In addition to children in various age groups such variables include the age, occupation, economic position and employment status of household head and spouse; housing characteristics such as location, type, size, central heating, tenure; seasonal dummies, trend and other time varying macro variables (e.g. the interest rate) etc.

An empirical rank-3 demand system is obtained by assuming $a_h(p_t)$, $b_h(p_t)$ and $l_h(p_t)$ to have the explicit forms corresponding to the Quadratic Logarithmic Almost Ideal Demand System (QUAIDS) (Banks et al. 1997)

$$a_h(p_t) = a_0(z^h) + \sum_i a_i(z^h)\, lnp_{it} + .5 \sum_i \sum_j \gamma_{ij}\, lnp_{it} lnp_{jt} \tag{4.8}$$

$$b_h(p_t) = \prod_i p_{it}^{\beta_i} \tag{4.9}$$

$$l_h(p_t) = \sum_i \lambda_i \, lnp_{it} \tag{4.10}$$

Replacing (4.8)-(4.10) in (4.7), the Marsallian form of the budget shares is given by

$$\omega_{hit} = a_{ih} + \sum_i \gamma_{ij} \, lnp_{jt} + \beta_{ih}[lnY_h^* - lnP_{ht}] + \frac{\lambda_{ih}}{b(p)}[lnY_h^* - lnP_{ht}]^2 \tag{4.11}$$

where $lnP_{ht} = a_{0h}$ and Y_h^* is the deflated budget of household h (in period t) from the first stage budgeting. In the case where $Y_h \neq Y_h^*$ then

$$lnP_{ht} = a_{0h} + \sum_i \left(a_{ih} + 0.5 \sum_i \gamma_{ij} \, lnp_{jt}\right) lnp_{it}$$

as in Banks et al (1997).

The household-specific parameters are assumed to be linear functions of household characteristics, $a_{ih} = a_i + \sum_\zeta a_{i\zeta} z_{\zeta h,}$ $\zeta = 1, \ldots, \mathrm{Z}$ and $a_{0h} = a_0 + \sum_n a_{0n} z_{nh,}$ $n = 1, \ldots, \mathrm{N},$ where N is a subset of Z that includes only demographic characteristics of the household[68]. The parameters a_{ih} reflect demographic substitution effects and the parameters a_{0h} reflect the marginal (log) cost of the n^{th} demographic characteristics at base prices. The "subsistence" (log) cost a_0 corresponds to the reference household defined by $z_{\zeta h}$, all ζ. Throughout the analysis a_0 is fixed at a level of the log of the expenditure of the poorest 1% of households in the sample, as suggested by Deaton and Muellbauer (1980). The estimation method adopted here is nonlinear SUR after the imposition of integrability conditions (i) $\sum_i a_{ih} = 1 \; all \; h,$ $\sum_i \gamma_{ij} = 0 \; all \; j, and \sum_i \beta_i = \sum_i \lambda_i$=0 for adding up, (ii) $\sum_j \gamma_{ij} = 0 \; all \; i$ for homogeneity and (iii) $\gamma_{ij} = \gamma_{ji} \; all \; i \; and \; j$ for symmetry.

Using z_0^h to denote the number of children in the family attending state schools (reference household), the benefit for the μ^{th} household can be measured by the money metric,

$$C\left(z_0^h; p, z_\mu^h, U\right) \equiv \frac{C\left[z_\mu^h, p, U^h\right]}{C\left[z_0^h, p, U^h\right]} \tag{4.12}$$

---

[68] Although linear in parameters, these functions can be nonlinear in household characteristics including square and product terms.

measuring the relative cost of a household with children in private education to reach the same utility level (at prices p) as an otherwise identical household with children in state education. As such (4.12) is a measure of the compensation a household would accept in order to give up its entitlement to free state schooling for its children and send his child into a private school. This compensation decreases with the cost of supplementing the minimum education provided free of charge by the state with out-of-pocket payments; and will obtain its minimum value (i.e. unity, indicating no benefit from state schooling) when the expenditure required to achieve a given utility level under a state-plus-supplementation regime is not higher than that required to achieve the same utility level under an all-private education regime.

Equation (4.12) resembles an "equivalence scale" – described earlier in this chapter – except for the fact that it shows the cost of children in private education, rather than the cost of children themselves. As such it is an index of welfare comparison, and thereby, subject to the usual Independent of Base (IB)[69] restriction (Lewbel, 1989; Blackorby and Donaldson, 1989) required to make such comparison meaningful - at least for utility levels above zero. In general, for a given household characteristic z^h IB holds when the cost function $C[z^h, p, U^h]$ can be written in the multiplicatively separable form $C_1(p, z^h)C_2(p, U^h)$, implying that $\partial lnC(.)/\partial U^h$ does not depend on the household characteristic in question. In the context of the QUAIDS model, for (4.12) to be IB, the (4.9) and (4.10) functions must be independent of whether children attend a private or state school. This is further discussed in the the results' section below.

4.4.3 Results

The results obtained from SUR estimation given by empirical specification (4.11) are presented below. Table 4.2 reports selected parameter estimates that are of interest to the issues raised in this chapter, together with the corresponding t-statistics. The remaining parameter estimates, which show the effect of other household characteristics included in the budget share equations through the a_{ih} function, are reported in Appendix C2. Table 4.2 also reports the estimated linear (β_i) and quadratic (λ_i) log expenditure parameters for

---

[69] The IB rule said that any monotonic transformation of utility must be independent of the household characteristics.

both “Food and Catering” and “Other” nondurable goods, the number of observations used, as well as the value of the objective function for the systems. The expenditure effects of household h vary with number of children and number of children in private schools.

Model I corresponds to the case where both child cost and the cost of child in a private school are estimated. Model II reports the same results obtained separately for households with children in pre-primary and primary education (ages 0-10) and children in secondary education (ages 11-16). The results of Model I show that each child costs 16.5% of the total expenditure of a household, while this cost increases to 24 % if the child attends private rather than state school. Thus, a household with a child attending a state school has a benefit equal to 7.5% of its total expenditure compared to an otherwise similar household with a child in private education. Stating the results using the concept of adult equivalence scale Model I suggests that relative to a couple without children, a couple with one child has an adult equivalence scale of 1.18[70], whereas in the case where the child attends a private school the scale increases to 1.27.

Table 4.2: Selected estimates and system statistics

	Model I		Model II	
	Coef.	t-ratio	Coef.	t-ratio
child cost	0.165	11.04		
child cost ages 0-10			0.112	6.89
child cost ages 11-16			0.235	13.5
child cost in private school	0.076	2.17		
child cost in private school ages 0-10			0.082	2.04
child cost in private school ages 11-16			-0.006	-0.13
β_i				
Food and Catering	-0.123	-21.81	-0.126	-22.29
Other nondurable	0.123	21.81	0.126	22.29
λ_i				
Food and Catering	0.012	4.06	0.013	4.49
Other nondurable	-0.012	-4.06	-0.013	-4.49
Number of Observations (N)	14619		14619	
Objective*N	14582		14581	

The results obtained separately for children in different age groups suggest that the child cost of younger children (ages 0-10) is lower than the cost of older children (ages 11-16), with the former being about half the size of the latter one. The benefit from attending state rather than private school, however, is observed only in the case of households with children in the first group, i.e. in pre-primary and primary education; while for children in

[70] These results conform to findings elsewhere in the literature (e.g. Pashardes, 1991, 1995).

secondary schools it appears that it makes no difference to the household cost whether they attend state or private school. More precisely, households with children in pre-primary and primary education enjoy a benefit from free state schooling equivalent to 8% of their budget; whereas the benefit of free state schooling for households with children in secondary education is zero.

