University College London

Department of Economics

EMPirical ESSays IN THE

ECONOMICS OF NeIGHBOURHOODS

AND E DUCATION

Stephen Anthony Gibbons

Thesis submitted for the degree of Ph.D.

2002
Abstract

This Thesis presents a series of empirical studies that investigate the community dimension to opportunity and achievement in Britain. It looks at the impacts of social interactions at the neighbourhood level, both in terms of the direct effects on educational attainments, and in terms of the impacts of implicit trade in community goods that takes place via the housing market. The research provides new empirical insights into neighbourhood effects in Britain using micro-econometric techniques and presents some new applications of semi-parametric methods. The first empirical Chapter explores the link between educational outcomes and neighbourhood of upbringing and finds evidence of direct neighbourhood impacts on early and adult achievements. Next, in Chapter 3 we use school performance data and postcode-matched spatial data to evaluate the importance of area-based factors and specific school level inputs in primary school production functions. Here we look for spatial clustering of primary school quality and urban effects on performance as evidence of school or neighbourhood social interactions. Chapter 4 investigates willingness to pay for the stock of neighbourhood human capital as a means to evaluating an upper bound to the benefits of neighbourhood and the community benefits of education. The next essay in Chapter 5 looks at a particular component of the demand for good neighbourhoods – the demand for high-performing primary schools – and shows that households are prepared to pay for technologies that improve primary school performance. Finally, Chapter 6 turns away from explicitly educational issues and explores the importance of crime and community disorder at the neighbourhood level, again using a property value model. We show that it is visible crimes signalling community disorder that matter the most, and that willingness to pay to avoid high-crime areas far exceeds the direct cost of these crimes.
Acknowledgements

I would like to thank Stephen Machin for supervision and support during the preparation of this thesis, and for his contribution to our joint work on the valuation of primary schools in England. I am also grateful for financial support from the ESRC, and to the Centre for Economic Performance and the Department of Geography and Environment at the London School of Economics for the time and facilities to complete the work. Thanks also to my colleagues at the LSE, and especially the staff who have provided computing and data support. As always, the comments of innumerable participants at seminars and conferences have been invaluable in guiding my research. And thanks to Rachel, Sam and Laura for their (almost) unlimited patience and tolerance.
## Contents

### List of main sections

<table>
<thead>
<tr>
<th>Contents</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>List of main sections</td>
<td>4</td>
</tr>
<tr>
<td>List of Tables</td>
<td>9</td>
</tr>
<tr>
<td>List of Figures</td>
<td>11</td>
</tr>
</tbody>
</table>

### 1 Introduction

1.1 Overview and context                                                | 13   |
1.2 Thesis structure                                                    | 16   |
1.3 The Chapters outlined                                               | 19   |
  1.3.1 Neighbourhood Effects on Educational Achievements               | 19   |
  1.3.2 The Neighbourhood Dimension to Primary School Performance       | 20   |
  1.3.3 Paying for Neighbourhood Human Capital                          | 21   |
  1.3.4 Valuing Primary Schools                                         | 22   |
  1.3.5 The Costs of Urban Property Crime                               | 23   |
1.4 Data sources                                                        | 23   |

### 2 Neighbourhood Effects on Educational Achievement

2.1 Introduction                                                        | 25   |
2.2 Literature and context                                              | 28   |
2.3 Models and empirical identification strategies                       | 32   |
  2.3.1 A simple model                                                  | 32   |
  2.3.2 Assumptions in alternative specifications                       | 35   |
  2.3.3 School-quality selection                                         | 38   |
  2.3.4 Social tenants and parental selection                            | 40   |
4.4 Description of the data ...........................................................................................161

4.5 Empirical methods ...............................................................................................163

4.5.1 Empirical problems in the measurement of neighbourhood value ..............163

4.5.2 Empirical model ..............................................................................................165

4.5.3 Instrumental variables identification strategy ..............................................168

4.6 Empirical Results ...............................................................................................170

4.6.1 Summary and assessment of the data .........................................................170

4.6.2 Implicit price of neighbourhood educational status ....................................173

4.6.3 Non-linearities in response ..........................................................................182

4.6.4 Comparing local income and education effects ..........................................184

4.6.5 Robustness to unobserved neighbourhood heterogeneity .........................185

4.6.6 Evidence for human capital externalities .....................................................188

4.6.7 Other evidence .............................................................................................189

4.7 Concluding discussion .......................................................................................190

4.8 Appendix A: Changes in educational composition .........................................194

4.9 Appendix B: Neighbourhood incomes .............................................................196

4.10 Appendix C: Estimating non-linearities .........................................................197

5 Valuing Primary Schools .......................................................................................199

5.1 Introduction ........................................................................................................199

5.2 English primary schools and the housing market ............................................203

5.2.1 The hedonic approach ................................................................................203

5.2.2 Does location matter for primary school admission? ................................206

5.3 The Empirical Approach ..................................................................................208

5.3.1 Empirical model ..........................................................................................208

5.3.2 Estimation strategy ......................................................................................209

5.3.3 Bandwidth choice ......................................................................................211

5.3.4 Further identification through Instrumental Variables ..............................211
Table 2-12: Neighbourhood education and family effects on women’s age 33 qualifications ..........................................................................................................86

Table 2-13: Neighbourhood education effects on non-social tenants, age 33 qualifications ..................................................................................................................................................88

Table 3-1: Rank adjacency statistics, Key Stage 2 performance measures, 1996-1999 mean ......................................................................................................................118

Table 3-2: School sensitivity to neighbourhood conditions by region, 1996-1999 ......121

Table 3-3: Key Stage 2 Performance in English Primary Schools, 1996-2001 ..............122

Table 3-4: School sensitivity to catchment area status index by region, 1996-1999 .....124

Table 3-5: Catchment area status and primary school performance, 1998 & 1999 ......126

Table 3-6: Catchment area attributes and primary school performance, 1998 & 1999...128

Table 3-7: Primary school interaction effects – spatial autocorrelations with J schools.131

Table 3-8: Resources, school structure and primary school performance, 1998 & 1999137

Table 3-9: Factor loadings and scoring coefficients for neighbourhood status ..........148

Table 3-10: Positive Eigenvalues of principal factors of catchment area status ..............148

Table 3-11: Relationship between local tenancy group incomes ......................................149

Table 4-1: Summary statistics for local education and property prices ............................171

Table 4-2: Property price response to neighbourhood educational composition: Postcode-sector data, by region, 1995 ..............................................................175

Table 4-3: Sensitivity of education parameter estimates to bandwidth choice ............179

Table 4-4: Property price response to local education: SML data, by region, 1997 ......180

Table 4-5: Results for London using property-level data, 2001..........................................182

Table 4-6: Robustness to neighbourhood heterogeneity, pooled regions ....................186

Table 4-7: Property price response to neighbourhood incomes, 1996 and 1999 ..........196

Table 5-1: Descriptive Statistics ...........................................................................................220

Table 5-2: Property price response to primary school performance: Postcode-sector data, by region ..........................................................................................................224

- 10 -
Table 5-3: Sensitivity of Property Price Models to Specification

Table 5-4: Comparison Of Coefficients From Cross-Local Authority Boundary Effects, SSE Estimator and Property-level Micro Data For Greater London Area

Table 6-1: Summary statistics

Table 6-2: London property prices and property crimes, 2001

Table 6-3: Robustness checks

Table 6-4: Postcode Sector Data, 2000-2001

Table 6-5: Alternative instruments for criminal damage

Table 6-6: Association between year on year changes in Police Force Area crime rates, 1997-1999

List of Figures

Figure 1-1: Relationship between social housing and area deprivation

Figure 2-1: Attainments and childhood neighbourhood; semi-parametric estimates

Figure 2-2: Attainments and childhood neighbourhood: semi-parametric estimates

Figure 2-3: Attainments of social tenants: semi-parametric estimates

Figure 2-4: Educational intergenerational mobility and neighbourhood

Figure 3-1: Illustrative pseudo-catchment area matching on Census Enumeration Districts for a school in Hackney, London

Figure 3-2: a) Quintiles of Key Stage 2 scores, by English County, 1999

Figure 3-3: a) Quintiles of Key Stage 2 scores, by Parliamentary Constituency

Figure 3-4: a) Clusters of below-median primary school performance (free-school meal adjusted, Local Moran I p-values)

Figure 3-5: Above-median primary performance clusters in London Constituencies (free-school meal adjusted, Local Moran I p-values)

Figure 3-6: Primary school performance and catchment area characteristics
Figure 3-7: Relationship between primary school performance, local property prices and social-tenant children, 1996-1999................................................................. 134

Figure 4-1: Smoothed Postcode-sector proportion with high qualifications: London and surrounding area 1991 ............................................................................................................. 174

Figure 4-2: Spatially averaged kernel regression line for the London area: ................... 184

Figure 4-3: Proportion with high qualifications (degrees, diplomas and teaching qualifications) 91-99 .............................................................................................................. 194

Figure 5-1: Example Postcode District, Postcode Sectors and Geographical Distribution of School Types ........................................................................................................ 218

Figure 5-2: Relationship Between Log House Prices in 1999 and Mean 1996-1998 Primary School Performance – Deviations From Postcode District Means ..222

Figure 5-3: Example House Price-Location Surface For London, From Smooth Spatial Effect Model ................................................................................................................. 223

Figure 6-1: Incidents of Burglary in a Dwelling per km², April 1999-March 2001 ........253

Figure 6-2: Incidents of Criminal Damage to Dwellings per km², April 1999-March 2001 ......................................................................................................................... 254

Figure 6-3: Log Property Prices, First 6 Months of 2001 ................................................. 255

Figure 6-4: Residual Log-Property Price Surface .............................................................. 256

Figure 6-5: Crime trends in Metropolitan Police Force Area, 1993-2001 ..................... 268

Figure 6-6: Association between local deviations in Criminal Damage density and local deviations in log property prices......................................................... 274
1 Introduction

1.1 Overview and context

Spatial concentration of economic and social disadvantage draws our attention to these problems at the individual level. Geography mediates the individual experience of deprivation to the public consciousness. Our perceptions of relative wealth and deprivation are what we see as we move between neighbourhoods in the city, as we move from region to region, and from nation to nation. At international and broader sub-national levels, these spatial inequalities arise partly through differences in natural endowments – the ‘first nature’ geography of terrain and natural resources – and partly through historical processes of accumulation that lead to spatial diversity in the skills and characteristics of populations as they migrate in search of economic opportunities. But at the intra-regional and intra-urban levels, many of the patterns we observe arise through the sorting of similar individuals into spatial clusters, partly because of shared desires, needs and information, but – probably more importantly – because they are constrained in their housing choices by their incomes and other resources. Differences in individual conditions that would be lost to the observer in a random assignment of individuals across space become vivid through this type of spatial segregation.

This spatial manifestation of inequality has become a key policy concern in Britain. Over £2800 billion was assigned to DETR\(^1\) controlled ‘area-based initaitives’ for 2000-2002, including nearly £1700 million from the Single Regeneration budget targeted at

\(^1\) The Department of Environment Transport and the Regions, subsequently the Department of Transport Local Government and the Regions, and subsequently the Office of the Deputy Prime Minister and the Department of Transport.
deprived areas (Hughes (2000)). On top of this, other Government departments have their own high-profile, area-focussed policies. In education, the Excellence in Cities programme tackles spatial inequalities in school performance, with a budget of £120 million to £300 million per year between 1999 and 2004 (Department for Education and Employment (2001)). The Crime Reduction Programme is to receive some £200 million over three years to tackle crime in high-crime areas, and a £96 million Phoenix fund operates to encourage local enterprise in deprived areas (Social Exclusion Unit (2001) p.9). But area-targeted policies are a blunt tool for dealing with individual disadvantages. They are predicated on the idea that there are specific gains or economies of scale from tackling these issues at the spatial-group level. Implicitly, these policy initiatives assume that area status is a public good, and that tackling disadvantage at the area level offers benefits to everyone – the residents, regardless of their personal circumstances, prospective residents, local employers and the rest of the economy.

A proposition underlying this thesis is that this type of spatial segregation is not just an outcome arising from spatial disparities in employment, housing quality, amenities and public services. It has external effects. The reason people care about the community in which they live is because social interactions with neighbours generates spillovers from group behaviour to individual well-being. These are classic non-pecuniary externalities of the type introduced in any text on public microeconomics, and have some of the characteristics of local public goods (e.g. Atkinson and Stiglitz (1980), p. 7 and Chapter 17). These social interaction-based externalities provide the link between spatial inequality as an outcome, and as a determinant of individual inequality. Importantly for some of the methods used in this thesis, individuals express their preferences over the characteristics of their neighbours through their housing choices.

If we consider inequality expressed as inequalities in welfare (utility, happiness or well-being), then the link from community to individual outcomes could be direct. There is, perhaps, a personal utility cost to living amongst unhappy neighbours. We can also
think in terms of a penalty for deviation from group behaviour, so that there is always an incentive to conform to the neighbourhood norm (Brock and Durlauf (2001)). But we also need to consider impacts from the status of a neighbourhood – average education, skills, incomes, employment, crime or whatever – to the productive or social capabilities of an individual. In particular we will be thinking about the effects of neighbourhoods on the acquisition of human capital in children, and the effects this has on the residential choices of parents.

Economics is a social science, and all of the processes we study are the product of social interactions at various levels of abstraction (Manski (2000)). Markets work by social interaction, whether this is at the level of the street market trader, or the brokers of financial derivatives. But the kind of social interactions we should have in mind here are the interactions that generate spillovers, or externalities in the economic sense. These interactions between individuals crystallise as intangible commodities that are not bought and sold explicitly in conventional markets. These commodities are not readily visible and are hard to quantify. But a defining characteristic is that they are, to some extent, geographically localised. This is corollary of their basis in social interactions, which are fundamentally spatial in character. For sure, in this age of high speed telecommunications, the net of social interactions is spread wider. Any yet there are deep issues around the codifiability of knowledge and experience that mean that much of what we share with others is shared within geographically localised spaces (see Leamer and Storper (2001) for discussion in relation to production technologies). It is on this basis that we look for evidence of social interactions and their value to individuals and society through neighbourhood effects.

These considerations are not new. Theoretical analyses of the how preferences over local public goods and community characteristics impact on spatial segregation date back to Tiebout (1956) and Schelling (1969), Schelling (1971), with more recent examples in Benabou (1996), Durlauf (1996), Epple and Platt (1998) and Fernandez and Rogerson
(1996). Others have analysed the impacts of segregation and social interactions on the individual (Coleman (1988), Becker (1996)) and on intergenerational mobility and individual inequality (Loury (1977), Borjas (1992), Kremer (1997)). A vast, largely US-focussed literature tries to measure these effects, with varying degrees of success (see Chapter 2 for references). The main aim and contribution of this Thesis is to provide empirical estimates of some these effects in the British case, to find improved tools for measurement that are appropriate to the British setting and data.

1.2 Thesis structure

The focus of this thesis is on processes that link individual characteristics and neighbourhood composition. These processes fall into two categories. The first is about the impact of neighbourhood on outcomes of individuals who live or grow up there – here we consider educational attainments in particular. The second is about the way in which neighbourhood characteristics determine the composition of the community through sorting processes. These are, obviously, interrelated. An aspect of community composition that is valued for its contributions to life outcomes will partly determine the equilibrium spatial distribution of community characteristics, assuming heterogeneous preferences and incomes.

Starting with the first category, Chapters 2 and 3, explore the impact of neighbourhood composition on educational attainments of local residents, each Chapter considering two aspects of this process. Chapter 2 looks at the impact of neighbourhoods on individual attainments by adulthood and focuses on identifying the impact of community education levels, over and above family influences. Chapter 3 deals with local effects on the attainments of younger pupils at primary school level in England, and investigates the contributions of schools and local conditions to school performance.

Turning to the second category, Chapters 4, 5 and 6 look for evidence of preferences over neighbourhood attributes, as revealed through property prices. These
chapters use the hedonic-price framework (Rosen (1974)) to elicit marginal willingness to
pay for neighbourhood characteristics. The innovation here is to develop an estimation
strategy based on deviations of prices and characteristics from smooth three-dimensional
surfaces that describe the more general trends in these variables over geographical space.
We apply semi-parametric techniques to estimation of the models. The technique gets rid
of a substantial amount of nuisance variation attributable to distance-to amenity type
factors. This reduces the potential for biases induced by unobserved general
neighbourhood effects.

If we think about popular conceptions of ‘good’ and ‘bad’ neighbourhoods in
Britain we probably think of three key things – education, crime, social housing. The
research here is motivated by these considerations. We start in Chapter 4 with a broad­
brush sketch of how households value community composition – specifically its
educational status. We characterise this by the proportion of highly qualified residents.
We treat neighbourhood education levels as an index of neighbourhood status or
deprivation, and use our estimates to infer the value households attach to life in educated
communities. Whereas Chapter 1 asked to what extent neighbourhood educational status
influenced residents outcomes, Chapter 4 asks how much households are prepared to pay
for these and other benefits of educated communities. In subsequent Chapters we look in
more detail at specific neighbourhood attributes.

So, Chapter 4 looks at the value of communities as ranked by their educational
status. We treat the proportion of educated neighbours as a measure of community human
capital, and as a barometer of community ‘quality’. To be more accurate, we look at the
discount on owner-occupier properties in neighbourhoods with high proportions in social
housing – essentially an exogenous, historically determine policy driven variable – and
translated this into a value attached to educated communities. This analysis is predicated
on an assumption that the proportion in social housing is a key predictor of community
status in Britain, within broader geographical areas. It is easy to show that this is not
unfounded. By factor analysis, we can generate an index of neighbourhood deprivation using a battery of 1991 Census characteristics; the proportion in social housing alone 'explains' up to 70% of the spatial variation in this index. The strength of the relationship clearly seen in Figure 1-1.

**Figure 1-1: Relationship between social housing and area deprivation**

![Figure 1-1: Relationship between social housing and area deprivation](image)

Variables are Postcode-sector means, in deviations from Postcode-district form, to illustrate the micro-geographic relationship.

The importance of this results is that it means we can use the proportion in social housing as a powerful control for socio-economic status in hedonic property value model, without inducing correlation between regressors and omitted neighbourhood attributes that simultaneously determine house prices and neighbourhood composition. In Chapters 5 and 6, we carry this through, along with the semi-parametric methods for removal of

---

2 Postcode sector proportions unemployed, highly qualified, long-term sick, economically active, lone parents, black, Indian/Asian, population density, average age.
general neighbourhood effects, to the analysis of two vital dimensions to community quality – schools and crime. School quality receives a lot of attention in the policy and media domains, and in Chapter 5 we look at how this affects house prices. This Chapter is one of the first to extend the US work on hedonic pricing of school quality to England, and is unique in showing the importance of local primary schooling in household preferences. Whereas Chapter 2 examines the impact of neighbourhoods on primary schools, Chapter 5 examines the impact of primary schools on the neighbourhood. Chapter 6 moves away from directly education-related issues and looks at how aversion to neighbourhood crime and disorder affects house prices, and the implied costs of crime in the London metropolitan area. Again, this is the first piece of evidence for Britain on this crucial policy-relevant question.

1.3 The Chapters outlined

1.3.1 Neighbourhood Effects on Educational Achievements

Area-targeted regeneration policy implicitly assumes that neighbourhoods make a difference to the prospects and achievements of individuals, especially children. Surprisingly, the evidence for this in the British context is sparse. To address this, we estimate the impact of a child's neighbourhood on his or her final educational attainments using data on British children who were teenagers during the 1970s. The paper is the first to look at the implications of neighbourhood influences for social mobility between generations in Britain. The focus is, however, quite specific: we ask whether the characteristics of a residential neighbourhood community for teenagers influence the final level of qualification they obtain. The emphasis is also on measurement of the size of

---

3 In the Conclusion in Chapter 7 we will draw together these two strands.
these effects, and on separating out the causal effect of neighbourhoods, rather than seeking firm explanations.

A novel feature of this study is that we consider the impact on social tenants of living in neighbourhoods with different characteristics. Differences between social tenants in their neighbourhood quality – residents incomes, education or wealth for example – are less related to their own incomes and resources than are differences between property owners' or private tenants' neighbourhoods. This is because social tenants' choices of residential location amongst council homes are less determined by their ability to pay for housing than are the choices of home-buyers and private renters. One view is that relationship we observe between childhood neighbourhood and adult attainments is purely attributable to parental resources, and the fact that more-educated, wealthier families live in more educated, wealthier neighbourhoods. If this were true, we would not expect to find a link between neighbourhood education levels and the eventual qualifications of social tenant children.

1.3.2 The Neighbourhood Dimension to Primary School Performance

Primary schooling can play a pivotal role in influencing an individual's ultimate educational attainments, skills and life chances. This is clear from existing academic research on the impact of early attainments on later success. It is also evident in parents' willingness to pay for quality in primary schooling, either through prep and pre-prep private schools or through house prices close to good state primaries. But primary schools, particularly in the urban environment, are closely linked to quite small neighbourhood communities. This means that they could have a strong role in mediating community inequalities to inequalities in individual educational attainments and life success. Local interactions between pupils and schools might also be important determinants of school success. Here we investigate some of these geographical aspects
of school performance, using data on most primary schools in England from 1996 to 1999.

Firstly, we ask whether good primary schools are clustered together geographically? If so, is this just due to similarity of pupil intake or teaching resources. Or is it due to more subtle social interactions within neighbourhoods. The kinds of interactions we have in mind here include role-model influences on children's behaviour and achievements, peer-group influences, and sharing of ideas and technologies between schools. Next we ask to what extent neighbourhood quality influences primary school performance measures, and what neighbourhood characteristics matter? Surprisingly there is little existing systematic evidence on this. Given our attention to sorting and selection issues, it is clear we must take some trouble to show that this relationship is not due to better-off parents choosing to live near better schools. Finally we look at whether more school resources overcome neighbourhood disadvantage? We tackle this problem by examining whether additional key resources – more teachers per pupil and higher Local Education Authority expenditures – affect school performance scores.

1.3.3 Paying for Neighbourhood Human Capital

Here we ask whether home-buyers pay more for property in ‘educationally-rich’ neighbourhoods than they do for a similar property in poorer, low-education neighbourhoods. We might expect this to be the case, perhaps because of lower crime rates and a better-maintained physical environment. Following on from Chapter 4, we argue that education in the community matters because of the influence this has on children's acquisition of education and life-skills. These effects include direct effects from adults to children through expectations, role models and skill transfers, alongside peer group effects that operate through interactions between children in the street and at school. The empirical work measures the property price premium attracted by higher-education communities. Our approach is to estimate how property prices change from one
neighbourhood to the next as residents educational status changes. Following the literature on hedonic prices in property markets, we take this to measure the value, in monetary terms, that a household places on improvements in the community. Assuming that it is education specifically that matters, then we can infer households' valuation of educational improvements in general. This leads us to a rough estimate of the local community benefits of improvements in education, expressed in monetary terms.

We address a number of important methodological issues that are relevant to amenity valuation in many applications. The first innovation is to develop a semi-parametric methodology for estimation of the implicit prices of local amenities, when the data contains limited information on neighbourhood and property characteristics. Our use of spatially accurate, non-parametrically estimated neighbourhood property price effects allows us to strip down the number of explanatory variables in the hedonic price function and work on variation within small geographical areas that are relatively homogenous. Combining this with an Instrumental Variables approach allows us to deal with the endogeneity of community characteristics in a property price equation, as implied by residential sorting processes.

1.3.4 Valuing Primary Schools

Parents of young children in England are clamouring to get their kids into good state primary schools. But high demand for good schools, coupled with policy that rations admissions to local residents, means that parents who move close to good schools must pay through higher house prices. And how much must they pay? This is the question we answer in this Chapter. Government policy has, at least in principle, made parental preference the key factor in primary school admissions. In practice, demand for places in schools that perform well outstrips the number of places available. Constraints on class sizes mean that it is no longer possible to increases school size to accommodate excess demand. As a result, places must be rationed on the basis of other criteria — most
importantly residential proximity. One of the few ways parents can increase the chances of admission to school of their choice is to move as close as possible to it. Indeed, we know of lots of anecdotal evidence to suggest that parents are prepared to move to try to secure admission to a good school, and that they pay a high premium for this privilege.

Although widely recognised and researched in the US academic arena, this issue has received much less attention in Britain, despite the popular interest. This research fills this gap. We measure the price premium attracted by neighbourhoods with observably better schools. The empirical approach is similar to Chapter 4. We estimate how property prices change from one neighbourhood to the next, and over time, as primary school performance changes.

1.3.5 The Costs of Urban Property Crime

Whilst earlier Chapters focus on the production of education in the local community, the last empirical chapter looks at crime and social disorder. We turn the empirical tools of Chapters 4 and 5 to measure the impact of neighbourhood crimes on property prices for a large metropolitan area – London. Again we deal carefully with issues to do with endogeneity issues and omitted neighbourhood amenities. The Chapter reflects on the extent to which we measure the direct costs of crime to the individual, or more intangible dis-benefits of community disorder

1.4 Data sources

There is nothing new in the idea of a role for geographical location in the determination of educational outcomes, or in the idea that educated communities are desirable commodities. But analysis of these issues in Britain – except through small-scale case studies – has been hampered by the lack of useful data. A principle contribution of this thesis is to undertake an empirical analysis of these relationships in Britain using large-scale datasets, which up until a few years ago would have been
infeasible. Confidentiality issues mean that individual or household level information with identifiers of residential location are rarely available. On exception is the National Child Development Survey, which tracks at intervals up to the present, a cohort of children born in 1958. This data set has residential address Postcodes and matched Census data available for 1974s and 1981, allowing us to track the impacts of neighbourhood in a child's teenage neighbourhood on their adult outcomes. This is what we do in Chapter 1. The NCDS offers the benefits of a rich individual and family data, plus a life-cycle perspective. But we need to trust that there has been no fundamental change in individual responses to neighbourhood conditions since the 1970s if the results are to have contemporary policy relevance. Subsequent chapters find other ways to investigate neighbourhood impacts using more recent data.

Up until the 1990s, the decennial Census provided the only information on local conditions in Britain, and despite being a rich data source, provides no information on local incomes, property prices or educational attainments of young residents. Since the mid-1990s, new data sources have become available which allow us to say something more interesting about the processes at work at the neighbourhood level. School performance 'league tables' published by the DfES allow mapping of school performance to geographical locations for the whole of England. The availability of information on Postcoded property transactions from the Land Registry means we can now track the geographical distribution of property prices at a locally disaggregated level. Since 1996, the marketing company CACI has produced a commercial dataset of local incomes based on their own surveys and Census information. It is these data sets, coupled with the Census, that form the backbone of the empirical analysis in this Thesis.
2 Neighbourhood Effects on Educational Achievement

2.1 Introduction

Does neighbourhood quality affect a child’s ultimate educational attainment? This is the central question addressed in this Chapter. Specifically, we consider whether or not the educational composition of the resident population in a neighbourhood makes a difference to the academic achievements of children who grow up there. We will focus on identifying this effect. For sociologists and psychologists, acceptance of the effect of neighbourhoods on behaviour, development and action follows naturally from social and psychological theory. The relevant empirical questions are more along the lines of “how big are the effects” and “through what channels are they mediated”. Economists, on the other hand, tend to be sceptical about the very existence of neighbourhood effects on attainments. We are more inclined to attribute apparent associations between neighbourhood and individual outcomes to family-based inputs and geographical sorting of like families, or to local school quality and funding. This study carefully compares the results of various econometric approaches to uncover evidence that neighbourhood does indeed matter, albeit in a relatively small way once we account for family and individual differences, school quality differences, and parental selection of residential neighbourhood.

The main results are based on data from the National Child Development Study (NCDS). This data set is unique amongst large British samples in observing individual educational outcomes, providing a rich description of family and school background, and identifying childhood neighbourhood residence to the neighbourhood level. From this, we can match data on neighbourhood characteristics from the 1971 and 1981 Census to the
cohort's residential addresses in childhood and early adulthood. The original NCDS birth cohort dates from 1958, so any results based on the experience of these children in the 1960s and 1970s have something of a historical flavour. A newer cohort survey, the 1970 British Cohort Study (BCS), has neither neighbourhood identifiers nor address Postcodes, so is less useful for our purposes. However, we can compare broader area effects and changes in overall intergenerational educational mobility between the NCDS and BCS.

Measurement of neighbourhood effects on individual outcomes is plagued by well-known empirical problems. The most serious issue arises from the sorting of families by resources into areas of differing residential quality, and the potential for like-minded parents to select neighbourhoods and schools on the basis of their anticipated effects on child outcomes. These factors mean that similar families tend to be spatially clustered. Separating contextual neighbourhood influences from the direct effect of family inputs is difficult. Saturating an empirical regression model with parental characteristics is a doomed strategy, since the precise operational neighbourhood group is rarely defined or known, and the relevant neighbourhood characteristic is measured with error. Parental characteristics can be good proxies for neighbourhood characteristics and tend to swamp background variation in measured neighbourhood attributes. Estimates obtained this way will most likely be small or imprecisely measured, since selection by parents on neighbourhood characteristics means that there may be little useful variation in neighbourhood quality, conditional on parental characteristics.

In practice, no single, non-experimental method can provide consistent estimates of the influence of a neighbourhood on a randomly assigned individual. The approach taken in this Chapter is to compare results from a number of empirical strategies. Firstly it explores the impact of adding and removing key factors in a traditional human capital production function with neighbourhood inputs. Secondly, we test for the presence of school selection bias in the estimates by using local variation in property characteristics to predict neighbourhood quality. This strategy assumes that property characteristics will be
unaffected if motivated parents or children converge on good quality schools. Thirdly, we treat social tenants as randomly assigned to neighbourhoods, relative to the selection processes that bias estimates of neighbourhood effects, and estimate the magnitude of effects on children in this group.

Few empirical studies attempt to separate out community influences on individual outcomes from school-based influences and most blur the distinction between peer group effects in the class-room and role model effects from adults. This study shows that the educational status of the community – as measured by the proportion of highly qualified adults – is the strongest available neighbourhood-level predictor of individual educational attainment, from amongst a selection of Census variables of the type commonly used to measure neighbourhood deprivation. Moreover, this educational status variable has an impact on individual attainments over-and-above its potential peer-group-related effects on local school performance. The existence of these non-schooling-related effects, and the impact of owner-occupier characteristics on social tenants, is suggestive of role-model effects operating through the formation of expectations based on observation of the local community.

The structure of the Chapter is as follows: Section 2.2 briefly reviews the existing literature, to set the work in context. Section 2.3 describes the estimation strategy in some detail. It starts with a simple linear human capital production function model, and uses this to develop various empirical strategies for identifying a structural neighbourhood effect on attainment. Section 2.4 describes the data set, sample and variables. Section 2.5 presents and discusses the empirical results on attainment of adult qualifications, and abilities and aspirations at compulsory school leaving age. Section 2.6 provides an overview of the implications of area-related effects for intergenerational educational mobility and inequality. Concluding remarks appear in Section 2.7.
2.2 Literature and context

Although there are earlier examples (e.g. Datcher (1982)), much of the recent interest in the effect of neighbourhoods on individual's educational and labour market outcomes stems from the work of Wilson (1987). Wilson argued that the increased concentration of poverty and worklessness in inner-city districts in the US has had an adverse effect on the behaviour and development of residents in these neighbourhoods. Wilson sees work, and the expectation of work, as central to a community's discipline, organisation and social cohesion. This idea of breakdown in organisation and social relations is often cast in terms of social capital (Coleman (1988); Coleman (1994)). Social capital extends the ideas of human capital to investments and changes in the systems of relationships in a community that facilitate individual action. In other strands of the literature, neighbourhood influences are explained in terms of Bronfenbrenner's ecological models of child development, in which neighbourhood provides one of numerous contexts for individual development (Bronfenbrenner (1979)). Other approaches, such as that of Sampson and Byron Groves (1989), refer to Shaw and Mackay's social disorganisation theory (Shaw and Mackay (1942)) in which low economic status, ethnic heterogeneity and residential mobility in the community lead to breakdown in social organisation and consequent crime and delinquency. Although much of the empirical research recognises and refers to these theories, the actual approach is usually ad-hoc, and seeks to find influences from various aspects of neighbourhood socioeconomic status on individual outcomes.

This empirical work, and the theoretical discussion of the mechanisms through which these effects are mediated, has been concentrated in the quantitative sociological literature. The range of outcomes analysed is wide: school drop outs, educational attainments, teenage pregnancies, drug and alcohol use, crime victimisation and offences, IQ in infancy, child maltreatment, infant mortality. Researchers' choice of operational
neighbourhood or community characteristics shows similar variety: neighbourhood income and poverty, composite socioeconomic status, occupational status, female-headed families, welfare receipt, joblessness, race, social housing, neighbourhood deprivation indices. This extensive literature is summarised in Gephart (1997) and Jencks and Mayer (1990). The majority of the studies use quite small samples on specially selected groups. Most do not focus on identification issues, beyond controlling for an ad-hoc set of parental characteristics.

Directly related to the current work, and using UK data, is the study by Garner and Raudenbush (1991). This uses data on 2500 young people leaving school from 1984-1986 in one Local Education Authority in Scotland, matched to 1981 Census data. Neighbourhood quality is measured by a deprivation score derived from 12 Census characteristics at Enumeration District level. The authors' estimates show that a 10th to 90th percentile change in neighbourhood deprivation relates to a change in attainments equivalent to around two O-level passes. The strength of their data is that the models can include school dummies to control for secondary school effects, plus primary school age test scores, alongside basic indicators of parental background. The disadvantage is that it focuses on one area and is not easily generalised.

Also focusing on educational outcomes, Kremer (1997) estimates that an additional year of mean Census tract education in the US increases individual education by around 0.14 years, but concludes that changes in residential segregation have little impact on inequality and intergenerational mobility. Casting neighbourhood effects in terms of ethnic group effects, Borjas (1992, 1995) finds an impact from mean ethnic group education levels on education years. Jensen and Seltzer (2000) use a small sample of Australian pupils from 1996 and find influences from neighbourhood income, unemployment or educational attainment on intentions to continue in education. Drawing a distinction between immediate and broader neighbourhood impacts, Overmann (2002) finds that the proportion in the community with vocational qualifications in the wider
neighbourhood increases drop out rates in a sub-sample of the Australian Youth Survey. Community vocational qualifications have a stronger impact than neighbourhood educational qualifications or incomes, but their impact is reversed in smaller micro-neighbourhoods. Duncan (1994) finds significant effects from neighbourhood incomes on white males and more affluent groups, but no effects on disadvantaged groups, similar to Datcher (1982), who finds significant income effects on years of education for whites only. A more extensive body of literature describes the effects of environment on children's behaviours and early attainments, for example Brooks-Gunn, et al. (1993), Chase-Lansdale, et al. (1997), and, for the UK, McCulloch and Joshi (2000).

None of these studies really attempts to assess the effect that spatial clustering of unobserved family or individual attributes has on measured neighbourhood effects. One way round the problem, given appropriate data is to estimate within-family models (Plotnick and Hoffman (1996), Aaronson (1998)). This approach removes unobservable, constant family effects by comparing siblings who grew up in different neighbourhoods. The first study finds no neighbourhood effects on incomes or post-secondary qualifications in the US, once allowance is made for family effects, but the authors do not look for impacts from neighbourhood educational composition. By contrast, Aaronson finds small but significant negative impacts from neighbourhood dropout rates on the probability of graduating from high school, both comparing across families and comparing siblings within families.

Technical discussion of the identification of neighbourhood effects is largely confined to the economics literature. Manski (1993) shows that endogenous neighbourhood effects, where the outcome of individuals is dependent on the average outcome in a local reference group, are not, in general, separately identifiable from dependence on unobserved on group characteristics. By contrast, Brock and Durlauf (2001) show that this type of social interaction effect is identified in a binary choice framework. This is because the mean group behaviour is a non-linear function of
individual and neighbourhood characteristics, so predicted group behaviour and individual characteristics are not linearly dependent in a probabilistic choice regression model. Nesheim (2002) also discusses parametric identification of neighbourhood effects in a model in which the empirical relationship between educational outcomes and local mean neighbourhood schooling attainments is determined by schooling as an input into human capital production, and by parental selection of residential neighbourhood. His approach requires estimation of parents' demand for schooling in a non-linear hedonic price function, and uses the non-linearities in this locational choice equation to provide instruments for neighbourhood quality. The approach adopted in our study is based on similar intuitions, but employs parental demand for property characteristics to predict variation in neighbourhood mean education levels which is exogenous to characteristics of residents in social housing.