Taken at face value, one explanation for the results above is that a household with a child in pre-primary and primary education does not pay as much to locate itself in the catchment area of a high quality state school as it would to send its child to an equal quality private school. Therefore, households with young children benefit from state schooling. In contrast, for a household with a child in secondary education this benefit is fully eroded because the extra housing cost required to access a high quality state school is the same as the fee of an equal quality private school.

Hence, buying quality in England's state schools through the housing market appears to represent a cheap alternative to private schooling but only for households with young children. Benefits from state schooling at primary level are also found by other researchers (i.e Gibbons and Machin, 2008). The fact that no benefits are associated with state schooling at secondary level obtained here needs to be further explored; however, this is beyond the scope of this thesis.

As regards the statistical fitting of the models estimated, the linear and quadratic log expenditure effects are both significant at 0.01 levels. "Food and Catering" appears to be necessity good (negative sign of β_i and positive λ_i) and "Other" nondurable goods to be luxury (positive sign of β_i and negative λ_i). Both β_i and λ_i must add to zero[71]. Diagnostic statistics are reported in Table 4.3. In the empirical analysis of this model there are only two share equations with the adding-up and homogeneity restrictions to be the following: i) adding-up: $\gamma 11 + \gamma 21 = 0$ and $\gamma 12 + \gamma 22 = 0$; ii) homogeneity: $\gamma 12 + \gamma 11 = 0$ and $\gamma 21 + \gamma 22 = 0$. From (i) and (ii) we have that $\gamma 11 = -\gamma 21$ and $\gamma 12 = -\gamma 11$, thus $\gamma 12 = \gamma 21$. The last restriction ($\gamma 12 = \gamma 21$) is the symmetry restriction. Consequently, one can test only for one of two restrictions; the symmetry or the homogeneity one. Homogeneity was rejected in both models at 0.05 levels. The failure of

[71] In both models the effects of child cost and child cost in private schools were insignificant when the education component were excluded from the total expenditure.

homogeneity is not unusual in static demand analysis (see, for example, Deaton and Muellbauer,1980; Blundell, Pashardes and Weber,1993).

The IB hypothesis was tested in each model by allowing the $\beta's$ and $\lambda's$ to vary between households with and without children as well as with children in private vs state schools. As seen from the last line in Table 4.3 this hypothesis was rejected in both models at 0.10 levels. This is consistent with evidence reported elsewhere (Pashardes, 1995; Dickens et al, 1993) and, as said earlier, it implies that money metrics of consumer welfare constructed from empirical demand analysis - including equivalence scales and the benefit from state schooling - can be valid only at base utility level (a_0 =log of the expenditure of the poorest 1% of households in the sample).

Table 4.3: Diagnostic Tests

	Model I		Model II	
Symmetry and Homogeneity (p-value)		0.046		0.0143
<u>Independence of Base (IB):</u>				
LM test {p-value}	8.87	0.0643	14.79	0.0634

The separability assumption of theoretical model can be tested empirically using Eq. (4.11) by defining $\delta_i = \sum_{s=1}^{2} \delta_{is}\bar{q}_s \;\; i = 1,2$, where $\bar{q}_s$, $s = 1,2$, are the quantities of durable goods and composite commodity education-and-housing which has been determined, and testing whether δ_{is} are zero for all i and s. Rejection of the null hypothesis would indicate that the nondurable goods in the lower stage cannot be separated from other commodity groups of the upper stage. The assumption of separability was rejected (null hypothesis p=0.000) in both models indicating that the nondurable goods cannot be separated from the other two groups in the upper stage.

4.5 Conclusion

Although today there are private markets for the provision of private goods such as education, large amounts of public funds are still being invested in this sector. A significant reason for the free provision of the publicly provided private goods by the state is the guarantee of minimum consumption by economically deprived social groups. However, depending on the sector, the design of the public provision scheme can allow for the eligible individuals to supplement their consumption with purchases from the private sector.

Taking into consideration that state education is provided free of charge only at some minimum level, consumers (for example in UK) who are not satisfied with this level of education can opt out of the state education system and send their children in private schools, in which case they pay fees; or supplement the free of charge minimum state schools, in which case they pay fees; or supplement the free of charge minimum state education with more state education paid through purchasing relatively more expensive accommodation in the catchment area of a high performing state school.

In this study the demand for education was modelled in the context of an integrable complete demand system, satisfying the fundamental principles of economic theory. The integrability of the demand system is necessary for the derivation of a money metric of welfare associated with free public provision. The derivation of a money measure that has been recovered from the parameter estimates of a demand system was used to measure the welfare (utility) which a household obtains from the consumption of publicly (free of charge) provided goods.

The theoretical model was based on the notion of separability and the two stage budgeting framework. The empirical analysis was applied to data drawn from the 2001-2007 Expenditure and Food Survey (EFS) of UK. The conclusion emerging from the empirical analysis is that households with children in pre-primary and primary education enjoy a large benefit from free state schooling something that is not true for households with children in secondary education. Particularly, on average a child who attends in a state school benefits the household around 7,5% of their total expenditure compared to a family having a child in a private sector.

The approach proposed in this study can be used for the theoretical and empirical valuation of freely provided private goods other than education, such as health care, income in kind and specialised services. Furthermore, as the data for the econometric analysis can be drawn from a household expenditure survey, the proposed methodology can be applied in different countries for reasons of comparison and/or for examining issues pertaining to the economic valuation of publicly provided private goods from a wider international perspective (e.g. in the context of the European Union). Finally, it can be used to break away from the need to resort to costly contingent valuation surveys that can be difficult to conduct in ways consistent with economic theory.

In terms of policy implications, the analysis in this study could contribute to the current debate about the reform of the education systems relating to the famous literature of private school vouchers programs. It could also guide policies towards a more informed and constructive treatment of the role which can be played by the private education sector.

Appendix C

C1. Total expenditure categories

The theoretical model proposes in the upper stage includes the three following broad categories of total expenditure:

- Composite Commodity Education-and-Housing (EDUHS)
- Nondurable Goods (NDUR)
- Durable Goods (DUR)

The broad category of Nondurables goods is used in the lower stage for the empirical analysis and included the following sub-categories:

- Food and Catering (FC)
- Fuel and Light (FL)
- Clothing and Footwear (CF)
- Transport and Communication (TRCOM)
- Household Goods and Services (HGS)
- Personal and Leisure Goods and Services (PLGS)

For the categories of FC, FL and CF the exact distinction of RPI[72] is followed. For the other three sub-categories all durable groups of items are excluded and a new category created (named "Durable goods") that is used in the first stage budgeting as a different broad category.

Specifically the sub-categories of TRCOM, HGS and PLGS as also the category of Durable Goods (DUR) contain the following:

[72] The Retail Prices Index (RPI) is the most familiar general purpose domestic measure of inflation in the United Kingdom. It is available continuously since June 1947. The RPI measures the average change in prices by calculating the change in price of a fixed basket of goods and services representing the items bought by all UK households. The RPI excludes households with the top 4 per cent of income and excludes around 20 per cent of pensioner households – those that derive at least three-quarters of their income from state pension or benefits (These households are excluded because they are considered to have 'atypical' spending patterns and therefore their inclusion would distort the overall average price movement of the RPI). All other pensioner households are included. Finally RPI uses the EFS to calculate the annual weights (http://www.statistics.gov.uk).

- Transport and Communication (TRCOM)

  - maintenance of motor vehicles
  - petrol and oil
  - vehicle tax and insurance
  - fares and other travel costs such as rail fares
  - bus and coach fares, and other travel costs

- Households Goods and Services (HGS)

  - household consumables
  - pet care
  - Postage
  - telephones- telemessages
  - domestic services
  - fees and subscription[73]

- Personal and Leisure Goods and Services (PLGS)

  - personal articles
  - chemists goods
  - personal services
  - audio-visual equipment
  - CDs and tape
  - Toys
  - photography and sports goods
  - books and newspaper
  - gardening product
  - television licence and rentals
  - entertainment and other recreation
  - foreign holidays
  - UK holidays

- Durable Goods (DUR)

  - furniture
  - furnishings
  - electrical appliances
  - other household equipment
  - purchase of motor vehicles

[73] Excluded fees on education

C1.1 Prices

For the sub-categories of Food and Catering, Fuel and Light and Clothing and Footwear the available monthly price indices are used, as provided by the Office of National Statistics. For the other three sub-categories monthly price indices are constructed using the available annual weights upon the specific groups' items in each category.

Annual weights are typically used in the construction of RPI because households spend more on some goods and services than on other. Thus, price-increases for certain items have a bigger effect on the overall change in the cost of the 'basket' than others. The weights for RPI reflect these varying degrees of importance. For example, on average, households spend as much on fruit as they do on postal services, so a five per cent increase in the price of fruit would have ten times as much effect on the total cost of the 'basket' when compared to a five per cent increase in postal charges. Each item in the index is weighted to reflect the proportion of household expenditure spent on the item.

Below we provide an example for the sub-category of "Transport and Communication" (TRCOM) and explain how the construction of new price indices had been developed. Prices and annual weights were available for the: (a) maintenance of motor vehicles-(mmv), (b) petrol and oil -(po),(c) vehicle tax and insurance-(vti) and (d) fares and other travel costs- (ftc).

Step 1: calculation of new weights (e.g for 2001)

- wmmv01= 21/(21+41+21+23)
- wpo01= 41/(21+41+21+23)
- wvti01= 21/(21+41+21+23)
- wftc01= 23/(21+41+21+23)

The above are the new calculated weights of the four groups of "TRCOM" for the year 2001. The calculation of the price of broad group "Transport and Communication" for the month January of 2001 was as follows:

Step 2: calculation of the new price index

Price TRCOM = (wmmv01*216) + (wpo01*225.3) + (wvti01*264.4) + (wftc01*188), where 216, 225.3, 264. 4 and 188 are the prices of the January 2001 provided by the Office of National Statistics for the four groups commodities.

This procedure was repeated for each month and year respectively for TRCOM, PHGS, PPLGS, PDUR, EDUHS[74] and OTHER. The group “OTHER” includes all the sub-categories of Nondurable goods apart from “Food and Catering” (FC).

The table below reports the prices for the FC, FL, CL, TRCOM, HGS, PLGS, POTHER, EDUHS and DUR.

Table C1.1: Prices

Variable	Mean	Std. Dev.	Min	Max
2001				
PFC	162.131	1.846	158.800	164.700
PFL	124.896	1.036	123.100	126.100
PCL	107.484	2.140	102.500	110.200
PTRCOM	223.337	2.430	218.505	227.846
PHGS	159.003	1.702	156.956	161.745
PPLGS	177.654	2.080	173.840	180.274
POTHER	173.243	1.116	171.123	174.563
PEDUHS	230.523	2.765	225.805	234.151
PDUR	129.856	1.208	127.051	131.465
N=2259				
2002				
PFC	164.857	0.306	164.300	165.300
PFL	128.763	0.705	127.700	129.900
PCL	102.402	1.936	97.600	104.200
PTRCOM	226.759	2.510	222.024	229.090
PHGS	164.600	1.957	162.409	167.652
PPLGS	187.201	2.685	182.831	190.743
POTHER	177.602	1.983	174.258	180.186
PEDUHS	219.512	1.237	217.415	220.685
PDUR	129.143	1.041	127.507	130.907
N=2056				
2003				
PFC	167.859	1.331	165.300	169.900
PFL	131.413	1.060	130.000	133.400
PCL	100.827	1.464	97.500	102.300
PTRCOM	238.530	2.321	233.190	240.983
PHGS	170.339	1.441	168.718	172.583
PPLGS	191.817	0.314	191.269	192.272
POTHER	184.082	0.936	181.894	185.178
PEDUHS	222.086	1.463	219.732	223.962
PDUR	128.820	0.902	127.466	130.238
N=2199				

	Mean	Std. Dev.	Min
2005			
PFC	172.942	0.676	171.600
PFL	159.724	5.868	152.800
PCL	95.703	1.123	92.700
PTRCOM	259.292	6.915	248.971
PHGS	177.992	1.318	176.178
PPLGS	196.454	0.734	195.034
POTHER	192.956	2.440	189.051
PEDUHS	233.121	1.897	229.721
PDUR	126.101	1.314	124.407
N=2037			
2006			
PFC	176.838	2.151	174.100
PFL	198.602	15.846	173.300
PCL	94.549	1.512	91.500
PTRCOM	269.996	5.269	263.552
PHGS	184.226	2.552	181.224
PPLGS	226.583	25.886	199.602
POTHER	210.891	10.990	197.551
PEDUHS	238.583	2.543	234.607
PDUR	127.524	1.684	125.707
N=2033			
2007			
PFC	184.439	2.951	180.000
PFL	212.875	6.906	206.100
PCL	93.751	1.283	91.300
PTRCOM	281.378	7.785	268.296
PHGS	188.246	2.184	185.983
PPLGS	208.323	1.559	205.903
POTHER	210.090	2.165	206.470
PEDUHS	252.128	3.124	247.068
PDUR	125.492	2.175	122.690
N=1919			

[74] For the combined variable (EDUHS) available prices and weights of “fees on education” were used as an index of education, while “rent” as an index of housing.

2004				
PFC	169.967	0.581	169.100	171.200
PFL	140.617	4.752	134.000	150.400
PCL	98.015	1.207	94.800	99.000
PTRCOM	248.614	3.971	242.397	254.670
PHGS	174.347	0.686	173.005	175.304
PPLGS	194.050	1.539	191.629	196.208
POTHER	188.598	1.908	185.456	191.683
PEDUHS	227.010	1.567	224.665	229.530
PDUR	128.316	1.823	125.514	130.491
N=2116				

C2. Tables

Table C2.1: Descriptive Statistics for all variables

Variable	Mean	Std. Dev.	Min	Max
<u>Shares:</u>				
SFC	0.268	0.107	0	0.928
SOTHER	0.732	0.107	0.072	1
Log household expenditure (Upper Stage)	6.138	0.448	4.554	7.432
Log household expenditure (Lower Stage)	5.554	0.546	2.895	7.033
<u>Survey years:</u>				
2001	0.155	0.361	0	1
2002	0.141	0.348	0	1
2003	0.150	0.357	0	1
2004	0.145	0.352	0	1
2005	0.139	0.346	0	1
2006	0.139	0.346	0	1
2007	0.131	0.338	0	1
<u>Quarters:</u>				
quarter1	0.243	0.429	0	1
quarter2	0.246	0.431	0	1
quarter3	0.255	0.436	0	1
quarter4	0.255	0.436	0	1
<u>Household characteristics:</u>				
Gas heating	0.766	0.423	0	1
Oil heating	0.111	0.314	0	1
Elecricity heating	0.049	0.216	0	1
Detached house	0.298	0.457	0	1
Semi-detached house	0.340	0.474	0	1
Terraced house	0.251	0.434	0	1
Total rooms>5	0.578	0.494	0	1
Number of vehicles	1.478	0.752	0	8
<u>Regions:</u>				
Mesyside	0.018	0.135	0	1
Yorkshire and Humberside	0.083	0.276	0	1
North East	0.039	0.193	0	1
North West	0.089	0.285	0	1
East Midlands	0.078	0.269	0	1
West Midlands	0.084	0.277	0	1
East Anglia	0.094	0.291	0	1
London	0.075	0.263	0	1
South East	0.143	0.350	0	1
South West				
Wales	0.050	0.218	0	1
Scotland	0.085	0.279	0	1
Northern Ireland	0.070	0.255	0	1
<u>Household reference person characteristics:</u>				
Married	0.796	0.403	0	1
White colour	0.666	0.472	0	1
Male	0.787	0.409	0	1
Age	45.028	11.578	18	65
Unemployed	0.019	0.135	0	1
Owner of the house	0.214	0.410	0	1
Income source: wages	0.751	0.432	0	1
Income source:Self employed	0.097	0.296	0	1