Another strand in the literature looks to quasi-experimental evidence on the effect of neighbourhoods, using random re-assignments of families to new neighbourhoods. Evidence from Chicago's Gatreaux Assisted Housing programme indicates that moves from the city to the suburbs reduces drop out rates and improves college enrolment (Rosenbaum, et al. (1988); Rosenbaum (1991)). Using data on the Moving to Opportunity programme in Boston, Katz, et al. (2001) find short run treatment-on-the-treated effects on behaviour, health and well-being. Although the experimental approach is not subject to the same sources of bias as regression based estimates, it is not easy to identify causal factors or to generalise the results.
2.3 Models and empirical identification strategies

2.3.1 A simple model

This section develops a minimal linear specification for estimation of neighbourhood or community effects on education. Standard economic models of human capital development focus on child-level production functions of the type:

\[
h^c = h(h^P, h^n, z, \psi, t; \beta)
\]  

(2-1)

where \( h^P \) represents parent's own human capital, \( h^n = E[h^P | j] \) measures neighbourhood or community inputs from the adult human capital stock in area \( j \), \( z \) represents school-based inputs, \( \psi \) represents individual innate abilities, \( t \) represents time or effort spent in direct parental involvement, and \( \beta \) is a vector parameterising the partial derivatives. A child's school quality depends on the catchment area community inputs \( h^n \), possibly on parent's own financial inputs \( s \) (most importantly if the parents decide to send their child to a school in the private sector), and on other factors like aggregate school expenditures and teaching quality inputs \( \mu \).

\[
z = z(h^n, \mu, s; \gamma)
\]  

(2-2)

Assume that parents first make a choice of locality of residence – say a County or city, and its corresponding local education authority – according to labour market opportunities, returns to skills in the local labour market and other exogenous factors. They then decide in which neighbourhood to live according to physical property characteristics and local amenities, the quality of local schools and the educational status of neighbouring adults. Since admission to schools in the state sector is generally based on residential location, parents can only choose \( h^n \) and \( \mu \) simultaneously with choices over housing and other local amenities. Parents observe all these inputs, but can only vary the inputs to \( z \), other than own expenditure, by changing their spatial location. Imagine
that the basis of parental choice of residential location is a family-level utility function with local consumption goods \( q \), human capital attainments of the average child \( h^c \), non-spatially related numeraire consumption good \( c \) and family-specific preference parameters \( \theta \):

\[
U = U(h^c, q, c; \theta) \quad (2-3)
\]

or, substituting the observable inputs into human capital and schooling:

\[
V(h^n, \mu, s, t, q; \theta, \beta, \gamma, \nu) \quad (2-4)
\]

Neighbourhoods are repositories of three community goods: housing and environmental services \( q \), community educational capital \( h^n \) and local school performance \( z \). A location is completely described by \( (h^n, z, q) \) and hence by \( (h^n, \mu, q) \) if the parameters of the school production function are identical within localities. Neighbourhood property prices are described by a hedonic price function:

\[
P = P(h^n, \mu, q) \quad (2-5)
\]

where the implicit prices \( P_h, P_\mu, \) and \( P_q \) are constant across neighbourhoods (within localities). The budget constraint faced by parents with \( k \) children, in the decision on residential location and human capital investments is:

\[
w = c + P(h^n, \mu, q) + ks + wkt \quad (2-6)
\]

where \( w \) is the permanent income stream from lifetime income, or those components of income that are available to finance or guarantee loans for purchase of property or long term expenditures on a child's education. Expenditures \( c, P(h^n, \mu, q) \), \( s \) are permanent streams of lifetime expenditures, and \( t \) is the proportion of life spent attending to a child's education. Maximisation of (2-4) subject to (2-6) gives the optimal choices of the arguments of the utility function in terms of permanent income (and hence \( h^P \)), the implicit prices, family size and demand parameters:
A family chooses a residential location which jointly satisfies their demands for \( h^n, \mu, q \). The distribution of the demand function parameters (\( \Phi_x \)) across families will depend on the distribution of parental preferences (\( \theta \)), and their knowledge and expectations of the parameters of their children's human capital production function (\( \beta \)) and school production function (\( \gamma \)). Families with a stronger preference for their child's education, or for whom neighbourhood status is more productive (either directly or through school peer groups) will choose higher educational status neighbourhoods\(^4\).

Parental demands for location-based characteristics drive the sorting of individuals into neighbourhoods by income and preferences. Clearly, neighbourhoods that have concentrations of high quality housing stock, or have good local schools will be populated by high wealth, high human capital households, assuming these are normal goods and that capital markets are imperfect. Exogenous variation in characteristics of the neighbourhood which are normal goods generates sorting along educational lines, even if there are no benefits from living in a high education neighbourhood. Parents with high demands \( h^n, \mu, \) and \( q \) will populate neighbourhoods with high stocks of these factors, leading to high correlation between the preferences, incomes and education of neighbours.

Using a linearised empirical representation of a simple Cobb-Douglas production function, we have:

\[
\ln h_i^c = \beta_1 \ln h_i^P + \beta_2 \ln z_i + \beta_3 \ln t_i + \beta_5 \ln \psi_i + x_i \beta_6 + \epsilon_i \tag{2-8}
\]

\(^4\) An alternative interpretation in a dynamic setting is that the distribution of parameters reflects differences in the discount rate applied by parents to future dynastic earnings or children's human capital in the utility function.
where \( x_i \) is a vector of other locational characteristics. Even if we agreed that this was complete specification, consistent estimation of \( \beta_1 \) in the human capital production function is hindered by the lack of precise empirical counterparts to its inputs and because all the inputs are subject to parental choice. Most efforts at estimating a human capital production function like (2-8) implicitly exploit substitution of the unobserved factors by linear approximations to their demand functions, or otherwise assume that ad-hoc inclusion of controls is sufficient to guarantee that \( E[\varepsilon_i | h^n_i] = 0 \). Indeed, this is the first empirical strategy used in this Chapter.

Note though, that our identification problem – where we want to know the impact of the human capital stock of adults in the neighbourhood – is not quite as severe as in Manski’s reflection problem (Manski (1993, 2000)). The reflection problem expresses the difficulty in identifying the impact of group behaviour on individual behaviour, where the individual is part of the group. In this case any unobserved components in (2-8) would contribute directly to mean group human capital. In our case, we break the direct link between individual unobserved components and group-mean behaviour by specifying that the influential group and the individual are from separate cohorts.

2.3.2 Assumptions in alternative specifications

2.3.2.1 Community or area models

Gephart (1997) refers to versions of (2-8) without any individual or parental controls as community models. A pure community model of educational attainment might maintain that parental inputs have no effects which are independent of the community. A communitarian social philosophy would support this kind of model, where existing community values provide “authoritative horizons” which fix the goals that individuals pursue, and communities define individual identity (Kymlicka (1990)). Sampson and Byron Groves (1989), for example, discuss and test a community-level model based on
social disorganisation theory. Community or area-only models restrict $\beta_2$ to $\beta_4$ in (2-8) to zero. This is not a structural model in the economic mould, but is appropriate if we believe that parental characteristics are either irrelevant to a child's education, or are completely determined by the characteristics of their community – or if we simply want a description of the data. However, OLS on this equation obviously fails to provide a consistent estimate of $\beta_1$ in the structure of equation (2-8) unless $\beta_2$ and $\beta_3$ are all structurally zero\(^5\). This is because family inputs not included in the estimating equation are all correlated with $h^n$ through the common preference parameters and income variables in the demand functions (2-7). Also, if the demand for neighbourhood status depends on child's abilities, consistent estimation of $\beta_1$ requires $\beta_5 = 0$\(^6\).

A community-based approach does not necessarily rule out separate effects from schools and neighbourhoods. A first generalisation of the community model that avoids including any parental characteristics in the empirical production function – and consequent attenuation of the estimate of $\beta_1$ if $h^n$ is measured with error, or parental characteristics are structurally dependent on $h^n$ – removes the restriction on $\beta_2$ and allows effects from school quality. As before, consistent estimation of $\beta_1$ and $\beta_2$ by OLS requires that $\beta_3$ to $\beta_5$ are all structurally zero. An interpretation of this is that parental preferences, education, incomes and child abilities have an effect on educational

\(^5\) We may also require $\beta_4 = 0$, if parental time (or effort) spent on children's education is related to their own human capital. With a budget constraint like (2-6), the loss in income from time and effort spent on a child can cancel out any benefits, so time will be unrelated to income or own education.

\(^6\) Tests of these restrictions in an empirical may just show that our measures of community inputs are imprecisely measured, either because we are measuring the wrong things, measuring the right things badly, or because they are measured at inappropriate levels of geographical aggregation.
outcomes, but only via the demand for schooling. In this case the human capital and school production functions form a recursive structure.

2.3.2.2 Community, schooling and parental background

Neighbourhood models which allow for parental background or individual effects are called contextual in the sociological literature. In principle, these are just reduced form versions of (2-8). If we substitute parental characteristics for the demand for school quality, we get a reduced form human capital production function in neighbourhood and family background and individual characteristics only, with other area controls $x_i$:

$$\ln h_i^c = (\beta_1 + \beta_2 \gamma_1) \ln h_i^n + f_i \beta + x_i' \beta_6 + \omega_i$$ (2-9)

The vector $f_i$ must include individual abilities, parental education, number of children and proxies for permanent incomes and preferences. Estimation of (2-9) gives a consistent estimate of the sum of the effects of neighbourhood on human capital production operating through peer-group effects at school ($\beta_2 \gamma_1$) and directly through other channels ($\beta_1$).  

With data on a child's school quality, we can enter $z$ directly. If parental human capital, incomes and child abilities are measurable, we need only substitute for unobserved parental time or dedication to a child's schooling. We shall assume that this...

---

7 A complication arises if school funding is dependent on local taxes, or otherwise on local wealth and human capital, since now school performance depends directly on $h_i^n$ through peer group effects, but also indirectly via expenditures on the school. In this case, estimation of (2-9) does not identify neighbourhood human capital externalities separately from effects of neighbourhood human capital on school funding. We must include controls for school expenditures if these vary within localities, or control for school quality directly. In the British setting, funding formulae for state school expenditures ensure that they are almost constant within Local Education Authorities, so fixed effects at this level are sufficient.
depends on parental preferences for their child's education and the number of children in
the family that we include in $f_j$. The empirical production function is now:

$$\ln h_i^c = \beta_1 \ln h_i^n + \beta_2 \ln z_i + f_j \beta + x_i \beta_6 + \omega_{2j}$$  \hspace{1cm} (2-10)

This is a fairly restrictive specification. More realistically, the marginal product of
neighbourhood and schooling in the human capital production function could depend on
observable characteristics – ability as measured by early test scores, parental skills,
education or other demographics. We might also believe that the effects of
neighbourhood or community are mediated via family characteristics. For example,
children may only benefit from highly educated neighbours if their parents are educated
enough to engage socially with the community. In these cases we may prefer an empirical
specification with neighbourhood-family interaction terms.

2.3.3 School-quality selection

2.3.3.1 Checks using school characteristics and child abilities

Clearly, unless we treat a model such as (2-10) as a complete specification, or
otherwise assume that the unobservables are conditionally independent of neighbourhood
quality, then estimation by Ordinary Least Squares regression will not consistently
estimate the structural parameter of interest, $\beta_1$. Unobserved components of individual
ability that are observed by the parent, child or school will generate selection bias in the
estimates of $\beta_1$ and $\beta_2$. This occurs if the demand for neighbourhood status and school
quality, and the production of human capital, is dependent on unobserved ability.
However, for current purposes, all we really want is a consistent estimate of $\beta_1$.

It is straightforward to assess the extent to which school selection effects bias our
estimates of the key parameter of interest, $\beta_1$, by using data on school performance,
some observable school characteristics, and any measure of individual ability. The most
likely scenario is that there are ability selection effects on schooling, with high ability
children attending better schools, or high-motivation parents pushing hard for admission to good schools. In this case, $\beta_2$ will be an upward biased estimate of the structural effect of school quality. Positive correlation between $h^n$ and $z$ implies that OLS estimates of $\beta_1$ will be inconsistent. Even without selection, $\beta_1$ will be inconsistent if $z$ is a noisy measure of school quality, since the estimated neighbourhood parameter may pick up unobserved components of school quality. A simple check is to include additional school characteristics which are proxies for $\mu$ in (2-2). Change in the estimate of $\beta_1$ would suggest that unobserved school quality and selection effects are influencing the measured direct neighbourhood effect. A further check is available, since we can compare the estimate of $\beta_1$ in (2-10) with and without exclusion restrictions on parental preference variables or controls for child abilities. Although these are not rigorous tests, the range of variation in estimates of $\beta_1$ under different specifications can be informative.

2.3.3.2 Identification using exogenous local characteristics

The structure outlined above suggests another approach to identifying the neighbourhood parameter $\beta_1$, separately from unobserved school quality selection effects. The demand for housing services, or local amenities ($q$) implies an equilibrium relationship between average local wealth and the average quantity of $q$ in the neighbourhood. Since average local wealth depends on average local education or human capital, we can write mean neighbourhood human capital as a function of $q$. Mean neighbourhood human capital will also depend exogenously on the proportion of households in social housing, since (almost by definition) households in social housing have lower incomes and lower educational attainments on average than owner-occupiers and private tenants. The neighbourhood mean human capital generating function is:
\[ \ln h_i^* = h^*(q, \pi, \xi_i) \]  

(2-11)

Estimation of (2-10) by Instrumental Variables, using neighbourhood property characteristics \( q \) and the proportion of social housing \( \pi \) as instruments, gives a consistent estimate of \( \beta_1 \) even if there are unobserved school quality components \( \mu \), under the assumption that \( E[\mu|q] = 0 \) and \( E[\mu|\pi] = 0 \). This requires that unobserved school quality factors do not depend on local property characteristics or the proportion of social housing, conditional on localities defined by \( x \) and other controls in (2-10).

2.3.4 Social tenants and parental selection

Focussing on neighbourhood as a factor in the human capital production function, we can re-write our human capital production function as:

\[ \ln h_i^c = \beta_1 \ln h_f^n + x_{if} \alpha + \eta_f + \varepsilon_i \]  

(2-12)

where \( h_i^c \) is the attainment of a child \( i \) in family \( f \), \( x_{if} \) is a vector of observed individual and family characteristics, \( \eta_f \) is an unobserved family specific effect, and \( \varepsilon_i \) is an individual specific error term. As discussed in Section 2.3.1, the potential problem in estimation of (2-12) is that the neighbourhood characteristic \( h_f^n \) is correlated with unobserved or badly measured characteristics of the families under observation. This correlation arises principally through the demand for community status, school quality, property characteristics, environmental characteristics and other local amenities, which leads to a dispersion of land rents and property prices across geographical space. This dispersion in land rents and property prices generates dispersion in family incomes, wealth and hence education across neighbourhoods.

We are interested here in the relationship between neighbourhood educational composition and the attainments of children. The relationship between the education of a parent in the sample, and the mean education in the neighbourhood can be written:
The regression coefficient in the regression of parental \( h_f^P \) on neighbourhood mean is one\(^8\). Running an OLS regression on a cross-section in (2-12) gives unbiased estimates of the parameter \( \beta_1 \), only if the parental characteristics are measured without error, and the unobserved family effects are uncorrelated with the neighbourhood measure or with other family characteristics. This is true even if households are randomly assigned to neighbourhoods.

However, assume we have two samples of individuals, the first group \( s \) randomly allocated to neighbourhoods, the second group \( o \) systematically sorted into neighbourhoods by education. We have the overall mean education in each neighbourhood, after assignment. From this we could infer the effect of neighbourhood composition on the randomly allocated group by regression of our outcome variable on neighbourhood composition without any parental controls. To formalise this, assume we have in each neighbourhood, proportion \( \pi \) families, who sort themselves into neighbourhoods on the basis of individual demands for some neighbourhood amenity \( q \). Assume we have another group of \((1-\pi)\) families who are allocated to neighbourhoods in a random way, or in such a way that their characteristics are uncorrelated with the neighbourhood amenity. The sorting processes relating individual characteristics to the neighbourhood can be written as:

\[
h_f^s = \mu^s + \zeta_f^s, \quad h_f^o = \mu^o + \phi q + \zeta_f^o
\]  

(2-14)

\(^8\) The neighbourhood mean is \( h_f^p = \frac{1}{J} \sum_{f=1}^{J} h_f \). The correlation coefficient between parental education and neighbourhood mean education is \( \rho_h = \frac{1}{\sigma^2_n(\sigma_n^2 + \sigma^2_{\omega})^{-1}} \).
\[ \eta_f^s = \mu^p + \xi_f^s, \quad \eta_f^o = \mu^p + \psi q^n + \xi_f^o \]  

The \( \mu \) are constants and \( \phi \) and \( \psi \) are parameters. The expected value of education in any neighbourhood \( n \) is then, from (2-14):

\[ E[h_f | n] = \pi \mu h_0 + \pi \phi q^n + (1 - \pi) \mu^{so} \]  

(2-16)

The covariance of the family and mean neighbourhood education for each group is:

\[ \text{Cov}(h^s, h^n) = 0, \quad \text{Cov}(h^o, h^n) = \pi \phi^2 \cdot \text{Var}(q^n) \]  

(2-17)

and the covariance of the unobserved family characteristic and the neighbourhood measure is:

\[ \text{Cov}(\eta^s, h^n) = 0, \quad \text{Cov}(\eta^o, h^n) = \pi \psi \phi \cdot \text{Var}(q^n) \]  

(2-18)

Under these assumptions, we can identify the effect of \( h^n \) on individuals in the \( s \) group without controlling for \( h_f^s \). If the sample used to construct the neighbourhood mean is sufficiently large, then the sample mean of \( h_f^s \) is the same in all neighbourhoods, so does not contribute to variation in \( h^n \) between neighbourhoods. This may be a strong assumption, since any sampling variation will mean the first conditions in (2-17) and (2-18) will not hold exactly. However, provided \( \text{Cov}(h^o, h^n) \) is large, so \( \text{Var}(h^n) \) is large, the regression coefficient derived from \( \text{Cov}(h^s, h^n)/\text{Var}(h^n) \), or \( \text{Cov}(\eta^s, h^n)/\text{Var}(h^n) \) will be near zero, so the bias in the OLS estimate of \( \beta_1 \) is negligible. Moreover, from (2-15), if we observe neighbourhood characteristics or local housing characteristics \( q^n \) which have value for group \( o \) only, we can use these as instruments for \( h^n = E[h_f | n] \) in an equation (2-12) estimated on group \( s \) only.

It is a plausible, though contentious, assumption that council tenants are randomly allocated to their neighbourhood. This will be true to the extent that the neighbourhood location of an allocated council home is largely unrelated to the resources and preferences
of the tenant. Under this assumption, council tenants provide a suitable group for the strategy described above. If a link between council tenants education and that of their neighbours' is attributable to a council policy of matching tenants, or parents' desire to be housed amongst similar tenants, then the characteristics of owner-occupiers in the neighbourhood will provide good instruments for the neighbourhood status of council tenants.

2.3.5 Local non-linearities in the neighbourhood effect

Local non-linearities in the relationship between neighbourhood and educational outcomes have implications for the long-run impact of neighbourhood on intergenerational mobility and inequality. A linear relationship with moderate slope suggests that the neighbourhood-educational process is mean reverting. If there are local non-linearities in the relationship, such as the threshold or contagion effects highlighted in Crane (1991), then there may be non-linearities in the familial intergenerational relationship. This can mean that neighbourhoods or dynasties from one end of the distribution of neighbourhood quality, may remain permanently separated from those at the other end (see Loury (1977); Benabou (1996) for example). A generalisation of the empirical specification (2-10) that allows for general nonlinearities in \( h_f^e \) is:

\[
    h_f^e = g\left(h_f^n\right) + \beta' f_t + \omega_t
\]

Parameter vector \( \beta \) can be estimated by semi-parametric methods, such as the partial linear model (Robinson (1988)). An estimate of the function \( g\left(h_f^n\right) \) is then obtained by a second stage kernel regression of \( h_f^e - \hat{\beta}' f_t \) on \( h_f^n \), at preset sequence of points \( \{h_f^n\} \). Details are in Appendix E.
2.4 Description of the data

2.4.1 The NCDS and Census data sets

Our empirical methods use the framework described in Section 2.3 to investigate the effects of mean neighbourhood education levels on educational attainments. The data is the British National Child Development Study (NCDS) This study follows the British cohort born between 3rd and 9th of March 1958, with follow-up surveys at age 7 (1965), age 11 (1969), age 16 (1974), age 23 (1981), and age 33 (1991). A further sweep is underway for 2001. The NCDS has been exploited in innumerable research papers in many disciplines. Census Ward identifiers for cohort member’s residential addresses are available for 1974 (and 1981), allowing us to match in British Census data from 1971 to 1974. Some Census data is already included in the NCDS files, though for 1974 this is all at Enumeration District (ED) and Local Authority (LA) level, and does not include neighbourhood education. Measures of neighbourhood education and other characteristics at the intermediate Ward level must be spliced in from the publicly available Ward-level Census statistics.