Table C2.2 : QUAIDS parameter estimates corresponding to household characteristics (for Model I)

Characteristics	Food and Catering		Other nondurable goods	
	Coef.	t-ratio	Coef.	t-ratio
Household with children	0.001	0.19	-0.001	-0.19
Number of vehicles	-0.014	-11.44	0.014	11.44
House with more than 5 rooms	0.003	1.44	-0.003	-1.44
Age of hrp	0.344	3.96	-0.344	-3.96
Age of hrp squared	-0.044	-3.69	0.044	3.69
Male (hrp)	-0.001	-0.65	0.001	0.65
Married (hrp)	-0.002	-0.76	0.002	0.76
White (hrp)	0.004	2.18	-0.004	-2.18
Unemployed (hrp)	0.006	1.1	-0.006	-1.1
Income source: wages	-0.008	-3.24	0.008	3.24
Income source: self employed	0.002	0.54	-0.002	-0.54
Gas heating	-0.002	-0.5	0.002	0.5
Oil heating	-0.004	-1.06	0.004	1.06
Elecricity heating	-0.012	-2.7	0.012	2.7
House:owned outright	-0.001	-0.31	0.001	0.31
Detached house	-0.005	-1.56	0.005	1.56
Semi-detached house	-0.010	-3.27	0.010	3.27
Terraced house	-0.009	-3.16	0.009	3.16
Mesyside	-0.005	-0.82	0.005	0.82
Yorkshire and Humberside	-0.010	-2.72	0.010	2.72
North East	-0.012	-2.51	0.012	2.51
North West	-0.008	-2.09	0.008	2.09
East Midlands	-0.003	-0.85	0.003	0.85
West Midlands	-0.004	-0.99	0.004	0.99
East Anglia	0.003	0.77	-0.003	-0.77
London	0.043	10.59	-0.043	-10.59
South East	0.008	2.37	-0.008	-2.37
Wales	-0.006	-1.37	0.006	1.37
Scotland	-0.001	-0.26	0.001	0.26
Northern Ireland	0.013	2.79	-0.013	-2.79
Second Quarter	-0.008	-3.76	0.008	3.76
Third Quarter	-0.004	-1.61	0.004	1.61
Fourth Quarter	0.003	1.55	-0.003	-1.55
$\gamma 11$ and $\gamma 22$	0.090	4.29	0.090	4.30
$\gamma 12$ and $\gamma 21$	-0.090	-4.29	-0.090	-4.30

Table C2.3 : QUAIDS parameter estimates corresponding to household characteristics (for Model II)

Characteristics	Food and Catering		Other nondurable goods	
	Coef.	t-ratio	Coef.	t-ratio
Household with children	0.001	0.23	-0.001	-0.23
Number of vehicles	-0.014	-11.4	0.014	11.4
House with more than 5 rooms	0.003	1.58	-0.003	-1.58
Age of hrp	0.325	3.74	-0.325	-3.74
Age of hrp squared	-0.042	-3.53	0.042	3.53
Male (hrp)	-0.002	-0.75	0.002	0.75
Married (hrp)	-0.001	-0.34	0.001	0.34
White (hrp)	0.004	2.17	-0.004	-2.17
Unemployed (hrp)	0.007	1.24	-0.007	-1.24
Income source: wages	-0.008	-3.38	0.008	3.38
Income source: self employed	0.002	0.54	-0.002	-0.54
Gas heating	-0.001	-0.42	0.001	0.42
Oil heating	-0.004	-1.01	0.004	1.01
Elecricity heating	-0.012	-2.67	0.012	2.67
House:owned outright	0.000	-0.1	0.000	0.1
Detached house	-0.005	-1.43	0.005	1.43
Semi-detached house	-0.009	-3.12	0.009	3.12
Terraced house	-0.009	-3.09	0.009	3.09
Mesyside	-0.006	-0.93	0.006	0.93
Yorkshire and Humberside	-0.010	-2.69	0.010	2.69
North East	-0.013	-2.64	0.013	2.64
North West	-0.008	-2.09	0.008	2.09
East Midlands	-0.004	-0.93	0.004	0.93
West Midlands	-0.004	-1.06	0.004	1.06
East Anglia	0.003	0.74	-0.003	-0.74
London	0.043	10.67	-0.043	-10.67
South East	0.008	2.41	-0.008	-2.41
Wales	-0.007	-1.51	0.007	1.51
Scotland	-0.001	-0.31	0.001	0.31
Northern Ireland	0.012	2.67	-0.012	-2.67
Second Quarter	-0.008	-3.77	0.008	3.77
Third Quarter	-0.003	-1.55	0.003	1.55
Fourth Quarter	0.003	1.57	-0.003	-1.57
$\gamma 11$ and $\gamma 22$	0.086	4.10	0.086	4.10
$\gamma 12$ and $\gamma 21$	-0.086	-4.10	-0.086	-4.10

Table C2.4: Estimation results of the sample selection model[a]: Composite commodity

	Rent Expenditure		Selection Equation[d]	
	Coef[b].	s.e	Coef[b].	s.e
Constant	4.394***	(0.092)	-0.235	(0.152)
Log total household Expenditure	0.096***	(0.010)	-0.083***	(0.019)
<u>Region (London)[c]:</u>				
Yorkshire and Humberside	-0.547***	(0.028)	-0.147***	(0.049)
Mesyside	-0.460***	(0.046)	-0.168**	(0.079)
North East	-0.636***	(0.040)	-0.396***	(0.065)
North West	-0.531***	(0.027)	-0.137***	(0.048)
East Midlands	-0.568***	(0.029)	-0.045	(0.051)
West Midlands	-0.408***	(0.029)	-0.189***	(0.050)
East Anglia	-0.345***	(0.026)	0.022	(0.046)
South East	-0.240***	(0.022)	0.047	(0.041)
South West	-0.407***	(0.025)	0.140***	(0.046)
Wales	-0.514***	(0.033)	-0.139**	(0.059)
Scotland	-0.467***	(0.044)	-0.495***	(0.085)
Northern Ireland	-0.256***	(0.044)	-0.400***	(0.081)
<u>Other Characteristics</u>				
<u>Total rooms (more than five)[c]:</u>				
House with 1 rooms	-0.323***	(0.066)	1.174***	(0.199)
House with 2 rooms	-0.261***	(0.037)	0.598***	(0.079)
House with 3 rooms	-0.104***	(0.026)	0.286***	(0.048)
House with 4 rooms	-0.031	(0.019)	0.188***	(0.033)
House with 5 rooms	-0.029*	(0.017)	0.090***	(0.028)
Number of economically active persons	-0.030***	(0.009)	0.000	(0.020)
Number of adults	0.102***	(0.011)	0.182***	(0.020)
Number of children	0.010	(0.007)	-0.081***	(0.011)
Council tax	0.018***	(0.001)	-0.001	(0.003)
Council water tax	-0.010	(0.019)	-0.005	(0.032)
Number of vehicles	0.010	(0.010)	-0.178***	(0.016)
<u>Heating type (other)[c]:</u>				
Electricity	0.069***	(0.023)	-0.033	(0.045)
Gas	0.150***	(0.019)	-0.261***	(0.034)
Oil	0.026	(0.032)	0.185***	(0.053)
<u>House Type (other)[c]:</u>				
Detached	-0.014	(0.029)	-0.298***	(0.046)
Semi-detached	-0.003	(0.022)	-0.335***	(0.035)
Terraced	-0.024	(0.017)	-0.170***	(0.032)
<u>Durables in the house:</u>				
freezer	-0.033	(0.022)	0.316***	(0.046)
microwave	-0.000	(0.017)	0.141***	(0.033)
dishwater	-0.107***	(0.018)	0.271***	(0.028)
<u>Source of Income (wages)[c]:</u>				
Investment	0.136*	(0.075)	0.023	(0.119)
Social security benefits	0.077***	(0.018)	0.162***	(0.034)
Other	-	-	0.884***	(0.076)
Self-employment	-	-	0.165***	(0.039)

Annuities	-	-	-0.039	(0.086)
Age of hrp	-	-	-0.035***	(0.001)
Survey Year (2001)[c]:				
2002	-0.009	(0.023)	0.002	(0.039)
2003	-0.045**	(0.022)	0.161***	(0.038)
2004	0.006	(0.023)	0.197***	(0.039)
2005	0.073***	(0.023)	0.246***	(0.039)
2006	0.074***	(0.023)	0.285***	(0.040)
2007	0.104***	(0.023)	0.333***	(0.040)

Notes: [a] The number of observations is 3786 for the rent expenditure regression (number of households that pay rent) and 35231 for the selection equation. The estimated standard error of the rent expenditure equation is 0.371. The estimated correlation between the errors of the rent expenditure and selection equations is -0.376 (s.e.=0.049) and the LR test for the independence of the two equations (ρ=0) gives a p-value equal to 0.000 (chi-squared statistic=53.26); [b]The symbols *, ** and *** denote statistical significance at 10%, 5% and 1%; [c] The variable in the brackets is excluded from the regression and is used as the benchmark for comparison.

Chapter 5

General Discussion and Contribution

5.1 Introduction

This thesis mainly focuses on the question regarding the effect of education on house prices in the UK. The estimation of the capitalization of local state school quality to house prices has been an object of a large body of literature based on hedonic analysis, especially in US (Brasington, 2000, 2002; Haurin and Brasington, 2005, 2006; 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; Cheshire and Sheppard, 2004; Rosenthal, 2003; Leech and Campos, 2003); this is probably because data about the location of individual households in the UK are not available at a sufficiently disaggregate level due to confidentiality.

The first chapter of the thesis follows the hedonic approach to estimating the capitalization of local state school quality to house prices (in the UK) and examines how the empirical results obtained from this approach can be affected by the measure used for capturing school quality. For this purpose the analysis in the thesis adopts the recently proposed Contextual Value-Added (CVA) indicator and examines how different components of this indicator are associated with house prices. Measures used for school quality in the literature vary from expenditure per pupil (e.g. Downes and Zabel, 2002), pupil/teacher ratio (e.g. Brasington, 1999), reading scores, value added measures scores (e.g. Hayes and Taylor, 1996; Downes and Zabel; 2002, Haurin and Brasington, 2006) and proficiency tests/final achievement (e.g Black, 1999).

The next two chapters of the thesis propose alternative (to hedonic) approaches to investigate the capitalization of local state school quality to house prices based on consumer demand analysis. The motivation behind these approaches is the fact that the observed increase in the price of houses that are located in the catchment area of high quality state schools expresses the household’s willingness to pay for the education of its children beyond the 'minimum' level provided free by the state. Using demand analysis one can recover the cost of education operating through housing. Chapter 3 of the thesis uses the demand analysis approach in order to estimate the shadow price of education relative to housing, as one tries to do with hedonic analysis; whereas Chapter 4 goes a step beyond the traditional hedonic analysis and tries to estimate the value consumers attach to the free of charge state schooling. The consumer demand approach draws into the analysis of the option for a household to opt out of the state education system by selecting private education (e.g. Buttin, 1998; Lankford, Lee and Wyckoff, 1995) as an alternative to locating itself in the catchment area of a high performing state school (e.g. Black, 1999; Gibbons and Machin, 2003).

5.2 Score vs Non Score Components

The literature on hedonic analysis of house prices is characterized by the use of a wide range of indices to measure school quality. Rosen and Fullerton (1977) were the first to argue that test scores are a 'better' measure of school quality, while subsequent research generally used them in studies of capitalization through housing. Over time many investigators have adopted various indices reflecting final achievement to measure school quality in house price equations. This trend has been aided by the finding that school inputs have little or no impact on student outcomes (see Hanushek, 1986). Yet, some studies still find school expenditure to be consistently capitalized into house prices (Brasington, 1999).

Educational economists have long been arguing that non-school factors can also contribute to student achievement (e.g. student characteristics, family, socioeconomic background, prior achievement, community characteristics, teaching quality or resources of school) and student achievements during a particular schooling period, i.e. the value added, can serve as a better measure of school quality itself (Mayer, 1997; Hanushek, 1992). This argument soon found its way to house price regressions, with some authors asking whether value added or final test score should be used as measures of school quality in hedonic analysis of house prices. However, available studies so far have been restricted to the case of US (Hayes and Taylor, 1996; Haurin and Brasington 2006, 2009). To my knowledge, Gibbons et al (2009) is the only one study examining the influence of simple value-added index on house prices in the UK.

In this thesis the impact of school quality on house prices is investigated using the Contextual Value-Added indicator (CVA) available for primary and secondary schools in England. This is the first time this indicator is studied, as it has only recently been introduced in order to account for student, family and socioeconomic characteristics affecting the progress made by pupils, unlike the simple Value-Added index of school quality which is used by Gibbons et al (2009). In this study, we use a unique database that covers regions throughout the whole of England and is constructed (during 2008) by combining three independent sources, solely for the purposes of this research.

The emphasis in the analysis is on: (i) finding an appropriate empirical specification to capture the aggregate effect of CVA and the effect of its score (school factor) and non-score (non-school factor) components on house prices; and (ii) highlighting the potential heterogeneity in these effects, between different levels of education and between neighbourhoods and spatially larger entities, like Local Authorities (LAs). The results confirm the strong positive effect of final score measures on house prices found elsewhere in the literature. In addition, the effect of CVA vary between different educational levels; positive and significant for primary schools and negative and significant for secondary schools, an empirical finding also reported by Haurin and Brasington (2006).

Furthermore, our results highlight the complexities of using broad measures of school quality, such as the CVA, in hedonic price regressions. Specifically, we find that the non-score component of CVA plays no role at the primary level of education, probably a reflection of the limited variation of prior achievement (the most important item in the non-score part of CVA) between pupils at this education level. In this case the CVA may not contain much more information than the final score itself. In contrast, the non-score part of CVA plays a significant role in the analysis of secondary schools, probably because prior achievement can vary substantially among pupils entering this level of education. Apparently, reaching a certain final score starting from a high initial level subtracts from the school's image of quality and, thereby, the prices of houses in its catchment area.

In terms of policy implications, the analysis of Chapter 2 suggests that the recently adopted practice of using CVA as a measure of school quality in England can encourage LAs to pay more attention to raising the non-score quality characteristics of their schools. This policy appears to conform to household preferences, expressed by their willingness to pay more (less) for houses in the catchment area of primary (secondary) schools in LAs with higher CVA average.