The ED is the smallest unit for which Census statistics are available. It is the ‘input’ geographical unit of the Census. Around 10 EDs make up a Census Ward. In 1971 there were around 18000 wards, with mean populations of around 3000. The distribution is, however, highly skewed with 50% of wards containing less than 1000 persons. One quarter have more than 4000 residents. The sample of Wards matched to the NCDS sample over-represents higher population Wards (unsurprisingly given the distribution of births), with median populations of about 8000. The number of wards represented in the base NCDS sample used in this study is 5479. Enumeration Districts are intended to encompass as many households as the Census Enumerator could cover on the Census day. Whilst they are probably ideal as a neighbourhood measure, some of the most useful
Census data is based on a 10% sample, so ED sample sizes are extremely low\textsuperscript{9}. We will use mostly Ward-level neighbourhood data. An exception here is an indicator of residence on or near a Local Authority estate, constructed from the ED-level data in the NCDS. Labour market controls are taken at Local Authority level, plus dummy variables for up to 64 counties of residence in 1974.

2.4.2 Description of the samples

Estimation of the models is based on samples of men and women from the 1991 NCDS sweep at age 33 who reported their highest educational qualifications. Parental characteristics for these adults come from the earlier childhood sweeps, along with information on their early abilities (age-7 test scores) and the performance and characteristics of the secondary school they attended at age 16. Attrition is a problem in the NCDS. Where possible, parental characteristics are those measured in 1974, but the latest information available from earlier sweeps replaces missing data to maintain sample sizes.\textsuperscript{10} All observations without Ward identifiers in 1974, or without adult qualification data are dropped. This gives us a maximal sample size of 4538 men and 4835 women. The sample size of each is reduced by around 900 once we drop observations without secondary school quality measures. A similar base sample is used to look at the earlier attainments of the NCDS cohort, focusing on the results of reading and maths tests carried out at age 16.

\textsuperscript{9} The issue of the correct choice of neighbourhood group arises frequently in the literature. Overman (2002) finds differing effects on high school drop out rates from occupational status operating at small and large neighbourhood definitions. Since we have no choice, we use wards. Experimenting on proxies for education, e.g. professional workers, at ED and Ward level shows that using either Ward or ED, unconditional on the other, gives similar results. Apparently, the choice is not so critical.

\textsuperscript{10} This never amounts to more than 5% of the responses for which residential address information is available.
2.4.3 Description of the variables

All the variables used in the results are defined in Appendix A. Our key regressor is the empirical counterpart to neighbourhood human capital $h^a$. Potential choices are social class variables, or the educational attainments of adults in the neighbourhood, or some composite of the two. We will focus on educational status only, since this seems the most natural choice in models of educational outcomes. The variable is the proportion of over-18s in each Ward with Ordinary National Diplomas, A-levels, Higher National Diplomas, degrees, higher degrees, professional qualifications or equivalent qualifications. Using this, we can readily evaluate the impact of the proportion of adult neighbours with these qualifications on the probability of a child attaining the same qualifications. Models of influences on teenage abilities use the natural log of the age-16 test scores as the dependent variable.

2.5 Empirical results and discussion

2.5.1 Community and area models

Table 2-1 shows marginal effects and t-statistics based on ordered probit estimation of the community model of Section 2.3.2.1, for boys only. The marginal effects show the percentage-point change in the probability of a child being in an educational outcome category occurring with each one percentage-point change in the explanatory variable. The Low/High categories are described in Appendix A. Control variables are a selection of key economic, community and geographical characteristics measured at Ward or Local Authority level, with County dummies. The results in Row 1 show what happens to children from similar neighbourhoods that differ in respect of the qualifications of residents. Looking at the results in Column (1), Table 2-1, it is clear that better educated men originated from high-education neighbourhoods.
Better neighbourhood education is the most important of the factors identified here: a one percentage point increase in the proportion of neighbours with A-levels or higher is associated with a 0.67% increase in the probability of a man having at least A-levels by age 33, and a 0.50% decrease in the probability of failing to gain any decent qualifications. These translate into elasticities at the sample mean of 0.23 and −0.31 respectively. Boys from the top decile of educated neighbourhoods (21.7% with A-levels+) were twelve percentage points more likely than those from the bottom decile (3.7%) to end up at age 33 with high qualifications (41% against 27%), but were nine percentage points less likely to end up with no or few qualifications (15% against 24%).

Table 2-1: Area-only models of men’s, age 33 qualifications

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>Mean</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Low</td>
<td>High</td>
<td>t</td>
</tr>
<tr>
<td>Highly qualified</td>
<td>-0.497</td>
<td>0.673</td>
<td>6.102</td>
</tr>
<tr>
<td>School quality</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Unskilled workers</td>
<td>0.259</td>
<td>-0.349</td>
<td>1.734</td>
</tr>
<tr>
<td>Unemployment rate</td>
<td>0.599</td>
<td>-0.809</td>
<td>2.164</td>
</tr>
<tr>
<td>Economically active men</td>
<td>-0.023</td>
<td>0.032</td>
<td>0.253</td>
</tr>
<tr>
<td>Economically active women</td>
<td>-0.038</td>
<td>0.051</td>
<td>0.749</td>
</tr>
<tr>
<td>One year migrants</td>
<td>-0.010</td>
<td>0.013</td>
<td>0.083</td>
</tr>
<tr>
<td>New com. immigrant</td>
<td>-0.045</td>
<td>0.061</td>
<td>0.619</td>
</tr>
<tr>
<td>Average dwelling size</td>
<td>-0.052</td>
<td>0.070</td>
<td>3.444</td>
</tr>
<tr>
<td>Lacking amenities</td>
<td>0.435</td>
<td>-0.588</td>
<td>1.808</td>
</tr>
<tr>
<td>Social housing</td>
<td>0.073</td>
<td>-0.099</td>
<td>2.516</td>
</tr>
<tr>
<td>City Ward</td>
<td>-0.034</td>
<td>0.046</td>
<td>2.552</td>
</tr>
<tr>
<td>High population</td>
<td>0.001</td>
<td>-0.002</td>
<td>0.112</td>
</tr>
<tr>
<td>Agricultural employment</td>
<td>0.237</td>
<td>-0.320</td>
<td>2.302</td>
</tr>
<tr>
<td>Mining, manufacturing emp.</td>
<td>-0.060</td>
<td>0.081</td>
<td>1.053</td>
</tr>
</tbody>
</table>

County effects: \( \chi^2_{60} = 75.17, P=0.090 \) \( \chi^2_{60} = 104.25, P=0.00 \)

Predicted group probability: 0.187, 0.338, 0.179, 0.343

Log likelihood: -4504.80, -3375.03

Pseudo R²: 0.044, 0.097

Sample size: 4538, 3637

Three category ordered probit estimates.

Including Ward professional and managerial employees reduces parameter on Ward education to –0.223/+0.299 (t=2.214) in Column (2), but is not itself highly significant (t=1.54).

Neighbourhood educational measure is proportion of adults in high qualification category in cohort member’s age-13 residential Ward. School quality is proportion of boys age 15 studying for GCEs at cohort members’ school.
Comparable results for girls are shown in Appendix B, Table 2-11. Most of what was said for boys in terms of the effect of their origin in the distribution of neighbourhoods applies to girls.

We get similar results if we replace the dependent variable with cumulative time in education. The elasticities on Ward education come out at 0.26 for men and women \((t = 6.5)\). The within-County \(R^2\)'s from these regressions imply that a maximum of 8% of the variation in time in education is associated with these neighbourhood attributes. Much of this will be attributable to parental background, not community effects. The explanatory role of neighbourhood in educational outcome would appear to be small, relative to other factors.

Even if we reject the community-only model of attainments, these results are interesting since they highlight the important associations between area characteristics and adult educational outcomes. *A priori*, we would expect fairly strong associations unconditional on parents' characteristics, because neighbourhood characteristics reflect the characteristics of the individuals' parents. So, although the results do not imply causality, they are useful from a policy point of view if what we want is area basis for targeting resources to the educationally disadvantaged.

Our second set of results, Column (2) in Table 2-1, repeats the analysis holding constant the quality of local schooling – specifically the proportion of own-sex 15 year olds studying for GCE O-Levels at the school attended by the child at age 16. Once we take account of the quality of secondary school attended, the effect of neighbourhood education falls by over one third to give elasticities of 0.14 on A-levels and -0.21 on low qualifications. The coefficient is still highly significant, and substantial considering we are only capturing effects over and above anything influencing the measured quality of the child's schooling at age 16. Even conditional on this measure of secondary school quality, teenagers in the top educational decile of neighbourhoods are 7.5 percentage
points more likely to end up with high qualifications than those at the bottom decile, and 5.6 percentage points less likely to end up with the lowest qualifications. Results for women in Appendix B, Table 2-11, show a similar pattern.

Comparable within-County regressions for time in education tell us that around 20% of the variation in education of men, and 21% for women, is attributable to secondary school quality and neighbourhood together. School quality alone accounts for about two-thirds of this, leaving components of neighbourhood unrelated to school quality to explain around 6% of the variation in time in education.

2.5.2 Community, area and parental background

2.5.2.1 Conditional or contextual effects

Table 2-2, changes the analysis to look at the impact on boys of our measure of neighbours’ educational qualifications, conditional on the characteristics of the cohort members’ parents and family. The choice of family characteristics included is in line with the basic human capital production function model presented in equation (2-9), and might be called contextual. The specification includes additional geographic controls for urban and high population wards, and high social housing Enumeration Districts, alongside local labour market measures and sixty County dummy variables. Other neighbour characteristics are excluded\(^{11}\) to focus on comparable outcome and neighbourhood issues.

The marginal effect of neighbourhood in Column (1), conditional on parental characteristics, is substantially lower than in Table 2-1. Most of the coefficients on parental characteristics are significant. Early test scores are highly significant and powerful predictors of attainments, as is well known from other studies e.g. Feinstein and Symons (1999). Area characteristics, other than the County dummies and local

---

\(^{11}\) Proportion of in professional occupations and average dwelling size attract negative statistically insignificant coefficients as additional regressors in qualifications or educational-years models.
unemployment rates have little effect. The coefficient on neighbourhood educational status corresponds to elasticities of +0.095/-0.13 on the high and low qualification categories – evaluated at mean neighbourhood education. A boy from the top decile of neighbourhoods is around five percentage points more likely to qualify with A-levels or higher than a boy from a similar family in the bottom decile. Another way of looking at this is that a move from the bottom decile to the top decile of neighbourhoods has an effect on educational attainments of similar magnitude to an extra 2 years of parental education.

Repeating the analysis on time in education gives an elasticity of 0.095 at mean non-compulsory years of education (the $t$ statistic is 3.66). The partial $R^2$ for our Ward human capital variable is 0.0024, against an overall within-County $R^2$ of 0.3008. It is worth noting that, by this calculation, only 0.8% of the variation of time in education attributable to childhood background is explained by teenage neighbourhood educational status!

2.5.2.2 Investigating school selection

As discussed in Section 2.3.3, the measured neighbourhood effect could be entirely attributable to school selection by parents and children, where the factors that drive selection are unobserved, and school quality has an effect on individual achievements. If these attributes are correlated with parental education they will also be correlated with the education of neighbours. Selection on schools by parents and children of unobservably different types (or selection on parents and unobserved pupil ability by schools) will lead to inconsistent estimates.
Table 2-2: Neighbourhood education and family effects on men's age-33 qualifications

<table>
<thead>
<tr>
<th>Qualification group:</th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Ward education</td>
<td>0.200</td>
<td>0.278</td>
<td>0.154</td>
<td>0.117</td>
<td></td>
<td></td>
</tr>
<tr>
<td>School quality</td>
<td>-</td>
<td>-</td>
<td>0.008</td>
<td>0.116</td>
<td></td>
<td></td>
</tr>
<tr>
<td>School type:</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Grammar</td>
<td>-</td>
<td>-</td>
<td>-0.080</td>
<td>0.116</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Secondary modern</td>
<td>0.007</td>
<td>-0.001</td>
<td>0.66</td>
<td>0.211</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Independent</td>
<td>-</td>
<td>-</td>
<td>-0.060</td>
<td>0.040</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Grant maintained</td>
<td>-</td>
<td>-</td>
<td>-0.088</td>
<td>0.029</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Other non-LEA</td>
<td>-</td>
<td>-</td>
<td>0.029</td>
<td>0.024</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mother education</td>
<td>-0.015</td>
<td>0.021</td>
<td>0.310</td>
<td>15.06</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Father education</td>
<td>-0.017</td>
<td>0.024</td>
<td>0.522</td>
<td>15.21</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Father's age</td>
<td>-0.001</td>
<td>0.002</td>
<td>0.169</td>
<td>30.54</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dwelling size</td>
<td>-0.009</td>
<td>0.012</td>
<td>0.137</td>
<td>4.978</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tenure:</td>
<td>-0.088</td>
<td>0.011</td>
<td>0.239</td>
<td>0.050</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tenure missing</td>
<td>0.082</td>
<td>-0.113</td>
<td>2.65</td>
<td>0.020</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Council tenant</td>
<td>0.058</td>
<td>-0.080</td>
<td>5.37</td>
<td>0.366</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Private rental</td>
<td>0.023</td>
<td>-0.032</td>
<td>1.23</td>
<td>0.050</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Other tenure</td>
<td>0.052</td>
<td>-0.072</td>
<td>2.17</td>
<td>0.036</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Parental interest</td>
<td>0.046</td>
<td>-0.063</td>
<td>5.15</td>
<td>1.042</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Log siblings</td>
<td>0.046</td>
<td>-0.063</td>
<td>5.15</td>
<td>1.042</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age-7 test scores</td>
<td>-0.627</td>
<td>0.868</td>
<td>20.49</td>
<td>0.530</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tests missing</td>
<td>-0.415</td>
<td>0.575</td>
<td>16.48</td>
<td>0.112</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Social housing</td>
<td>-0.010</td>
<td>0.014</td>
<td>0.96</td>
<td>0.384</td>
<td></td>
<td></td>
</tr>
<tr>
<td>City Ward</td>
<td>-0.014</td>
<td>0.090</td>
<td>1.14</td>
<td>0.410</td>
<td></td>
<td></td>
</tr>
<tr>
<td>High population</td>
<td>-0.099</td>
<td>0.013</td>
<td>0.78</td>
<td>0.249</td>
<td></td>
<td></td>
</tr>
<tr>
<td>LA agriculture</td>
<td>0.102</td>
<td>-0.142</td>
<td>1.24</td>
<td>0.027</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Unemployment</td>
<td>0.702</td>
<td>-0.973</td>
<td>1.96</td>
<td>0.040</td>
<td></td>
<td></td>
</tr>
<tr>
<td>County effects</td>
<td>0.178</td>
<td>0.343</td>
<td>0.343</td>
<td>0.342</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>-2962.24</td>
<td>-2920.75</td>
<td>-2908.15</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pseudo R²</td>
<td>0.207</td>
<td>0.218</td>
<td>0.222</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sample size</td>
<td>3637</td>
<td>3637</td>
<td>3637</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Three category ordered probit estimates.

Neighbourhood educational measure is proportion of adults in high qualification category in cohort member's age-13 residential Ward. School performance is proportion of boys age 15 studying for GCEs.

Estimation of model of first column using full available sample of 4538 men gives marginal effect from neighbourhood education of -0.196 on the low category and 0.310 on the high category (t = 4.306, p-value = 0.0000). The coefficient is not significantly different from estimate on the smaller subsample for which school performance is observable.

Replacing County dummies with 1974 LEA dummies gives marginal effect -0.184/0.24 (t=3.149, P=0.002) in Column (3).
If we believe the simple model of Section 2.3.2.1, and we trust that the data provides good measures of parental characteristics, then selection on school quality is not an issue. Of course, this strategy alone is unconvincing. Hence, Columns (2) and (3) indirectly test the assumption that there are no omitted variables which are correlated with neighbourhood education levels and with the quality of local schooling. Firstly Column (2) includes the measure of local schooling quality as a regressor. We expect a fall in the neighbourhood effect, because part of the influence of the educational status of the neighbourhood will operate through peer group effects in school\(^{12}\). In fact, the size of the estimated impact of neighbourhood is reduced by only around 20%, and remains significant at the 1% level. Around 20% of the neighbourhood effect could be attributable to selection on schooling or school peer-group effects. This finding of the importance of neighbourhood factors over and above secondary schooling is in line with Gamer and Raudenbush (1991).

Certainly, the proportion of boys studying for GCE-O levels is a crude quality measure – though not unlike the performance measures used in the school league tables of the last decade. Our neighbourhood effect could be picking up residual, unobserved characteristics of local secondary schooling. Secondly, if there is selection bias on school quality then the estimate of entire parameter vector is inconsistent. To assess how far this affects our main parameter of interest, Column (3) adds in dummy variables for school type. We would expect attenuation of the coefficients on any variables that are correlated with unobserved differences in school performance across school type categories, and so the coefficient on school quality is nearly halved; yet the estimated neighbourhood effect is almost unchanged. We interpret this as showing that there are no important unobserved components of school quality that affect adult attainments and are correlated with

\(^{12}\) Column A estimates \((\beta_1 + \beta_2 \gamma_1)\) in equation (2-9), whereas Column (2) estimates \((\beta_1)\).
neighbourhood education level. Selection on school characteristics does not bias the estimate of the influence of neighbourhood status. Similar results for girls are shown in Appendix B.

2.5.2.3 Robustness to changes in specification

Table 2-3, Rows 1-3 show that estimates are not unduly sensitive to the inclusion of additional paternal and maternal characteristics – income, social class and socioeconomic group, where these are available. The first three Columns, Rows 1-3 show estimated marginal effects and t-statistics for our neighbourhood effect, using the models of Table 2-2, Column (1) applied to the sample available with observations on these additional family characteristics. The next three columns show the estimates when the additional controls are included.

<table>
<thead>
<tr>
<th>Additional control variables</th>
<th>Without additional controls</th>
<th>With controls</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Low</td>
<td>High</td>
</tr>
<tr>
<td>Family income dummies¹</td>
<td>-0.169</td>
<td>0.237</td>
</tr>
<tr>
<td>χ²₃₁=1502 P=0.000, N = 2525</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Father and mother’s social class¹</td>
<td>-0.184</td>
<td>0.256</td>
</tr>
<tr>
<td>χ²₂₃=35.65 P=0.001, N = 3345</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Father &amp; mother’s SEG²</td>
<td>-0.171</td>
<td>0.230</td>
</tr>
<tr>
<td>χ²₃₁=1338.1 P=0.000, N=3392</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Parental interest dummies²</td>
<td>-0.208</td>
<td>0.278</td>
</tr>
<tr>
<td>χ²₁₁=141.8 P=0.000, N=3637</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Early attainments (mean of age 7 tests)²</td>
<td>-0.189</td>
<td>0.246</td>
</tr>
<tr>
<td>Reading ability at age 11</td>
<td>-0.166</td>
<td>0.223</td>
</tr>
<tr>
<td>t=15.406</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Three category ordered probit estimates.
1. Other regressors as in Table 2-2, Column (1)
2. Other regressors as in Table 2-2, Column (2)

The estimates on the neighbourhood effect coefficients are statistically comparable to those in the original models, and are significant to at least the 5% level. In the worst case, including 32 socioeconomic group dummies attenuates the coefficient by over 25%,
but then current parental employment skill group is not predetermined and may itself depend on past neighbourhood human capital formation.