5.3 Estimating Capitalization from Demand Analysis

Applying hedonic analysis to estimate the capitalisation of local state school quality to house prices is difficult in countries like the UK, where family expenditure surveys do not reveal information about the location of households so as to enable one to associate house prices with the quality of the local state school. Chapter 3 of this thesis proposes a method based on the theory of consumer demand analysis (Deaton and Muellbauer, 1980) to circumvent this problem. This approach is motivated by the argument put forward in the context of the widely researched issue of state vs private school selection, that communities with a high proportion of students enrolled in private schools can be less willing to pay taxes to finance state schools (e.g. Goldhaber, 1999). As a result, families may locate in areas where local authority spending on education and property tax rates (hence, housing costs) are high enough to match their educational desires; or select a location outside these areas and pay out-of-pocket for the education of their children by enrolling them to private schools.

The proposed theoretical model in this thesis is based on the notion of separability and the two-stage budgeting framework (Deaton and Muellbauer, 1980); and using Translog and Quadratic Logarithmic functional forms a consumer demand system is derived allowing for the estimation of the shadow price of education relative to housing. The empirical analysis relies on the fact that households with children in private education report their education and housing expenditures separately, while those with children in state education report only housing expenditures that may also embody education costs in the form of higher house price. The unobserved education component of the joint composite education-and-housing expenditure is approximated by Heckman-type methods (Heckman, 1979). The model is applied to data drawn from the UK Family Expenditure Surveys (FES) for the years 1994 -1997 where information about the purchase value of the property of the participating households is available.The results show a positive and significant education component in the observed education-and-housing expenditure with the relative price of education to housing to vary between 0.4 to 2, largely depending on its conditioning on other characteristics of the household and the quantities of the nondurable goods included in the household budget (total expenditure).

In addition, a usual hedonic analysis is performed based on both parametric and semi-parametric methods, where the notional education expenditure estimated by Heckman type methods – a willingness to pay thought to be associated with high school quality - is included among the explanatory variables. The results suggest a non-linear relationship, with middle-income households being more willing to pay for the education of their children through housing (i.e. having a higher elasticity of the notional education expenditure with respect to education-and-housing). Thus, Chapter 3 of this thesis shows that using widely available data in Household Expenditure Surveys in the context of demand analysis one can estimate the shadow price of education relative to housing, but also compute a measure of school quality that can be used in hedonic analysis to estimate the capitalization of school quality into house prices.

5.4 What is the Value of Free State Schooling?

During the past decades there has been a strong debate about the most appropriate method of mass educational provision. Central to this debate has been the ability of the state to provide «quality» education at any given time and place. Governments intervene in education by regulating its content (prepare curriculum and testing), its demand (through laws concerning compulsory attendance) and through the determination of funding and provision. Funding can be either direct through free state schooling or indirect through the use of education vouchers (e.g.Ladd, 1992; Epple and Romaro 1998; Neshyba, 1999, 2000). At the heart of the debate about state provision of education is the efficiency-equity trade-off (e.g Epple and Romano, 1996a, 1996b; Blomquist and Christiansen, 1995, 1999).

Existing empirical literature on public provision of education is mainly conducted at an abstract level, not helping researchers to understand the value households attach to this provision and how it can be estimated from widely propagated statistical data and econometric methods. Apart from the implementation of hedonic analysis methods and contingent valuation surveys (e.g. Brookshire and Coursey,1987) to the best of our knowledge there are no consumer valuations studies of the public provision of education in the literature based on economic theory tools. The approach proposed in Chapter 4 aspires to fill this gap by providing economists with a new tool that can be used for this purpose. Its contribution lies on modelling the economic valuation of freely provided state education in the context of a complete demand system satisfying the fundamental principles of economic theory.

The theoretical model is analysed in the context of consumer theory, using the notions of implicit separability and two-stage budgeting (Deaton and Muellbauer,1980), allowing one to derive and compute (from the parameters of an integrable complete demand system from a known utility/cost function (Banks at all, 1997)) a money metric of utility from the consumption of a free of charge private good. This money metric can be estimated using widely available household data - drawn from household expenditure surveys routinely available in many countries, including UK Expenditure and Food Survey (EFS).

The conclusion emerging from the empirical analysis is that on average households with a child in a state school benefit by an amount equivalent to around 7,5% of total budget compared to an otherwise identical household with a child in a private education. However, further analysis shows that this benefit is associated with attendance in pre-primary and primary education alone; the empirical results do not suggest that the same holds true for households with children in secondary education. Speculating about this result one can argue that households that select a level of education for their children in secondary education comparable to that provided by a private school pay as much to locate themselves in the catchment area of good quality state school as the fee required by the private school.

The methodology proposed for the valuation of free state schooling can be applied to value other free of charge private goods (e.g. health) and, also, in different countries for reasons of comparison and/or for examining issues pertaining to the economic valuation of publicly provided private goods from a wider international perspective (e.g. in the context of the European Union). It can also be an alternative to costly contingent valuation surveys.

5.5 Conclusion and Suggestions for Further Research

The relationship between housing prices and the quality of local state schools has been a topic of interest among economists examining educational issues more generally. Measurement of both the motivation of local authorities to offer higher quality education to their constituency and the willingness of household to pay for this education is based upon evidence showing a positive impact of educational output on house values.

In recent years the epicentre of debate on hedonic valuations of the house price capitalization of state school quality, has been the determination of the school performance measure that best encapsulate education quality as well as the interplay between this and other affecting factors on house prices.

This thesis attempts to address these questions by adopting the newly introduced - and popular among educational economists - CVA index and emphases the importance of choosing an appropriate empirical specification and highlighting the potential heterogeneity in the effects of the various components of this index. The results confirm that the score component of the CVA associated positively with house prices, while the non-score components are insignificant for primary and significant but negative for secondary schools. The main limitation of the analysis in the thesis has been the lack of information to break down the CVA index into all its constituent components (prior achievement, final score, socio-economic background, pupil specific characteristics etc) so as to be able to say more about which of these components are valued most by households at different education levels. This, together with drawing more detailed data on household-, school- and neighbourhood-specific characteristics is an interesting line of future research that can help understand better the motivation of local authorities to offer high quality education and the willingness of household to pay for this quality.

The use of family expenditure survey data to associate house prices with school quality, by approximating the latter with an estimate of the willingness of households to pay for the education of their children, is an approach attempted for the first time in this thesis. While departing from traditional hedonic analysis the proposed approach can be a used in countries where the location of households is not disclosed in order to estimate measures capturing the willingness of households to pay for education through housing: the shadow price of education relative to housing and the effect of 'notional' education expenditure on house prices. Both of these measures can be estimated for households grouped by many socio-economic characteristics of policy interest, such as demographics, income, location etc. Although not attempted in the thesis, analysis of the proposed measures involving different socio-economic groups or analysis at different regions can be a useful empirical extension, as it can provide information about the comparative strength of 'notional' demand for education among these groups, thereby aiding the design of education policies targeting this demand.

The last issue taken up in this thesis moves a step further in seeking a way to estimate the willingness of households to pay for higher quality education and attempts to model demand for education in the context of an integrable complete demand system so to derive a money metric of the free state schooling. The idea here, again the first proposed in the literature, is that the correct value of free state schooling should not include the extra cost undertaken by households in order to secure access to better quality state schools. The results suggest that only at pre-primary and primary state schools households can get education of private-school quality at a lower cost. Clearly, this is a first result using a new approach. Theoretical issues need to be resolved and extensive empirical investigation should be performed before one can reach convincing results about the value of free state schooling to consumers. Nevertheless, the methodology proposed in the thesis can be a first step to filling a gap in the literature concerning the use of econometric techniques for the valuation of education and other private goods freely provided by the state. As such it can be useful in the context of the debate about school vouchers programs and, more generally, in examining questions about the optimal public-private mix in the provision of private goods.

Bibliography

Atkinson A., J. Gomulka and N. Stern (1990), “Spending on Alcohol: Evidence from the Family Expenditure Survey 1970-1983”, Economic Journal, 100, 808-827.

Banks J., R. Blundel and A. Lewbel (1997), “Quadratic Engel Curves and Consumer Demand”, Review of Economics and Statistics, 4, 527-539.

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

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

Barzel Y. (1973), “Private Schools and Public School Finance”, The Journal of Political Economy, 81 , 174-186.

Becker Gary S. (1964), Human Capital: A theoretical and empirical Analysis, New York: Columbia University Press for the National Bureau of Economic Research.

Behrman J., Z. Hrubec, P. Taubman and T. Wales (1980), “Socioeconomic Success: A study of the effects of genetic endowments, family environment, and schooling”, North-Holland publishing company Amsterdam-New York-Oxford.

Behrman J., and P. Taubman (1986), “Birth Order, Schooling, and Earnings”, Journal of Labour Economics, 4, 121-145.

Behrman J., R. Pollak and P. Taubman, (1989), “Family Resources, Family Size, and Access to Financing for College”, The Journal of Political Economy, 97, 398-419.

Behrman J. (1997), “Mother’s Schooling and Child Education: A Survey”, University of

Besley T. and S. Coate (1991), “Public Provision of Private Goods and Redistribution”, American Economic Review, 81, 979-984.

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

Blackorby C., D. Primont and R. Russel (1978), Duality, separability, and functional structure: Theory and economic applications, North-Holland, Amsterdam.

Blackorby C. and A. Shorrocks (1996), eds, Collected Works of W.M. Gorman: Separability and aggregation, Clarendon Press, Oxford.

Blanden J. and S. Machin (2004), “Educational Inequality and the Expansion of UK Higher Education”, Scottish Journal of Political Economy, Special Issue on the Economics of Education, 54, 230-249

Blomquist S. and V. Christiansen (1995), “Public Provision of Private Goods as a Redistributive Device in an Optimum Income Tax Model”, Scandinavian Journal of Economics 97, 547-567.

Blomquist S. and V. Christiansen (1999), “The Political Economy of Publicly Provided Private Goods”, Journal of Public Economics, 73, 31-54.

Blundell R and I. Walker (1982), “Modelling the Joint Determination of Household Labour Supplies and Commodity Demands”, The Economic Journal, 92, 351-364.

Blundell R and I. Walker (1986), “A Life-Cycle Consistent Empirical Model of Family Labour Supply Using Cross-Section Data”, The Review of Economic Studies, 53, 539-558.

Blundell R, L. Dearden and B. Sianesi (2001), “Estimating the Returns to Education: Models, Methods and Results”, Centre for Economics and Education, LSE.

Blundell R, P. Pashardes and G. Weber (1993), “What do we learn about consumer demand patterns from micro-data”, American Economic Review, 83, 570-597.

Blundell R. and A. Lewbel (1991), “The Information Content of Equivalence Scales”, Journal of Econometrics, 50, 46-68.

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

Bogart T. W. and A. Cromwell (2000), “How Much is a Neighbourhood School Worth?”, Journal of urban Economics, 47, 280-305.

Bowels S. (1972), “Schooling and Income Inequality from Generation to Generation”, The Journal of Political Economy, 80,219-251.

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 (2006), “Educational Outcomes and House Values: A test of the Value Added Approach”, Journal of Regional Science, 46, 2, 245-268.

Brasington D. (2006), “School Choice and the Flights to Private Schools: To What Extent are Public and Private Schools Substitutes?”, Working Paper 2006-04, Luisiana State University (http://www.bus.lsu.edu/economics/papers/pap06_04.pdf).

Brookshire D. S. and D. Coursey (1987), “Measuring the Value of a Public Good: An Empirical Comparison of Elicitation Procedures", American Economic Review, 77, 554-566.

Browining M. and C. Meghir (1991), “The Effects of Male and Female Supply on Commodity Demands”, Econometrica, 59, 925-951.

Buttin R. J. and J. Cordes (1998), “School Choice in California: Who Chooses Private Schools?”, Journal of Urban Economics, 44, 110-134.

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 (1998), “Estimating the Demand for Housing , Land, and Neighbourhood Characteristics”, Oxford Bulletin of Economics and Statistics", 60, 357-382.

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

Chevalier A., C. Harmon, I. Walker and Y. Zhu (2004), “Does Education Raise Productivity, or Just Reflect it?”, Economic Journal,114, 499-517.

Christensen L., D. Jorgenson and L. Lau (1975), “Transcendental Logarithmic Utility Functions”, American Economic Review, 65, 367-383.

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.

Clinch J. and A. Murphy (2001), “Modelling winners and losers in contingent valuation of public goods: appropriate welfare measures and econometric analysis”. The Economic Journal, 111, 420-442.

Deaton A. and J. Muellbauer (1980a), Economics and Consumer Behavior, Cambridge University Press, New York.

Deaton A. and J. Muellbauer (1980b), “An Almost Ideal Demand System”, American Economic Review, 70, 312-326.

Deaton A. and J. Muellbauer (1986), “On Measuring Child Costs: With Applications to Poor Countries”, Journal of Political Economy, 94, 720-744.

Dickens R., V. Fry and P. Pashardes (1993), “Non-Linearities and Equivalence Scales”, The Economic Journal, 103, 359-368.

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.

Epple D. and R. Romano (1996a), “Public provision of private goods”, Journal of Political Economy, 104, 57-84.

Epple D. and R. Romano (1996b), "Ends against the middle: Determining public service provision when there are private alternatives", Journal of Public Economics, 62, 297-325.

Epple D. and R. Romano (1998), “Competition between Private and Public Schools, Vouchers, and Peer-Group Effects”, American Economic Review, 88(1), 33-62.

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

Fack G. And J. Grenet (2010), “When Do Better Schools Raise Housing Prices?”, Evidence from Paris Public and Private Schools, Journal of Public Economics, 94, 59-77.

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.

Fernandez R. and R. Rogerson (2003), "Equity and Resources: An Analysis of Education Finance Systems", Journal of Political Economy, Vol.111, No.4.

Ferreyra M. (2007), “Estimating the Effects of Private School Vouchers in Multidistrict Economies”, The American Economic Review, 87, 789-817.

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: Admission 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.

Goldhaber D. (1999), “An Endogenous Model of Public School Expenditures and Private School Enrolment”, Journal of Urban Economics, 46, 106-128.

Gorman T. (1976), “Tricks with utility functions”, in (M.Artis and R. Nobay, eds), Essays in Economics, Cambridge: Cambridge University Press.

Gorman T. (1981), Some Engel curves, in Essays in the Theory and Measurement of Consumer Behaviour in Honor of Richard Stone, ed. A. Deaton, Cambridge University Press.

Greene W. H. (2003), Econometric Analysis, 5th Ed., Prentice Hall, New Jersey.

Gronau R. (1974), “Wage Comparisons: A Selectivity Bias”, Journal of Political Economy, 82, 1119-1143.

Hamilton B.W., M. Macauley (1991), “Determinants and Consequence of the Private-Public School Choice”, Journal of Urban Economics, 29, 282-294.

Hanemann W.M. (1994), “Valuing the Environment Through Contingent Valuation”, Journal of Economic Perspectives, 8, 19-43.