Row 4 checks the sensitivity to exclusion of parental interest dummies, once we control for the quality of school attended as in Table 2-2, Column (2). Again, parental interest dummies might proxy neighbourhood rather than family effects: a community where some parents are visibly supportive of their children’s schooling may encourage this behaviour in others. They are, however, useful indicators of parental qualities that motivate selection on higher-education neighbourhoods. Those parents with high interest in their child’s education will be those that seek out the returns to good neighbourhoods. Nevertheless, removing these controls from the original specification pushes the coefficient on neighbourhood up by less than 25%. Similarly, Row 5 shows the impact from removing the age-7 academic ability controls from the model with parental background and schooling. This increases the coefficient by only 10%. Adding in an age-11 reading test score decreases the coefficient by around 10%. Given these figures, it is unlikely that selection on unobserved individual abilities accounts for a substantial proportion of the estimated neighbourhood effect. Section 2.5.4 investigates these issues further.

2.5.2.4 Heterogeneity in returns

We find no evidence that the relationship between neighbourhood and educational attainments differs across ability groups. Interaction terms between our neighbourhood education measure and age-7 ability quartile dummies were insignificant. The same is true for interactions between school performance and ability. On the basis of this evidence it seems that concerns about ability selection effects, or complementarities between neighbourhood and ability are unfounded, unless selection is on components of ability that are unrelated to early test scores.
What we do find though is that interactions between neighbourhood and parental interest categories, and interactions between school quality and parental interest categories are important\(^\text{13}\). This is evidence that parental interest affects the returns to neighbourhood or school performance in the human capital production function. Alternatively, the returns might influence parental interest – these are observationally equivalent here. Allowing for this heterogeneity, the marginal effect of neighbourhood in the modal group is \(-0.27/0.37\), but ranges from \(+0.45/-0.20\) for those who expressed little interest and did not read to their child, up to \(-0.75/1.01\). In contrast, there seems to be little heterogeneity in the returns to neighbourhood or school quality across parental education groups. Interaction terms were generally positive – implying that children of better educated parents benefit more – but not statistically significant. This is in line with the non-parametric estimates of the relationship between adult educational attainments and parental-neighbourhood education, presented later in Section 2.6.1.2.

### 2.5.3 Predicting neighbourhood status from housing type

Section 2.3.3.2 suggested that property characteristics and the proportion in social housing provide suitable instruments for neighbourhood status, if we want to purge our estimates of school selection effects. These neighbourhood characteristics are unchanged by parental selection on school quality, at least in the short run.

Table 2-4 shows 2-step Instrumental Variables estimates with a first stage linear regression and second stage ordered probit as in Table 2-2. Instruments are the Ward-proportions of Local Authority tenants, and property size\(^\text{14}\).

---

\(^{13}\) The interaction terms are significant for both neighbourhood (\(\chi^2_{11} = 19.7\) p-value = 0.05) and school performance (\(\chi^2_{11} = 22.8\), p-value = 0.02).

\(^{14}\) Proportion with more than seven rooms.
## Table 2-4: Neighbourhood education and family effects on men – 2-step estimates

<table>
<thead>
<tr>
<th>Qualification group:</th>
<th>Low</th>
<th>High</th>
<th></th>
<th>Low</th>
<th>High</th>
<th></th>
<th>Low</th>
<th>High</th>
<th></th>
<th>Mean</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Ward education</strong></td>
<td>-0.242</td>
<td>0.336</td>
<td>2.48</td>
<td>-0.198</td>
<td>0.266</td>
<td>2.01</td>
<td>-0.219</td>
<td>0.289</td>
<td>2.17</td>
<td>0.117</td>
</tr>
<tr>
<td><strong>School quality</strong></td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-1.37</td>
<td>0.184</td>
<td>9.05</td>
<td>-0.074</td>
<td>0.098</td>
<td>3.53</td>
<td>0.278</td>
</tr>
<tr>
<td><strong>School type:</strong></td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Grammar</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-0.080</td>
<td>0.106</td>
<td>4.25</td>
<td>0.116</td>
</tr>
<tr>
<td>Secondary modern</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>0.008</td>
<td>-0.010</td>
<td>0.67</td>
<td>0.211</td>
</tr>
<tr>
<td>Independent</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-0.061</td>
<td>0.081</td>
<td>2.14</td>
<td>0.040</td>
</tr>
<tr>
<td>Grant maintained</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-0.087</td>
<td>0.115</td>
<td>2.61</td>
<td>0.029</td>
</tr>
<tr>
<td>Other non-LEA</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>0.029</td>
<td>-0.039</td>
<td>1.10</td>
<td>0.024</td>
</tr>
<tr>
<td>Mother education</td>
<td>-0.015</td>
<td>0.021</td>
<td>4.61</td>
<td>-0.013</td>
<td>0.018</td>
<td>3.95</td>
<td>-0.012</td>
<td>0.016</td>
<td>3.66</td>
<td>15.06</td>
</tr>
<tr>
<td>Father education</td>
<td>-0.017</td>
<td>0.025</td>
<td>6.39</td>
<td>-0.015</td>
<td>0.021</td>
<td>5.55</td>
<td>-0.015</td>
<td>0.020</td>
<td>5.42</td>
<td>15.21</td>
</tr>
<tr>
<td>Father’s age</td>
<td>-0.001</td>
<td>0.002</td>
<td>1.91</td>
<td>-0.001</td>
<td>0.005</td>
<td>1.69</td>
<td>-0.001</td>
<td>0.001</td>
<td>1.65</td>
<td>30.54</td>
</tr>
<tr>
<td>Dwelling size</td>
<td>-0.008</td>
<td>0.017</td>
<td>2.17</td>
<td>-0.005</td>
<td>0.007</td>
<td>1.29</td>
<td>-0.004</td>
<td>0.006</td>
<td>1.05</td>
<td>4.978</td>
</tr>
<tr>
<td><strong>Tenure:</strong></td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Tenure missing</td>
<td>0.082</td>
<td>-0.114</td>
<td>3.07</td>
<td>0.083</td>
<td>-0.111</td>
<td>3.01</td>
<td>0.088</td>
<td>-0.117</td>
<td>3.15</td>
<td>0.020</td>
</tr>
<tr>
<td>Council tenant</td>
<td>0.058</td>
<td>-0.080</td>
<td>5.05</td>
<td>0.060</td>
<td>-0.080</td>
<td>5.21</td>
<td>0.058</td>
<td>-0.077</td>
<td>5.04</td>
<td>0.366</td>
</tr>
<tr>
<td>Private rental</td>
<td>0.023</td>
<td>-0.032</td>
<td>1.13</td>
<td>0.022</td>
<td>-0.029</td>
<td>1.05</td>
<td>0.018</td>
<td>-0.023</td>
<td>0.86</td>
<td>0.050</td>
</tr>
<tr>
<td>Other tenure</td>
<td>0.051</td>
<td>-0.070</td>
<td>2.28</td>
<td>0.049</td>
<td>-0.066</td>
<td>2.22</td>
<td>0.049</td>
<td>-0.065</td>
<td>2.21</td>
<td>0.036</td>
</tr>
<tr>
<td>Parental interest</td>
<td>0.169</td>
<td>43.45</td>
<td>-0.000</td>
<td>0.145</td>
<td>43.48</td>
<td>-0.000</td>
<td>0.139</td>
<td>39.46</td>
<td>-0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>Family size (In kids)</td>
<td>0.045</td>
<td>-0.063</td>
<td>5.03</td>
<td>0.039</td>
<td>-0.052</td>
<td>4.25</td>
<td>0.039</td>
<td>-0.052</td>
<td>4.25</td>
<td>1.042</td>
</tr>
<tr>
<td><strong>Age-7 test scores</strong></td>
<td>-0.626</td>
<td>0.868</td>
<td>22.03</td>
<td>-0.577</td>
<td>0.77</td>
<td>19.88</td>
<td>-0.558</td>
<td>0.74</td>
<td>18.77</td>
<td>0.530</td>
</tr>
<tr>
<td><strong>Age-7 tests missing</strong></td>
<td>-0.415</td>
<td>0.576</td>
<td>18.19</td>
<td>-3.82</td>
<td>0.51</td>
<td>16.38</td>
<td>-0.372</td>
<td>0.49</td>
<td>15.69</td>
<td>0.112</td>
</tr>
<tr>
<td><strong>LA estate</strong></td>
<td>-0.012</td>
<td>0.016</td>
<td>1.04</td>
<td>-0.014</td>
<td>0.019</td>
<td>1.22</td>
<td>-0.014</td>
<td>0.020</td>
<td>1.26</td>
<td>0.384</td>
</tr>
<tr>
<td><strong>City Ward</strong></td>
<td>-0.014</td>
<td>0.019</td>
<td>1.21</td>
<td>-0.015</td>
<td>0.021</td>
<td>1.32</td>
<td>-0.015</td>
<td>0.019</td>
<td>1.26</td>
<td>0.410</td>
</tr>
<tr>
<td><strong>High population</strong></td>
<td>-0.008</td>
<td>0.012</td>
<td>0.74</td>
<td>-0.007</td>
<td>0.008</td>
<td>0.66</td>
<td>-0.008</td>
<td>0.011</td>
<td>0.71</td>
<td>0.249</td>
</tr>
<tr>
<td><strong>LA agriculture</strong></td>
<td>0.100</td>
<td>-0.138</td>
<td>1.11</td>
<td>0.109</td>
<td>-0.145</td>
<td>1.19</td>
<td>0.108</td>
<td>-0.145</td>
<td>1.18</td>
<td>0.027</td>
</tr>
<tr>
<td><strong>LA unemployment</strong></td>
<td>0.667</td>
<td>-0.926</td>
<td>2.00</td>
<td>0.617</td>
<td>-0.829</td>
<td>1.83</td>
<td>0.596</td>
<td>-0.789</td>
<td>1.75</td>
<td>0.040</td>
</tr>
<tr>
<td><strong>County effects</strong></td>
<td>0.780</td>
<td>0.0585</td>
<td>2964.61</td>
<td>0.48</td>
<td>0.0204</td>
<td>1190.61</td>
<td>0.0000</td>
<td>0.786</td>
<td>0.343</td>
<td>1.333</td>
</tr>
<tr>
<td>Exogeneity test</td>
<td>0.333</td>
<td>0.570</td>
<td>2668.08</td>
<td>0.336</td>
<td>0.562</td>
<td>661.98</td>
<td>0.0000</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Insts. in 1st step</td>
<td>0.178</td>
<td>0.343</td>
<td>3637</td>
<td>0.178</td>
<td>0.343</td>
<td>3637</td>
<td>0.178</td>
<td>0.343</td>
<td>3637</td>
<td></td>
</tr>
</tbody>
</table>

Three category ordered probit estimates.

Neighbourhood educational measure is predicted proportion of adults in high qualification category in cohort member’s age-13 residential Ward. Ward proportions with more than seven rooms, council tenants used as instruments.

Other variables as Table 2-2.
We assume these are uncorrelated with educational attainments, conditional on
neighbourhood educational status, parental property size and tenancy group. An LM
(Sargan) test does not reject the overidentifying restrictions on neighbourhood Local
Authority housing and house size in the second step equation. At the same time, these
instruments are highly significant in the first step equation (P<0.001). Standard errors in
Table 2-4 are corrected using the method of Murphy and Topel (1985).

Using this method, the point estimates of the marginal effect of neighbourhood on
male educational attainments are substantially higher – by around 20%. This is probably
because the IV method corrects for sampling error in the 10% Census sample. The
estimates may also be slightly higher because the instruments are better predictors of
long-run neighbourhood status than the education variable, so estimates are purged of
transient variation in educational status on the night of the Census.

Based on these figures, a family moving from a neighbourhood at the bottom decile
of qualifications to a neighbourhood at the top decile would increase the probability of
their children gaining these qualifications by over six percentage points, unconditional on
neighbourhood schools (31.7% versus 37.8%). The elasticity for the probability of
attaining high qualifications with respect to the neighbourhood proportion with high
qualifications is 0.11 at the sample mean. The comparable elasticity for failure to gain
anything above CSE grade 2 is about -0.15. Continuing the assessment of school quality
effects, we see, in moving from Column (1) to Column (2), that school selection probably
accounts for little more than 14% of the neighbourhood coefficient. Controlling for
observed school quality and type gives a neighbourhood elasticity of high educational
attainment of +0.10/-0.14 at the mean, for boys.

---

15 Our instruments are taken from the 100% sample. This measurement error issue is discussed in
more detail in Appendix C.
2.5.4 Social tenants’ attainments

We turn now to the approach outlined in Section 2.3.4 using families who were reported as council tenants in both 1969 and 1974. Our assumption is that the neighbourhood status of any socially housed tenant is unrelated to their family resources — relative to other socially housed children — and that most of the variation in their neighbourhood status is driven by the proportion of social tenants and the status of neighbouring owner-occupiers. Table 2-5 shows the marginal effects from ordered probit regressions for a pooled sample of men and women. In Column (1), which includes area, labour market and County controls, we see effects of similar magnitude to those from the full sample with parental controls. A one percentage point shift in the proportion of neighbours with A Levels and above increases the probability of the child of a council tenant gaining A Levels by 0.25%. This is equivalent to an elasticity at the mean of around 0.13. The same shift in neighbourhood status lowers the chance of ending up without any formal qualifications by 0.34%. Again, the elasticity is around 0.1. A 10th percentile to 90th percentile move through the population distribution of neighbourhoods would increase the probability of a social tenant gaining A-levels from to 15.6% to 20.1%. As before, we assess whether this relationship is mediated through the school environment and allow for observable school selection effects (Column (2)). The key coefficient is still substantial and significant. The elasticity of the attainment of high qualifications with respect to neighbourhood educational status is still around 0.1.
Table 2-5: Neighbourhood education effects on social tenants, age 33 qualifications

<table>
<thead>
<tr>
<th>Qualification group:</th>
<th>(1) Low</th>
<th>(1) High</th>
<th>(2) Low</th>
<th>(2) High</th>
<th>(3) 2-step Low</th>
<th>(3) 2-step High</th>
<th>(4) 2-step Low</th>
<th>(4) 2-step High</th>
<th>Mean</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ward education</td>
<td>-0.345</td>
<td>0.251</td>
<td>2.70</td>
<td>-0.244</td>
<td>0.178</td>
<td>2.11</td>
<td>-0.726</td>
<td>0.528</td>
<td>2.47</td>
</tr>
<tr>
<td>School quality</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-0.191</td>
<td>0.140</td>
<td>6.25</td>
<td>-</td>
<td>-</td>
<td>-0.365</td>
</tr>
<tr>
<td>Early attainments</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-0.934</td>
<td>0.686</td>
<td>19.11</td>
<td>-</td>
<td>-</td>
<td>-0.876</td>
</tr>
<tr>
<td>Missing</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-0.500</td>
<td>0.366</td>
<td>13.29</td>
<td>-</td>
<td>-</td>
<td>-0.455</td>
</tr>
<tr>
<td>Male</td>
<td>-0.078</td>
<td>0.056</td>
<td>5.20</td>
<td>-0.076</td>
<td>0.056</td>
<td>5.56</td>
<td>-0.077</td>
<td>0.056</td>
<td>5.17</td>
</tr>
<tr>
<td>Local social housing</td>
<td>-0.001</td>
<td>0.005</td>
<td>0.37</td>
<td>-0.000</td>
<td>0.000</td>
<td>0.01</td>
<td>-0.018</td>
<td>0.013</td>
<td>0.99</td>
</tr>
<tr>
<td>City Ward</td>
<td>-0.012</td>
<td>0.009</td>
<td>0.57</td>
<td>-0.025</td>
<td>0.019</td>
<td>1.30</td>
<td>-0.015</td>
<td>0.011</td>
<td>0.75</td>
</tr>
<tr>
<td>High population</td>
<td>-0.018</td>
<td>0.013</td>
<td>0.38</td>
<td>-0.016</td>
<td>0.012</td>
<td>0.84</td>
<td>-0.014</td>
<td>0.010</td>
<td>0.68</td>
</tr>
<tr>
<td>LA agriculture</td>
<td>0.020</td>
<td>-0.015</td>
<td>0.13</td>
<td>0.009</td>
<td>-0.007</td>
<td>0.06</td>
<td>0.010</td>
<td>-0.007</td>
<td>0.05</td>
</tr>
<tr>
<td>LA unemployment</td>
<td>1.334</td>
<td>-0.970</td>
<td>2.31</td>
<td>1.051</td>
<td>-0.770</td>
<td>1.95</td>
<td>1.018</td>
<td>-0.741</td>
<td>1.97</td>
</tr>
<tr>
<td>County effects</td>
<td>χ²₀₀=110.99, P=0.0000</td>
<td>χ²₀₀=94.44, P=0.0030</td>
<td>χ²₀₀=102.71, P=0.0005</td>
<td>χ²₀₀=97.61, P=0.0015</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sargan test</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Predicted group prob.</td>
<td>0.324</td>
<td>0.179</td>
<td>0.324</td>
<td>0.179</td>
<td>0.324</td>
<td>0.178</td>
<td>0.325</td>
<td>0.179</td>
<td></td>
</tr>
<tr>
<td>Log likelihood</td>
<td>-2789.71</td>
<td>-2542.99</td>
<td>-2790.06</td>
<td>-2537.15</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pseudo R²</td>
<td>0.029</td>
<td>0.115</td>
<td>0.029</td>
<td>0.117</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Sample size = 2818.

Predicted neighbourhood education uses proportion of owner-occupier properties with one-two rooms, seven or more rooms, lacking various amenities.

School quality is proportion of own sex, age 15 studying for GCEs, predicted by school type and pupil-teacher ratio to remove catchment area effects on school quality (but not peer group effects or correlation between pupil performance due to selection by schools or by parents). Exclusion of early attainments measure tests sensitivity to selection on ability to high performing schools, by schools or parents - without early attainment control estimates in Column (4) are -0.706/4.48 for school quality parameter, -0.628/5.09 for neighbourhood parameter. Neighbourhood parameter is only slightly sensitive, whilst the school performance parameter almost doubles.

Inclusion of parental education gives marginal neighbourhood effect of -0.29/+0.21 (t = 2.263) in Column (1), -0.730/+0.529 (t = 2.496) in Column (3).

Inclusion of parental interest dummies gives marginal neighbourhood effect of -0.24/+0.17 (t = 1.984) in Column (1), -0.608/+0.438 (t = 2.164) in Column (3).

Neighbourhood instruments have p-value < 0.0001% in prediction equations. Inclusion of Ward-proportion in council housing in instrument set gives estimates of neighbourhood marginal effect almost identical to non-instrumented estimate: e.g. +0.185/-0.254 with school quality control.
So far in this section, we have used raw neighbourhood and school quality. Our identifying assumption here was that the allocation of council tenants to neighbourhoods is unrelated to the tenants' own education, incomes, or concern for their child's education.\textsuperscript{16} Even if there is some correlation between parental characteristics and neighbourhood within the socially housed group, it is unlikely that these characteristics are correlated with their non-socially housed neighbours. Again, from the discussion in Section 2.3.4, we can use home-owner characteristics as instruments for social tenants neighbourhood. Table 2-5, Columns (2) and (3) present 2-step IV models for social tenants, using home-owner property size and housing amenities as instruments. The overidentifying restrictions test as acceptable (p-value=0.352). More direct evidence in favour of home-owner's characteristics as instruments, is that they are uncorrelated with council tenants' parental education, which we know affects their children's attainments\textsuperscript{17}. Just as we found with the full sample results, the IV estimates are substantially higher – in fact more than double what we get using the raw neighbourhood variable. The implied elasticities at the mean are 0.260 on attaining A-levels plus, -0.197 on less than CSE

\textsuperscript{16} A basic test of this assumption is to regress parent's education on neighbourhood educational status (with local and County controls). We find that the education of council tenants was not completely unrelated to the neighbourhood in the 1970s. The elasticity between parents’ mean years of education and the Ward-proportion of adults with A-levels and above is around 0.004 (s.e. 0.0015, N = 2818) for council tenants. This could mean that local authorities matched tenants in terms of their education or incomes, or that better educated tenants pushed for accommodation in better neighbourhoods. Alternatively, the correlation may reflect effects from neighbourhood persisting from the previous generation. Nevertheless, the correlation is weak compared to that for home-owners: the elasticity between home-owners' education and neighbourhood status is over ten times higher than that for council tenants, at 0.041 (s.e. 0.002).

\textsuperscript{17} The F-statistic for the joint test of the coefficients on the five instruments in regression of log mean parental education on these, and the other characteristics in Table 2-5 is 0.63, with a p-value of 0.68.
grade 1. This change is more than we would expect after correcting for sampling variation alone, and suggests that other factors are at work.

One possibility is that it is the educational status of owner-occupiers at the boundary between areas of social and non-social that matters for social tenants’ educational outcomes. This would be consistent with role-model or expectations-related effects on human capital accumulation. In this case, the education of neighbouring home-owners may be a better measure of social tenants’ perceptions of more advantaged neighbours than overall Ward-level averages. Variation in the education of non-socially housed residents results in variation in neighbourhood composition along the social–non-social housing boundary; variation in the relative proportions of social and non-social tenants does not, except in very small clusters of social housing. This view has some further empirical support: including the Ward-proportion in social housing in the instrument set reduces the estimated marginal effects to the values obtained with the raw neighbourhood measure. A similar interpretation is that the IV estimates are, what the programme evaluation literature describes as, Local Average Treatment Effects (LATE) – see Angrist and Imbens (1994). If the structural parameter varies across individuals, or changes over the distribution of neighbourhoods, then the LATE parameter estimate is interpretable as the average effect over the range predicted by the instruments, or for the sub-group affected by variation in the instruments.

We can test the robustness of these results using school quality measures. If the assumption of exogeneity of home-owner property characteristics is correct, we should expect little change once these are included. Comparing Column (3) of Table 2-5 with Column (4), which includes a predicted school quality measure, confirms our expectations. The measured impact of neighbourhood is almost unchanged18.

18 Since school performance is potentially endogenous, we predict from school type, and pupil-teacher ratio, and include age-7 attainments. This is imperfect, since class sizes and school choice
2.5.5 Non-linearities in response

The parametric approach of the previous results restricts the functional form of the empirical attainment-neighbourhood relationship. Figure 2-1 to Figure 2-3 enrich the analysis using the semi-parametric procedure of Section 2.3.5, by allowing for general non-linearities. Figure 2-1 (a), shows how the probability of attaining A-levels or higher qualifications by age 33 increases as the educational status of a teenager’s neighbourhood increases (and 10% confidence intervals). The likelihood of becoming highly qualified increases steadily as the proportion of highly qualified neighbours increases. There is no indication here that being brought up in neighbourhoods at either end of the distribution of qualification levels makes things disproportionately better or worse – the relationship is predominantly linear. Figure 2-1 (b) shows the relationship once we control for family background. There is more evidence of non-linearities, but most of the neighbourhood effect seems to be on individuals originating in the upper-middle of the distribution, not at either end. Looking now at lower qualifications, Figure 2-2 shows the impact on proportion of the cohort achieving only the lowest levels of qualifications. The relationships mirror those in Figure 2-1. Next, Figure 2-3 shows the same relationship for social tenants only, to minimize parental selection effects without using parental controls. The overall impression is similar to that in the other Figures.

The key story from these Figures is that marginal improvements in residential neighbourhood reduce the probability of failure in the educational system throughout the distribution of neighbourhood educational status. Nothing here indicates that the extremes of neighbourhood deprivation or privilege matter disproportionately.

are in part determined by selection on unobserved ability abilities. In fact, using the raw performance measures or removing the age-7 ability controls has little impact on the parameter of interest. Similarly Including parental education or interest in the social tenants’ models has only a small impact on the neighbourhood coefficient. See the notes at the foot of Table 2-5.
Figure 2-1: Attainments and childhood neighbourhood; semi-parametric estimates

a) County controls only

![Graph showing semi-parametric estimates with county controls only.]

b) Controlling for parental and other locational characteristics

![Graph showing semi-parametric estimates with controls for parental and other locational characteristics.]

Figures show kernel regressions of age 33 qualifications on childhood neighbourhood status, with 10% confidence intervals. Controls in figure 1b) are parental education, family size, father’s age, residential tenure, rooms in family home, mean age 7 test scores, Local Authority agricultural employment, unemployment rate and County dummies. Epanachikov kernel, bandwidth by Silverman’s rule. 10th, 50th, 90th percentiles of neighbourhood measure = 0.037, 0.095, 0.216. N=9279.
Figure 2-2: Attainments and childhood neighbourhood: semi-parametric estimates

a) County controls only

b) Controlling for parental and other locational characteristics

Figures show kernel regressions of age 33 qualifications on childhood neighbourhood status, with 10% confidence intervals. Controls in figure 1b) are parental education, family size, father’s age, residential tenure, rooms in family home, mean age 7 test scores, Local Authority agricultural employment, unemployment rate and County dummies. Epanachikov kernel, bandwidth by Silverman’s rule. 10th, 50th 90th percentiles of neighbourhood measure = 0.037, 0.095, 0.216. N=9279. 10th, 50th 90th percentiles of neighbourhood measure = 0.037, 0.095, 0.216.
Figure 2-3: Attainments of social tenants: semi-parametric estimates

a) A Levels and higher, County controls only

Figures show kernel regressions of age 33 qualifications indicator on childhood neighbourhood status. Epanachikov kernel, bandwidth by Silverman’s rule. 10% pointwise confidence intervals shown. 10th, 50th, 90th percentiles of neighbourhood measure = 0.030, 0.077, 0.175. N = 3418.
2.5.6 Effects on abilities and attitudes

So far, we have focussed on effects of childhood neighbourhood on individual educational attainments by age 33. But neighbourhood could operate on educational choices from teenage years to adulthood via numerous channels – direct influences on abilities and aptitudes, motivation to achieve qualifications, incentives on whether to stay on at school, drop out or whether to continue to higher education. We can shed some light on this by looking at measured academic abilities, rather than on an individual’s educational trajectory, using the standard tests given to the NCDS cohort members at age 16.

Table 2-6 presents some results for this exercise, for a pooled sample of men and women. All estimates are conditional on a school quality measure and scores in comparable tests at age-11. Columns (1) to (4) show the parameter estimates for all tenancy groups, in reading and maths test, with and without parental background controls. Introducing parental controls halves the estimated neighbourhood effects. The small elasticities are unsurprising given we are measuring the effect of teenage neighbourhood on gains in test scores between age-11 and age-16. Still, children from the 90th percentile in the distribution of neighbourhoods achieved reading test scores that were on average 1.3% higher than those of children from the 10th percentile. For mathematics, the gain was around 3.6%. Columns (5) to (8) show similar estimates for council tenants only. Introducing parental controls has a fairly small effect on the neighbourhood parameter, supporting the assumptions of Section 2.3.4. Children from social housing at the 90th percentile of neighbourhoods achieved reading scores that were 2.8% better, and maths scores which were 5.8% better than their counterparts living in neighbourhoods at the 10th percentile. Since we control for age-11 test scores and secondary school performance, these figures are net of any impacts on individual primary school age achievements, or on average pupil attainments in secondary school.
### Table 2-6: Teenage attainments of NCDS cohort, age 16 tests, conditional on age 11 tests and schooling

<table>
<thead>
<tr>
<th>All tenancy groups</th>
<th>Social tenants only</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Reading</td>
</tr>
<tr>
<td></td>
<td>(1)</td>
</tr>
<tr>
<td>1971 Ward education</td>
<td>0.160 (0.032)</td>
</tr>
<tr>
<td>High school quality</td>
<td>0.054 (0.007)</td>
</tr>
<tr>
<td>Age 11 test log-score</td>
<td>0.567 (0.012)</td>
</tr>
<tr>
<td>Male</td>
<td>0.011 (0.005)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>1971 County effects</th>
<th>( \chi^2 )=164</th>
<th>( \chi^2 )=12</th>
<th>( \chi^2 )=131</th>
<th>( \chi^2 )=193</th>
</tr>
</thead>
<tbody>
<tr>
<td>Log family size</td>
<td>-0.052</td>
<td>-0.052</td>
<td>-0.052</td>
<td>-0.052</td>
</tr>
<tr>
<td>Rooms in family home</td>
<td>-0.000</td>
<td>-0.002</td>
<td>-0.002</td>
<td>-0.002</td>
</tr>
<tr>
<td>Father’s age</td>
<td>-0.001</td>
<td>0.0015</td>
<td>-0.001</td>
<td>0.0015</td>
</tr>
<tr>
<td>Father’s education</td>
<td>( \chi^2 )=5.5</td>
<td>( \chi^2 )=38.6</td>
<td>( \chi^2 )=2.7</td>
<td>( \chi^2 )=20.0</td>
</tr>
<tr>
<td>Mother’s education</td>
<td>( \chi^2 )=14.3</td>
<td>( \chi^2 )=32.0</td>
<td>( \chi^2 )=11.3</td>
<td>( \chi^2 )=4.8</td>
</tr>
<tr>
<td>Tenancy group</td>
<td>( \chi^2 )=26.2</td>
<td>( \chi^2 )=46.0</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

Sample size: 8304 8205 3250 3189

Dependent variable is natural log of test scores.
Models are interval regression, to allow for upper and lower censoring.
Instrumenting school performance with school type and pupil-teacher ratio leaves results virtually unchanged.

The age-16 sweep asks whether the child expects to study for A-levels or equivalent, so we have an opportunity to explore the influence of neighbourhoods on educational expectations. Using similar specifications to Table 2-6, it turns out that school performance, not neighbourhood affects aspirations for girls. Yet neighbourhood status has a significant positive influence on educational aspirations of boys, over and above secondary school quality. Boys from neighbourhoods at the 90th percentile were 3.3 percentage points more likely to expect to study for A-levels than those at the 10th percentile – a relative shift of just under 10%. This gender difference could measure a real difference in behavioural response, but may just reflect the weakness of our neighbourhood educational status variable as a proxy for female role models.
2.6 An overview of the spatial contribution to educational attainment

2.6.1 Intergenerational mobility and neighbour correlations in the NCDS

All the approaches adopted above suggest that neighbourhood human capital levels have some impact on educational attainments, albeit relatively small once we allow for parental effects. We have focused on relatively small local communities defined by Census wards, and have looked for ways of getting robust and plausible estimates of the impact of these communities. This section presents summary results which characterise the relative importance of geographical areas and parents on educational outcomes. The focus is on parental and area-level educational status only, in terms of educational mobility as measured by the traditional immobility parameter – see Atkinson (1981), Dearden, et al. (1997), and Solon (1989, 1999) – and on inter-neighbour correlations in educational outcomes.

Kremer (1997) argues that the existence of neighbourhood effects has little impact on equilibrium inequality or intergenerational mobility, largely because most of the distribution in earnings is not explained by family background factors and because neighbourhoods and families are not permanently linked. Nevertheless, if we are focussing on what can be explained, it is reasonable to ask: a) what proportion of the variance in educational attainments could be attributable to neighbourhood and b) what proportion of the persistence in economic status across generations of the same family is attributable to neighbourhood or area of upbringing? Table 2-7 answers the first question, based on the correlation between outcome educational attainments and the attainments of other cohort members originating from the same Census Ward (at age 16). Table 2-8 answers the second question for various definitions of neighbourhood – Census Enumeration District, Ward, and County.
2.6.1.1 Inter-neighbour correlations

We can get an upper bound to the influence of neighbourhood by the correlation between an individual's adult attainments and those of his or her neighbours as a child. This approach has been applied in the sibling-correlations literature to give an upper bound to the impact of family background (Solon, et al. (1991)) and has also been used in the neighbourhood context by Solon, et al. (2000) to bound neighbourhood effects on education in the US. Table 2-7 presents the regression coefficients $\lambda$ from the model

$$\hat{h}_{ij} = \lambda \hat{h}_{j-1} + x_j' \beta + \omega_i$$  \hspace{1cm} (2-20)

where $\hat{h}_{ij}$ is years in education for individual $i$, resident in Ward $j$ at age 16, standardised to unit variance. Variable $\hat{h}_{j-1}$ is the mean years of education for other children brought up in Ward $j$ standardised to unit variance. Since $Var(\hat{h}_i) = Var(\hat{h}_j)$ the coefficient $\lambda$ is a consistent estimate of the correlation between an individual's education and that for the average other child in the childhood Ward of residence. This is clearly an upper bound on direct neighbourhood effects, since it includes correlation in outcomes attributable to correlations between neighbours' family background characteristics.

Table 2-7: Inter-neighbour correlations in age-33 education

<table>
<thead>
<tr>
<th>Controls for gender only</th>
<th>Controls for County in 1971 and gender</th>
<th>Controls for County in 1971, gender and parents' education</th>
</tr>
</thead>
<tbody>
<tr>
<td>Unweighted</td>
<td>0.107 (0.016)</td>
<td>0.087 (0.016)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.043 (0.014)</td>
</tr>
<tr>
<td>Weighted by number in</td>
<td>0.159 (0.019)</td>
<td>0.128 (0.020)</td>
</tr>
<tr>
<td>Ward</td>
<td></td>
<td>0.065 (0.018)</td>
</tr>
<tr>
<td>Sample size</td>
<td>8237</td>
<td>8237</td>
</tr>
</tbody>
</table>


The OLS estimates in the first Row 1 of the first column in Table 2-7 suggest an inter-neighbour correlation of, at most, 0.11. Following Solon, et al. (2000), the estimates
in Row 2 weight the regressions by the number of other cohort members in each Ward, because wards with more observed residents provide more information. This pushes up the estimate to 0.16. Moving across the Columns of Table 2-7, the neighbour correlation falls as we introduce County dummies, and then again once parental education is included. Conditional on these factors, the correlation between outcome education years for a child and his or her average neighbourhood peer is low – between 0.043 and 0.065. One way of looking at this is that a child could expect to increase his or her time in education by between 3.2 to 11.8 weeks if brought up amongst children destined to stay in education for 1.4 years longer than average.

These are weaker neighbourhood effects on education than those found in the US by Solon, et al. (2000), using US PSID data from the 1970s. They find unadjusted correlations of 0.153-0.192, falling to 0.062-0.104 when adjusted for parental income. But then their neighbours lived in much closer proximity to ours – perhaps within the same block in urban areas – making direct comparison with Ward-based results difficult\textsuperscript{19}. However, the unadjusted correlations are stronger than those measured by Raaum, et al. (2001) for residents in smaller neighbourhoods in Norway, using Census data. These authors get an unadjusted figure is 0.067, falling to 0.022 when adjusted for parental education. These country differences indicate that the institutional setting has some influence on neighbourhood-based inequalities in educational outcomes. Measures such as the Gini coefficient usually rank the US above Britain and both above Norway in terms of income inequality (Forster and Pellizzari (2000)). These parallel results for inter-neighbour correlations suggest that neighbourhood of origin has some role to play in driving differences in patterns of individual inequality between nations.

\textsuperscript{19} Note though that standardisation of our variables to unit variance makes the inter-neighbour correlations insensitive to increases in neighbourhood-area induced by aggregation. We can aggregate the right hand side of (2-20) and still get a consistent estimate of $\beta$. 

- 70 -
2.6.1.2 Intergenerational mobility

Looking now at neighbourhoods role in intergenerational mobility, Table 2-8 shows estimates of the parameters in the model:

\[ \tilde{h}_i^c = \rho_1 \tilde{h}_i^n + \rho_2 \tilde{h}_i^p + \epsilon_i \]  

(2-21)

In equation (2-21) the tilde indicates standardised, unit variance, zero mean transformations. The first row of the Table constrains \( \rho_1 \) to zero, the second row constrains \( \rho_2 \) to zero, rows three and four present the unconstrained parameters. The dependent variable is either standardised years of education, or a dummy indicating attainment of A-level qualifications or higher at age 33.\(^{20}\) In the probit case, the marginal effect just gives the effect of a one standard deviation change in the explanatory variable on the probability of gaining A-levels, since we cannot observe or adjust for changes in the variance of the underlying latent child's human capital.

The basic educational years mobility parameter in Row 1 is in line with other estimates in the literature. Looking at Row 2, Column (1), the parameter estimate of 0.263 means that children who end up 0.263 standard deviations above the mean in the distribution of time in education come from neighbourhoods which were one standard deviation above the mean. The influence of area diminishes as the definition broadens from Enumeration District to Ward to County. Note that the Ward-level estimates in Column (7), which include County dummies, are barely different from those in Column (3). County level effects have a minimal role in intergenerational educational mobility, conditional on neighbourhood. The Probit marginal effects in the even Columns show the same pattern.

\(^{20}\) Standardisation of the variables ensures that the coefficients are unaffected by general changes in the variance of human capital across generations. An asymptotically equivalent approach for the single regressor case is to rescale the parameter estimates by multiplying by \( \sigma^p/\sigma^c \) (see Solon, 1999).
Table 2-8: Area effects on intergenerational educational mobility

<table>
<thead>
<tr>
<th></th>
<th>ED(^1)</th>
<th>Ward</th>
<th>County</th>
<th>Ward</th>
<th>County</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Years (1) A Levels+ (2)</td>
<td>Years (3) A Levels+ (4)</td>
<td>Years (5) A Levels+ (6)</td>
<td>Years (7) A Levels+ (8)</td>
<td></td>
</tr>
<tr>
<td>Standardised parental education only</td>
<td>0.382 (0.011) 0.143 (0.006)</td>
<td>0.382 (0.011) 0.143 (0.006)</td>
<td>0.382 (0.011) 0.143 (0.006)</td>
<td>0.379 (0.011) 0.142 (0.006)</td>
<td></td>
</tr>
<tr>
<td>Standardised area measure only</td>
<td>0.263 (0.011) 0.098 (0.005)</td>
<td>0.222 (0.010) 0.087 (0.005)</td>
<td>0.084 (0.010) 0.032 (0.005)</td>
<td>0.221 (0.011) 0.087 (0.005)</td>
<td></td>
</tr>
<tr>
<td>Standardised parental education</td>
<td>0.333 (0.011) 0.124 (0.006)</td>
<td>0.349 (0.011) 0.129 (0.006)</td>
<td>0.377 (0.011) 0.141 (0.006)</td>
<td>0.348 (0.011) 0.130 (0.006)</td>
<td></td>
</tr>
<tr>
<td>Standardised area</td>
<td>0.159 (0.011) 0.064 (0.005)</td>
<td>0.130 (0.010) 0.056 (0.005)</td>
<td>0.047 (0.009) 0.018 (0.005)</td>
<td>0.130 (0.011) 0.056 (0.005)</td>
<td></td>
</tr>
<tr>
<td>Percentage of familial intergenerational mobility attributable to area</td>
<td>11212 9279</td>
<td>11212 9279</td>
<td>11212 9279</td>
<td>11212 9279</td>
<td></td>
</tr>
<tr>
<td>Percentage of area effect attributable to parents</td>
<td>12.8% 13.2%</td>
<td>8.6% 9.8%</td>
<td>1.3% 1.4%</td>
<td>8.2% 8.5%</td>
<td></td>
</tr>
</tbody>
</table>

Proportion with A-levels plus in 1991 is 32.9%.