Hanushek E. (1986), “The Economics of Schooling: Production and Efficiency in the Public Schools”, The Journal of Economic Literature, 24, 1141-1177.

Hanushek E. (1992), “The Trade off between Child Quantity and Quality”, Journal of Political Economy, 100, 84-117.

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

Harmon C. and I. Walker (1995), “Estimates of the Economic Return to Schooling for the UK”, The American Economic Review, 85, 1278-1286.

Harmon C., H. Oosterbeek and I. Walker (2000), “The Returns to Education, A Review of Evidence Issues and Deficiencies in the Literature”, Centre for Economics and Education, LSE.

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

Haurin D. R. and D. Brasington (2009), “Parents, Peers, or School Inputs: and Which Components of School Outcomes Are Capitalized into House Value?”, Regional Science and Urban Economics, 39, 523-529.

Hausman J.A, W. Newey and J. Powell (1995), “Nonlinear Errors in Variables: Estimation of Some Engel Curves”, Journal of Econometrics, 65, 205-233.

Haveman R. and B. Wolfe (1995), “The Determinants of Children’s Attainments: A Review of Methods and Findings”, Journal of Economic Literature, 33, 1829-1878.

Hayes K. J. and L. Taylor (1996), “Neighborhhod School Characteristics: What Signals Quality to Homebuyers?”, Federal Reserve Bank of Dallas Economic Review, 3, 2-9.

Heckman J. (1976), “The Common Structure of Statistical Models of Truncation, Sample Selection, and Limited Dependent Variables and a Simple Estimator for Such Models”, The Annals of Economic and Social Measurement, 5, 475-492.

Heckman J. (1979), “Sample selection bias as a specification error”, Econometrica, Vol. 47,pp. 153-161.

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

Ireland N. (1990), “The Mix of Social and Private Provision of Goods and Services”, Journal of Public Economics, 43, pp. 201-219.

Kalmijn M. (1994), “Mother’s Occupational Status and Children’s Schooling”, America sociological Review, 59, 257-275.

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

Ladd H. (2002), “School Vouchers: A critical View”, The Journal of Economic Perspectives, 16, 3-24.

Lankford H. R. and J. H. Wyckoff (1992), “Primary and Secondary School Choice among Public and Religious Alternatives”, Economics of Education Review, 11, 317-337.

Lankford H. R., Lee S. E. and Wyckoff H. J. (1995), "An Analysis of Elementary and Secondary School Choice", Journal of Urban Economics, 38, 236-251.

Leech D. and 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, 135-154.

Levy G. (2005), “The Politics of Public Provision of Education”, The Quarterly Journal of Economics, 120, 1507-1534.

Lewbel A. (1986), “Additive Separability and Equivalence Scales”, Econometrics, 54, 219-222.

Lewbel A. (1990), “Full Rank Demand Systems", International Economic Review, 31, 289-300.

Lewbel A. (1991a), “The rank of demand systems: Theory and Nonparametric Estimation”, Econometrica, 59, 711-730.

Lewbel A. (1991b), “Cost of Characteristics Indices and Household Equivalence Scales”, European Economic Review, 35, 1277-1293.

Long E. J. and E. F. Toma (1988), “The Determinants of Private School Attendance, 1970-1980”, The Review of Economics and Statistics, 70, 351-357.

Lyssiotou P., P. Pashardes and T. Stengos (2004), “Estimates of the Black Economy based on consumer demand approaches”, The Economic Journal, 114, 622-640.

Martinez-Mora F. (2006), “The Existence of Non-Elite Private Schools”, Journal of Public Economics, 90, 1505-1518.

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

McCaffrey D. F., K. Lockwood, T. Louis and L. Hamilton (2004), "Models for Value-Added Modelling of Teacher Effects", Journal of Educational and Behavioural Statistics, 29, 67-101.

Muellbauer J. (1974), "Household Composition, Engel Curves and Welfare Comparisons between Households: A Duality Approach", European Economic Review, 5, 103-122.

Neal D. (2002), "How Vouchers Could Change the Market of Education", Journal of Economic Perspectives, 16, 24-44.

Nechyba T. (1999), "School Finance Induced Migration and Stratification Patterns: The Impact of Private School Vouchers", Journal of Public Economic Theory, 1, 5-50.

Nechyba T. (2000), "Mobility, Targeting, and Private Schools Vouchers", The American Economic Review, 90, 130-146.

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, 82, 957-971.

Pashardes P. (1991), "Contemporaneous and Intertemporal Child Costs", Journal of Public Economics, Vol. 45, pp. 191-213.

Pashardes P. (1995), "Equivalence Scales in a Rank-3 Demand System", Journal of Public Economics, 58, 143-58.

Pashardes P., P. Nearchou and N. Empora (2004), "Economics of education in Cyprus: characteristics, quality and returns" Economic Policy Paper no. 01-04, Economics Research Centre, University of Cyprus

Pollak R and T. Wales (1992), Demand System Specification and Estimation, New York: Oxford University Press.

Pollak R. (1969), "Conditional Demand Functions and Consumption Theory", Quarterly Journals of Economics, 83, 70-78.

Pollak R. and T. Wales (1978), "Estimation of Complete Demand Systems form Household Budget Data: the Linear and Quadratic Expenditure Systems", American Economic Review, 68, 348-359.

Pollak R. and T. Wales (1979), "Welfare Comparisons and Equivalence Scales", American Economic Review, 69, 216-221.

Pollakowski H. (1973), "The Effects of Property Taxes and Local Spending on Property Values: A comment and Further Results", The Journal of Political Economy, 81, 994-1003.

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

Robinson M. (1988), “Root-N consistent Semi-parametric Regression”, Econometrica, 56, 931-954.

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

Rosen H. and D. Fullerton (1977), “A Note n Local Tax Rates, Public Benefit Levels and Property Values”, The Journal of Political Economy, 85, 433-440.

Rosenthal L. (2003), “The Value of Secondary School Quality", Oxford Bulletin of Economics and Statistics”, 65, 329-355.

Saunders L (1999), “A Brief History of Educational Value-Added: How did we get to where we are?”, School Effectiveness and School Improvement, 10, 233-256.

Sheppard S. (1999), Hedonic Analysis of Housing Market in: P. Cheshire (ed.), Handbook of Urban and Regional Economics, Elsevier Science.

Smith K. V. and J. C. Huang (1995), “Can markets value air quality? A meta-analysis of Hedonic Property value models”, Journal of Political Economy, 103, 209-277.

Sonstelie J. (1982), “The Welfare Cost of Free Public Schools”, The Journal of Political Economy, 90 (4), 794-808.

Sonstelie J. and P. Portney (1980), “Gross Rents and Market Values: Testing the Implications of Tiebout’s Hypothesis”, Journal of Urban Economics, 7, 102-118.

Stiglitz J.E. (1974), “The Demand for Education in Public and Private School Systems”, Journal of Public Economics, 3, 349-385.

Stone J. (1954), The Measurement of Consumers’ Expenditure and Behaviour in the United Kingdom, 1920-1938, Vol.1, Cambridge University Press.

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

Theil H. (1965), “The Information Approach to Demand Analysis”, Econometrica, 33, 67-87.

Tiebout C. (1956), “A Pure Theory of Local Expenditures”, Journal of Political Economy, 64, 416-124.

Trostel P. , I. Walker and P. Woolley (2002), “Estimates of the Economic Return to Schooling for 28 Countries”, Labour Economics, 9, 1-16.

Verbeek M. (2002), A Guide to Modern Econometrics, Wiley, Chichester.

Walker I. and Y. Zhu (2003), “Education, Earnings and Productivity: Recent UK Evidence”, Labour Market Trends, UK.