Note: 1. Enumeration District (ED) human capital stock is proxied by the proportion in professional and managerial jobs, since we have no data on educational attainments at ED level.
Unsurprisingly, once we include both neighbourhood and parental education in the equations (Rows 3 and 4), the independent effects of parental education and neighbourhood reduce – by the percentages shown in Rows 5 and 6. Row 5 shows the proportion of the intergenerational mobility parameter that we can attribute to neighbourhood status. Row 6 shows the proportion of the association between neighbourhood and educational attainment that is explained by parental education and sorting effects – around 35-40%. At most neighbourhood status contributes 13% to the intergenerational immobility. These results suggest a relatively weak contribution of between-area variation to overall patterns of inequality. This is in line with the findings of Berthoud (2001) who shows that between-postcode-sector variation accounts for less than 10% of the total variance in household incomes in Britain (the postcode sector is a geographical unit of similar size to a Ward). At the regional level, this figure drops to 2.7%.

Generalising (2-19), we can allow for a non-linear relationship between child’s, parent’s and mean neighbourhood educational attainments:

\[ \tilde{h}_i^c = g(\tilde{h}_i^n, \tilde{h}_i^p) + \epsilon_i \]  

(2-22)

This unrestricted function can be estimated by 2-regressor kernel regression. The results are shown in Figure 2-4. We have here a clear illustration of the relative effects of parental background and neighbourhood on educational attainment. Looking at the regression surface, there is little evidence of complementarities between parental and neighbourhood education: children from parents at the 90th percentile in the educational distribution can expect to end up roughly two and a half deciles above their counterparts at the 10th percentile, regardless of neighbourhood status. Parental effects are highly

---

21 See Appendix F for details of this computation.

22 This was what we found in Section 5.2.4.
non-linear – over half of this impact occurs within the top quintile of the parental distribution. Compared to parental education, neighbourhood has a relatively small impact, shifting children of similar parentage up the educational distribution by around one decile on average. Reading along the median outcome contour, we can see that a child at the top of the distribution of neighbourhood educational status can expect to reach median educational attainments, even if their parents’ educational score lies at the 35th percentile. By contrast, someone in a neighbourhood at the very bottom of the distribution must have parents educated to about the 85th percentile.

Figure 2-4: Educational intergenerational mobility and neighbourhood

2.6.2 Has much changed since the 1970s?

2.6.2.1 Changes in the area educational effects

Ideally, we would like to compare these results for children raised in the 1960s and 1970s with a later cohort. Unfortunately, the more recent cohort in the 1970 British
Cohort Study (BCS), has no codes for neighbourhood area of residence. The smallest geographical area to which we can assign to a child’s residential address is District Health Authority (DHA) at age 10, which can be matched to 1981 Census data. We can find some point of comparison here with the earlier cohort by comparing area education effects at DHA level in the BCS with area effects at Local Education Authority level in the NCDS. These are similar in terms of aggregation level.

Table 2-9 shows coefficients in the intergenerational mobility equation for the NCDS and BCS. Outcomes are at age 33 for the NCDS, but age 26 for the BCS. Looking at the results for years of education in Row 1, it seems that educational mobility changed little between the 1970s and 1980s. The immobility parameter changes little between Columns (1) and (5). In terms of higher qualifications, mobility appears to decrease: although the probit marginal effect estimates are higher in Columns (2) and (6), the relative effect at the mean increases from 0.44 to 0.51. Regional controls in Columns (3), (4), (7), (8) have little impact. Comparison of the 1980s and 1970s area effects reveals virtually no change in the association between area and educational outcome between the decades. Summarising, using the limited area data available for comparison, there is no evidence of weakening of area effects on educational attainments, but some mixed evidence of a strengthening of the link with parental education. This last result is surprising, but is consistent with findings in Blanden, et al. (2001) that income mobility in Britain decreased between the 1970s and 1980s.

---

23 This result is not robust to alternative ways of cutting the data. If we look at the children from the top 17% of the parental education distribution for the BCS and NCDS (the proportion of families with degree-educated fathers in the BCS and with non-missing data), we find that the probability of gaining a degree increased from 31.6% to 49.4% between cohorts. At the same time the probability of a child from the bottom 83% gaining a degree increased from 8.2% to 15%. Relatively speaking, the less educated fared better. A simple index of immobility (the sum of diagonal cell proportions, minus the sum of off diagonals) decreases from 0.63 to 0.58.
Table 2-9: Area effects on intergenerational educational mobility: comparison between 1970s and 1980s

<table>
<thead>
<tr>
<th></th>
<th>NCDS, 1970s</th>
<th></th>
<th>BCS, 1980s</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>LEA</td>
<td>LEA</td>
<td>Region</td>
<td>DHA</td>
</tr>
<tr>
<td></td>
<td>Years</td>
<td>A Levels+</td>
<td>Years</td>
<td>A Levels+</td>
</tr>
<tr>
<td>Standardised parental education only</td>
<td>0.384 (0.011)</td>
<td>0.145 (0.006)</td>
<td>0.383 (0.011)</td>
<td>0.145 (0.006)</td>
</tr>
<tr>
<td>Standardised area measure only</td>
<td>0.106 (0.010)</td>
<td>0.040 (0.005)</td>
<td>0.119 (0.014)</td>
<td>0.048 (0.007)</td>
</tr>
<tr>
<td>Standardised parental education</td>
<td>0.377 (0.011)</td>
<td>0.143 (0.006)</td>
<td>0.378 (0.011)</td>
<td>0.143 (0.006)</td>
</tr>
<tr>
<td>Standardised area</td>
<td>0.057 (0.009)</td>
<td>0.023 (0.005)</td>
<td>0.066 (0.013)</td>
<td>0.029 (0.007)</td>
</tr>
<tr>
<td>Percentage of intergenerational mobility attributable to area</td>
<td>1.8%</td>
<td>1.4%</td>
<td>1.3%</td>
<td>1.4%</td>
</tr>
<tr>
<td>Percentage of area effect attributable to parents</td>
<td>6.2%</td>
<td>42.5%</td>
<td>44.5%</td>
<td>39.5%</td>
</tr>
</tbody>
</table>

Proportion with A-levels plus in BCS, 1996 is 39.2%.
2.6.2.2 Segregation and spatial educational inequalities

Appendix G in this Chapter discusses changes in the spatial distribution of education using Census data from 1971, 1981 and 1991. The evidence from the Census data is that there has been virtually no change in the variance of the distribution of education across wards, though this does not deny that neighbourhoods may have changed rank in the distribution. The consensus from the US is that there has been increased segregation, and increase in the correlation between the characteristics of individuals and their neighbours (see Kremer (1997); or Gephart (1997) for references). In conjunction with structural neighbourhood effects, this would imply increasing inequality and intergenerational immobility, though Kremer argues that these effects are small. Increasing segregation implies a decrease in the variance of education within neighbourhoods and a widening of the distribution of mean education across neighbourhoods, holding the overall variance constant. This follows from the decomposition of variance:

\[ \text{Var}(E[h_i | n]) = \text{Var}(h_i) - E[\text{Var}(h_i | n)] \]  

(2-23)

For a given overall variance in education across individuals, an increase in the correlation between individuals' education within neighbourhoods \( n \) implies a decrease in the variance within neighbourhoods \( E[\text{Var}(h_i | n)] \), hence an increase in the variance of the mean across neighbourhoods \( \text{Var}(E[h_i | n]) \). Appendix G shows, however, that there has been virtually no such increase from 1971 to 1991 in Britain, once we correct for exogenous changes in \( \text{Var}(h_i) \) attributable to an increase in average educational achievements. From this evidence, neighbourhood effects are likely to have contributed little to any increases in inequality and social immobility. This does not detract from their potential importance in the static cross-sectional distribution.
2.7 Summary and concluding remarks

Children’s academic attainments are sensitive to community influences. The methods in this study focus on identifying a relationship between the proportion of highly qualified adults in a child’s neighbourhood and his or her educational attainments. The results show that the association between community attainments and child attainments is robust, under different empirical strategies that compensate for parental selection on schools and neighbourhood. In particular, children of social tenants brought up in Britain in the 1970s were influenced by the proportion of highly qualified adults in their neighbourhood, and with those components of neighbourhood educational status which are correlated with the physical characteristics of owner-occupied housing. These effects are at least as large as the effects estimated on the population as a whole. The fact that social tenants benefit is in contrast with the policy conclusions in Duncan (1994), who suggests that the weakness of effects on disadvantaged groups means that policy to redistribute resources between wealthy and poor neighbourhoods may have adverse effects.

Variation in educational status predicted from owner-occupied housing structure appears to have a stronger influence on the adult attainments of social tenants than does variation attributable to the proportion of social tenants. This could indicate that children resident in social housing are especially sensitive to the quality of the local residential community, outside their estate. The sensitivity of social tenants to home-owner characteristics – and the persistence of neighbourhood effects over and above measures of secondary school quality – supports the collective socialisation or adult role model effects in the sociological literature (Jencks and Mayer (1990)), the importance of social capital (Coleman (1988)) if this is related to the neighbourhood stock of human capital, or the influence of expectations formation within the local community (Roemer and Wets (1994), Streufert (1991)). Broadly speaking, the results provide evidence of educational
spillover effects from the community to the individual. The influence of community can be traced back to scores on attainment tests administered at age 16.

Reviewing the estimates in this study, we conclude that the probability of a child attaining high qualifications responds to the community proportion with these qualifications with an elasticity of around 0.1 in the average neighbourhood. The effect could be up to four-times greater than this in the centre of the distribution of neighbourhoods and weaker in the tails. Children of social tenants in the 1970s were similarly sensitive, though IV estimates that correct for measurement error, parental and school-based selection are substantially higher, with elasticities of around 0.26. Table 2-10 below summarises the key findings:

| Table 2-10: Summary impacts of teenage Ward proportion with higher qualifications |
|----------------------------------|-----------------|----------------|
|                                  | Elasticity at mean neighbourhood | 10th to 90th percentile change in neighbourhood |
| All tenancy groups, conditional on neighbourhood, area and school quality | 0.14 | -0.21 | 7.5 p.p. | -5.6 p.p. |
| All tenancy groups, conditional on background, area and school (IV) | 0.10 | -0.14 | 5.2 p.p. | -3.9 p.p. |
| Social tenants, conditional on school, area and early attainments | 0.10 | -0.07 | 2.6 p.p. | -3.5 p.p. |
| Social tenants, IV from owner occupied property | 0.26 | -0.20 | 7.1 p.p. | -9.8 p.p. |

Nothing in the results indicates that community effects are mediated via family circumstances, although there are significant interactions with indicators of parental interest in a child’s education. We can read this as meaning that parental interest affects the returns to neighbourhood and school performance in the human capital production.
function, or that the returns affect parental interest – these are observationally equivalent in the data available here.

Although the evidence is based on children who were teenagers some thirty years ago, there is no reason to believe that the underlying structural relationships will have changed in the intervening period. Comparison of broader area effects on children raised in the 1970s and 1980s using two different cohort studies provides no evidence of a weakening of the area components of intergenerational educational mobility.

The overall finding of this Chapter is that neighbourhoods do influence outcomes, regardless of family resources. In particular, children's educational attainments are sensitive to the adult educational composition of their neighbourhood. But we find nothing to contradict the general consensus that neighbourhoods determine only a small proportion of the variation in individual outcomes, and that family background matters more. Correlation between total time in education and the education of others from the same child-hood Ward is relatively weak: an upper bound on the inter-neighbour correlation in educational outcomes is 0.16. A more conservative estimate places this at around 0.07. These inter-neighbour spillover effects in educational attainment do imply higher benefits from tackling educational disadvantage at the neighbourhood level, rather than on an individual or family basis. But, the evidence from this Chapter is that these additional benefits are quite small.
### 2.8 Appendix A

#### 2.8.1 Dependent variables

This Appendix describes the dependent variables used in the main tables. Tables under each heading present summary statistics for the base samples used in the regressions.

*Highest qualification at age 33*: A 3-category qualifications variable derived from the 1991 cohort member interview, taking the values:

<table>
<thead>
<tr>
<th>Category</th>
<th>Description</th>
<th>Men</th>
<th>Women</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low qualifications</td>
<td>No qualifications, or qualifications below CSE-grade 1</td>
<td>18.7%</td>
<td>24.1%</td>
</tr>
<tr>
<td>Mid qualifications</td>
<td>O-levels, CSE grades 1 or more, lower and intermediate vocational qualifications (City and Guilds, BTEC etc.)</td>
<td>47.5%</td>
<td>43.9%</td>
</tr>
<tr>
<td>High qualifications</td>
<td>A-levels, higher vocational qualifications (Higher National Diplomas, teaching and nursing qualifications), First Degrees and equivalent, or Higher Degrees</td>
<td>33.8%</td>
<td>32.1%</td>
</tr>
</tbody>
</table>

*NCDS cohort member test scores*: Age 16 abilities are measured by a reading comprehension test devised by National Foundation for Educational Research, specifically for the NCDS and a mathematics comprehension test constructed by the same organisation. The reading test is heavily left skewed with the first quartile at 0.63.

<table>
<thead>
<tr>
<th></th>
<th>Observations</th>
<th>Mean</th>
<th>S.D.</th>
<th>Minimum</th>
<th>Maximum</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reading</td>
<td>8304</td>
<td>0.743</td>
<td>0.187</td>
<td>0.029</td>
<td>1</td>
</tr>
<tr>
<td>Arithmetic</td>
<td>8277</td>
<td>0.427</td>
<td>0.225</td>
<td>0</td>
<td>1</td>
</tr>
</tbody>
</table>

#### 2.8.2 Family background variables

*Education*: Education of the persons recorded as the father and mother figure in 1974. Codes as age left full-time education from 1974 sweep, or derived from data on whether father figure in 1965 stayed on at school and age left school. These education variables are recoded as a 3-category variable: 15 and under, 16-18, 18+ for use in the early attainment models.

*Family size*: Coded as maximum children under 21 recorded over all sweeps of the NCDS up to 1974. Enters as natural logarithms in regressions.

*House size*: Number of rooms in family residence recorded in 1974, or 1969 if missing, or 1965 if missing from both.
Observations | Mean | S.D. | Minimum | Maximum
---|---|---|---|---
Father education | 11336 | 15.14 | 1.99 | 12 | 32
Mother education | 11336 | 15.02 | 1.46 | 12 | 23
Number of children | 11336 | 3.29 | 1.70 | 1 | 14
Father's age | 11336 | 30.59 | 6.25 | 12.99 | 78.00
House rooms | 11336 | 4.93 | 1.39 | 1 | 32

*Parental interest in their child's education:* coded as a 12-category, non-ordered variable. The categories are:

<table>
<thead>
<tr>
<th>Parental interest</th>
<th>Read to child age 7</th>
<th>Proportion</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mother or Father very interested, all years</td>
<td>No</td>
<td>1.78</td>
</tr>
<tr>
<td>Mother or Father very interested, all years</td>
<td>Yes</td>
<td>3.56</td>
</tr>
<tr>
<td>Mother or Father very interested, two sweeps</td>
<td>No</td>
<td>8.57</td>
</tr>
<tr>
<td>Mother or Father very interested, two sweeps</td>
<td>Yes</td>
<td>14.35</td>
</tr>
<tr>
<td>Mother or Father very interested, one sweep</td>
<td>No</td>
<td>16.21</td>
</tr>
<tr>
<td>Mother or Father very interested, one sweep</td>
<td>Yes</td>
<td>20.13</td>
</tr>
<tr>
<td>Some, or various interest recorded</td>
<td>No</td>
<td>11.12</td>
</tr>
<tr>
<td>Some, or various interest recorded</td>
<td>Yes</td>
<td>11.43</td>
</tr>
<tr>
<td>Mother or Father very interested, one sweep</td>
<td>Unknown</td>
<td>4.69</td>
</tr>
<tr>
<td>Some, various or little interest recorded</td>
<td>Unknown</td>
<td>4.04</td>
</tr>
<tr>
<td>Little interest by mother and father for at least two sweeps</td>
<td>No</td>
<td>2.36</td>
</tr>
<tr>
<td>Little interest by mother and father for at least two sweeps</td>
<td>Yes</td>
<td>1.76</td>
</tr>
</tbody>
</table>

*Tenancy group:* recorded in 1974, or 1969 if missing, or 1965 if missing from both.

<table>
<thead>
<tr>
<th>Tenancy group</th>
<th>Proportion</th>
</tr>
</thead>
<tbody>
<tr>
<td>Owner-occupier</td>
<td>3.89%</td>
</tr>
<tr>
<td>Local Authority tenant</td>
<td>48.52%</td>
</tr>
<tr>
<td>Private tenant</td>
<td>39.08%</td>
</tr>
<tr>
<td>Tied, or other accommodation</td>
<td>5.02%</td>
</tr>
<tr>
<td>Missing</td>
<td>3.89%</td>
</tr>
</tbody>
</table>

### 2.8.3 Schooling variables

*School quality:* measured as the proportion of boys or girls aged 15 in the child's secondary school studying for GCE and SCE O-Levels.

<table>
<thead>
<tr>
<th>Observations</th>
<th>Mean</th>
<th>S.D.</th>
<th>Minimum</th>
<th>Maximum</th>
</tr>
</thead>
<tbody>
<tr>
<td>Boys</td>
<td>9007</td>
<td>0.259</td>
<td>0.332</td>
<td>0</td>
</tr>
<tr>
<td>Girls</td>
<td>9012</td>
<td>0.259</td>
<td>0.333</td>
<td>0</td>
</tr>
</tbody>
</table>

*Secondary school type:* at age 16 is categorised as follows:

<table>
<thead>
<tr>
<th>Proportion</th>
</tr>
</thead>
<tbody>
<tr>
<td>Comprehensive</td>
</tr>
<tr>
<td>Grammar</td>
</tr>
<tr>
<td>Secondary modern</td>
</tr>
<tr>
<td>Independent</td>
</tr>
</tbody>
</table>
Grant maintained 2.4%
Other, non-lea 2.5%

### 2.8.4 Area variables

#### 1971

<table>
<thead>
<tr>
<th>Variable</th>
<th>Mean</th>
<th>S.D.</th>
<th>Minimum</th>
<th>Maximum</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ward proportion with A levels +</td>
<td>0.114</td>
<td>0.078</td>
<td>0.000</td>
<td>0.733</td>
</tr>
<tr>
<td>Ward proportion unskilled</td>
<td>0.074</td>
<td>0.044</td>
<td>0.000</td>
<td>0.050</td>
</tr>
<tr>
<td>Ward unemployment</td>
<td>0.041</td>
<td>0.022</td>
<td>0.002</td>
<td>0.353</td>
</tr>
<tr>
<td>Ward proportion males econ active</td>
<td>0.606</td>
<td>0.051</td>
<td>0.250</td>
<td>1.000</td>
</tr>
<tr>
<td>Ward proportion females econ active</td>
<td>0.575</td>
<td>0.121</td>
<td>0.000</td>
<td>1.000</td>
</tr>
<tr>
<td>Ward New-Commonwealth immigrants</td>
<td>0.029</td>
<td>0.075</td>
<td>0.000</td>
<td>0.776</td>
</tr>
<tr>
<td>Ward proportion 1 year migrants</td>
<td>0.093</td>
<td>0.045</td>
<td>0.000</td>
<td>0.667</td>
</tr>
<tr>
<td>Ward mean rooms per household</td>
<td>4.849</td>
<td>0.566</td>
<td>2.190</td>
<td>7.750</td>
</tr>
<tr>
<td>Ward proportion lacking toilet or bath</td>
<td>0.029</td>
<td>0.029</td>
<td>0.000</td>
<td>0.222</td>
</tr>
<tr>
<td>Ward social housing</td>
<td>0.354</td>
<td>0.266</td>
<td>0.000</td>
<td>1.000</td>
</tr>
<tr>
<td>Local Authority agricultural employment</td>
<td>0.025</td>
<td>0.057</td>
<td>0.000</td>
<td>0.663</td>
</tr>
<tr>
<td>Local Authority mining &amp; manufacturing</td>
<td>0.374</td>
<td>0.130</td>
<td>0.000</td>
<td>0.725</td>
</tr>
<tr>
<td>Total population (1000s)</td>
<td>9.858</td>
<td>8.070</td>
<td>0.278</td>
<td>82.276</td>
</tr>
</tbody>
</table>

#### 2.8.4.1 Persistent labour market and LEA funding controls

I make the assumption that local labour markets and local educational expenditures are effective at County level, and use dummy variables to indicate County of residence at age 16 (or County of residence at age 23 for the models of NCDS member's children's attainments). This implies County of origin effects in the determination of adult human capital and early academic achievements.

As controls for educational spending, these effects are consistent with uniform spending within local education authorities in Great Britain, except in the metropolitan areas where spending may be more localised. Using LEA dummies makes little difference. Initial estimates were made using measures of LEA spending on secondary education in 1974 (child age 16), but the results suggest that this measure is endogenous due to prescriptive allocation of educational resources, with significant negative correlations between educational attainments and LEA spending.

The claim that County of origin dummies are adequate controls for local labour markets rests on the assumption that the cohort members are at least mobile within counties, between childhood and age 33. If labour markets are more localised, and there are significant fixed costs to moving from the parental neighbourhood, then the neighbourhood educational variable may still be endogenous. Low local demand for skills during childhood may mean low local demand for skills in adulthood if individuals are geographically immobile. This will almost certainly be true in rural communities where agricultural employment is relatively
high. In addition to the County dummies I include from the Census, the proportion of workers in agricultural employment in the Local Authority area.

2.8.4.2 Population, urban and rural effects

We have good reason to believe that the magnitude of the influences of neighbourhood will vary with neighbourhood population or population density, and that there will be differences between rural and urban areas. A richer network of contacts and proximity to others in a densely populated area may lead to stronger influences from others in the neighbourhood. Then again, the complexity of influences in a highly populated neighbourhood may break down any intergenerational links, whilst simpler structures in sparsely populated rural communities may encourage them. Differences in environmental influences between highly urbanised, metropolitan areas and others may influence occupational and educational choices.

Unfortunately it is not possible to construct exact population density measures with the available 1971 Census data. Instead I include a high (top 20%) Ward population indicator, which will proxy high population density if Ward areas are less variable than the populations. The regressions also include a city indicator, which defines all those Census districts listed as ‘C.B.’ (major cities and towns in England and Wales), ‘L.B.’ (London Boroughs), and ‘Cities’ in Scotland. Around 40% of the NCDS sample were resident in such districts in 1971. Unsurprisingly the correlation between the city and high population indicators is fairly high (0.43 in the whole NCDS sample, 0.49 in the selected male subsample).

Children brought up in agricultural areas will have obvious incentives to continue in agricultural work. Failure to control for agricultural employment will bias results in favour of finding neighbourhood effects on educational choices. One solution would be to exclude all those in agricultural jobs from the sample, at the expense of sample representativeness. The alternative, adopted in this Chapter, is simply to include a variable indicating the proportion of workers in agricultural employment in the childhood Local Authority area, taken from the 1971 Census. In fact, only 2% of the male subsample are in agricultural employment at age 33 and these can be omitted without changing the results significantly.
## 2.9 Appendix B: Results for women

### Table 2-11: Area-only models of women’s, age 33 qualifications

<table>
<thead>
<tr>
<th></th>
<th>I (Low)</th>
<th>I (High)</th>
<th>Mean</th>
<th>J (Low)</th>
<th>J (High)</th>
<th>Mean</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ward education</td>
<td>-0.570</td>
<td>0.655</td>
<td>6.236</td>
<td>-0.440</td>
<td>0.500</td>
<td>4.476</td>
</tr>
<tr>
<td>School quality</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-0.393</td>
<td>0.445</td>
<td>21.024</td>
</tr>
<tr>
<td>Unskilled workers</td>
<td>0.259</td>
<td>-0.297</td>
<td>1.466</td>
<td>0.170</td>
<td>-0.193</td>
<td>0.883</td>
</tr>
<tr>
<td>Unemployment rate</td>
<td>0.912</td>
<td>-1.047</td>
<td>2.515</td>
<td>0.591</td>
<td>-0.670</td>
<td>1.593</td>
</tr>
<tr>
<td>Economically active males</td>
<td>0.028</td>
<td>-0.032</td>
<td>0.288</td>
<td>-0.052</td>
<td>0.057</td>
<td>0.485</td>
</tr>
<tr>
<td>Economically active females</td>
<td>0.070</td>
<td>-0.081</td>
<td>1.249</td>
<td>0.033</td>
<td>-0.037</td>
<td>0.555</td>
</tr>
<tr>
<td>One year migrants</td>
<td>0.056</td>
<td>-0.065</td>
<td>0.512</td>
<td>0.023</td>
<td>-0.026</td>
<td>0.196</td>
</tr>
<tr>
<td>New com. immigrant</td>
<td>-0.147</td>
<td>0.169</td>
<td>1.845</td>
<td>-0.160</td>
<td>0.181</td>
<td>1.867</td>
</tr>
<tr>
<td>Average dwelling size</td>
<td>-0.004</td>
<td>0.005</td>
<td>0.271</td>
<td>0.014</td>
<td>-0.016</td>
<td>0.805</td>
</tr>
<tr>
<td>Household lacking amenities</td>
<td>0.537</td>
<td>-0.617</td>
<td>1.872</td>
<td>0.610</td>
<td>-0.692</td>
<td>2.022</td>
</tr>
<tr>
<td>Social housing</td>
<td>0.119</td>
<td>-0.136</td>
<td>3.515</td>
<td>0.113</td>
<td>-0.128</td>
<td>3.167</td>
</tr>
<tr>
<td>City Ward</td>
<td>-0.027</td>
<td>0.031</td>
<td>1.831</td>
<td>-0.005</td>
<td>0.006</td>
<td>0.319</td>
</tr>
<tr>
<td>High population</td>
<td>-0.001</td>
<td>0.002</td>
<td>0.098</td>
<td>-0.001</td>
<td>0.002</td>
<td>0.094</td>
</tr>
<tr>
<td>Agricultural employment</td>
<td>-0.009</td>
<td>0.010</td>
<td>0.074</td>
<td>0.000</td>
<td>-0.000</td>
<td>0.004</td>
</tr>
<tr>
<td>Mining, manufacturing</td>
<td>-0.008</td>
<td>0.009</td>
<td>0.128</td>
<td>0.015</td>
<td>-0.017</td>
<td>0.225</td>
</tr>
<tr>
<td>County effects</td>
<td></td>
<td></td>
<td></td>
<td>$\chi^2_{60} = 97.17, P = 0.002$</td>
<td>$\chi^2_{60} = 114.52, P = 0.000$</td>
<td></td>
</tr>
<tr>
<td>Predicted group probability</td>
<td>0.240</td>
<td>0.321</td>
<td></td>
<td>0.232</td>
<td>0.329</td>
<td></td>
</tr>
<tr>
<td>Log likelihood</td>
<td>-4967.89</td>
<td>0.002</td>
<td></td>
<td>-3794.69</td>
<td>0.000</td>
<td></td>
</tr>
<tr>
<td>Pseudo R²</td>
<td>0.039</td>
<td>0.096</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sample size</td>
<td>4835</td>
<td>3937</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Including Ward professional and managerial employees reduces parameter on Ward education to $0.278 \pm 0.315$ in Column (2) ($t=2.335$). Coefficient on professional workers is almost identical ($t=2.265$).

Neighbourhood educational measure is proportion of adults in high qualification category in cohort member’s age-13 residential Ward. School quality is proportion of girls age 15 studying for GCEs at cohort members’ school.
Table 2-12: Neighbourhood education and family effects on women's age 33 qualifications

<table>
<thead>
<tr>
<th>Qualification group</th>
<th>I Low</th>
<th>I High</th>
<th>t</th>
<th>II Low</th>
<th>II High</th>
<th>t</th>
<th>III Low</th>
<th>III High</th>
<th>t</th>
<th>Mean</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ward education</td>
<td>-0.147</td>
<td>0.176</td>
<td>2.61</td>
<td>-0.133</td>
<td>0.154</td>
<td>2.02</td>
<td>-0.142</td>
<td>0.163</td>
<td>2.15</td>
<td>0.117</td>
</tr>
<tr>
<td>School quality</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-0.174</td>
<td>0.202</td>
<td>10.24</td>
<td>-0.093</td>
<td>0.107</td>
<td>4.10</td>
<td>0.272</td>
</tr>
<tr>
<td>School type:</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Grammar</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-0.093</td>
<td>0.106</td>
<td>5.04</td>
<td>0.138</td>
</tr>
<tr>
<td>Secondary modern</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>0.012</td>
<td>-0.014</td>
<td>1.02</td>
<td>0.211</td>
</tr>
<tr>
<td>Independent</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-0.043</td>
<td>0.049</td>
<td>1.52</td>
<td>0.033</td>
</tr>
<tr>
<td>Grant maintained</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-0.113</td>
<td>0.129</td>
<td>2.97</td>
<td>0.026</td>
</tr>
<tr>
<td>Other non-LEA</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>0.036</td>
<td>-0.041</td>
<td>1.06</td>
<td>0.019</td>
</tr>
<tr>
<td>Mother education</td>
<td>-0.024</td>
<td>0.028</td>
<td>6.10</td>
<td>-0.022</td>
<td>0.025</td>
<td>5.51</td>
<td>-0.021</td>
<td>0.024</td>
<td>5.34</td>
<td>15.11</td>
</tr>
<tr>
<td>Father education</td>
<td>-0.013</td>
<td>0.016</td>
<td>4.46</td>
<td>-0.010</td>
<td>0.012</td>
<td>3.42</td>
<td>-0.010</td>
<td>0.011</td>
<td>3.21</td>
<td>15.22</td>
</tr>
<tr>
<td>Father's age</td>
<td>-0.001</td>
<td>0.002</td>
<td>1.83</td>
<td>-0.001</td>
<td>0.001</td>
<td>1.07</td>
<td>-0.001</td>
<td>0.001</td>
<td>1.03</td>
<td>30.69</td>
</tr>
<tr>
<td>Dwelling size</td>
<td>-0.004</td>
<td>0.004</td>
<td>1.10</td>
<td>-0.001</td>
<td>0.002</td>
<td>0.43</td>
<td>-0.001</td>
<td>0.001</td>
<td>0.34</td>
<td>4.978</td>
</tr>
<tr>
<td>Tenure:</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Tenure missing</td>
<td>0.089</td>
<td>-0.106</td>
<td>3.22</td>
<td>0.093</td>
<td>-0.108</td>
<td>3.11</td>
<td>0.092</td>
<td>-0.105</td>
<td>3.12</td>
<td>0.021</td>
</tr>
<tr>
<td>Council tenant</td>
<td>0.081</td>
<td>-0.097</td>
<td>7.67</td>
<td>0.087</td>
<td>-0.101</td>
<td>7.22</td>
<td>0.087</td>
<td>-0.100</td>
<td>7.15</td>
<td>0.380</td>
</tr>
<tr>
<td>Private rental</td>
<td>0.048</td>
<td>-0.057</td>
<td>2.21</td>
<td>0.048</td>
<td>-0.056</td>
<td>2.25</td>
<td>0.047</td>
<td>-0.053</td>
<td>2.18</td>
<td>0.043</td>
</tr>
<tr>
<td>Other tenure</td>
<td>0.022</td>
<td>-0.026</td>
<td>0.96</td>
<td>0.027</td>
<td>-0.032</td>
<td>1.15</td>
<td>0.028</td>
<td>-0.032</td>
<td>1.16</td>
<td>0.034</td>
</tr>
<tr>
<td>Parental interest</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Log siblings</td>
<td>0.054</td>
<td>-0.064</td>
<td>5.69</td>
<td>0.052</td>
<td>-0.060</td>
<td>5.38</td>
<td>0.052</td>
<td>-0.059</td>
<td>5.40</td>
<td>1.056</td>
</tr>
<tr>
<td>Age-7 test scores</td>
<td>-0.690</td>
<td>0.823</td>
<td>19.28</td>
<td>-0.616</td>
<td>0.714</td>
<td>17.01</td>
<td>-0.591</td>
<td>0.677</td>
<td>16.23</td>
<td>0.545</td>
</tr>
<tr>
<td>Missing</td>
<td>-0.394</td>
<td>0.470</td>
<td>13.93</td>
<td>-0.347</td>
<td>0.402</td>
<td>12.17</td>
<td>-0.331</td>
<td>0.379</td>
<td>11.55</td>
<td>0.102</td>
</tr>
<tr>
<td>Social housing</td>
<td>0.023</td>
<td>-0.028</td>
<td>0.16</td>
<td>0.002</td>
<td>-0.002</td>
<td>0.16</td>
<td>0.002</td>
<td>-0.002</td>
<td>0.135</td>
<td>0.400</td>
</tr>
<tr>
<td>City Ward</td>
<td>-0.015</td>
<td>0.018</td>
<td>1.30</td>
<td>-0.010</td>
<td>0.011</td>
<td>0.76</td>
<td>-0.012</td>
<td>0.014</td>
<td>0.96</td>
<td>0.405</td>
</tr>
<tr>
<td>High population</td>
<td>-0.028</td>
<td>0.033</td>
<td>2.15</td>
<td>-0.024</td>
<td>0.028</td>
<td>1.92</td>
<td>-0.023</td>
<td>0.027</td>
<td>1.83</td>
<td>0.241</td>
</tr>
<tr>
<td>LA agriculture</td>
<td>-0.075</td>
<td>0.090</td>
<td>1.09</td>
<td>-0.079</td>
<td>0.091</td>
<td>0.76</td>
<td>-0.092</td>
<td>0.106</td>
<td>1.06</td>
<td>0.026</td>
</tr>
<tr>
<td>Unemployment</td>
<td>0.819</td>
<td>-0.977</td>
<td>2.28</td>
<td>0.748</td>
<td>-0.868</td>
<td>2.05</td>
<td>0.735</td>
<td>-0.842</td>
<td>2.00</td>
<td>0.041</td>
</tr>
<tr>
<td>County effects</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Mean</td>
<td>0.232</td>
<td>0.328</td>
<td>0.233</td>
<td>0.328</td>
<td>0.233</td>
<td>0.328</td>
<td>-3317.91</td>
<td>-3266.22</td>
<td>-3249.81</td>
<td></td>
</tr>
<tr>
<td>Log likelihood</td>
<td>0.210</td>
<td>0.222</td>
<td>0.226</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Sample size</td>
<td>3937</td>
<td>3937</td>
<td>3937</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

Neighbourhood educational measure is proportion of adults in high qualification category in cohort member's age-13 residential Ward.

School performance is proportion of girls age 15 studying for GCEs.

Other variables as in Table 2-2.

Estimation of model of first column using full available sample of 4835 women gives marginal effect from neighbourhood education of -0.129 on the low category and 0.150 on the high category (t = 2.158, p-value = 0.031). The coefficient is not significantly different from estimate on the smaller subsample for which school performance is observable.
2.10 Appendix C: Sampling error in the 10% Census sample

A potential drawback with the Census data from the 10% sample is the sampling error when analysing at low levels of disaggregation. The results from comparison of the 100% and 10% samples in the full Census sample are worrying: regression of comparable employment rates from the 100% sample on the 10% sample gives a coefficient of 0.348, implying that some 65% of the sample variance in employment rates is noise! Fortunately, the NCDS matched sub-sample is biased toward higher population wards. Repeating on this sample gives a regression coefficient of 0.83. Some 17% of the sample variation may attributable to sampling error. The maximum variance attributable to measurement error in the education variable can be derived, for a given sample size. The underlying qualifications variable is dichotomous and the maximum standard error of the mean in a Ward occurs when all the true variance is within-Ward. Since the proportion of the population with A levels or degrees in 1971 was 0.114, the maximum variance within-Ward is 0.101. Based on samples of only 75 over-18s (the median in the Census), the ratio of the square of the standard error of the mean to the actual variance in the data is around 0.22. By this calculation, at most, 22% of the variance is noise. Regression of 100% on 10% derived measures of unemployment rates and employee/total employment ratios from the 1991 Census also suggests that the Ward-level 10% samples give estimates that are 20% down on their true value. Assuming these figures are correct, we would expect any regression coefficients on this to be downward biased by at least 17%. Additional correlated variables in the regression will attenuate the coefficient further.

2.11 Appendix D: Non-socially housed group results

For completeness, Table 2-13 shows comparable results for the non-socially housed group. As expected, given the sorting of this group into neighbourhoods along educational lines, the basic coefficients are much higher than for the socially housed group.
Table 2-13: Neighbourhood education effects on non-social tenants, age 33 qualifications

<table>
<thead>
<tr>
<th>Qualification group</th>
<th>I</th>
<th>II</th>
<th>III</th>
<th>IV</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Low High</td>
<td>Low High</td>
<td>Low High</td>
<td>Low High</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Mean</td>
</tr>
<tr>
<td>Ward education</td>
<td>-0.576 1.02</td>
<td>-0.317 0.553</td>
<td>-0.459 0.820</td>
<td>-0.268 0.468</td>
</tr>
<tr>
<td>School quality</td>
<td>- - -</td>
<td>-0.199 0.347</td>
<td>-0.008 0.014</td>
<td>-0.017 0.030</td>
</tr>
<tr>
<td>Early attainments</td>
<td>- - -</td>
<td>-0.567 0.990</td>
<td>-0.352 0.615</td>
<td>-0.309 0.539</td>
</tr>
<tr>
<td>Missing</td>
<td>- - -</td>
<td>-0.352 0.615</td>
<td>-0.309 0.539</td>
<td>0.137 0.428</td>
</tr>
<tr>
<td>Male</td>
<td>-0.007 0.012</td>
<td>-0.004 0.006</td>
<td>-0.008 0.014</td>
<td>0.017 0.030</td>
</tr>
<tr>
<td>City Ward</td>
<td>-0.009 0.016</td>
<td>-0.013 0.022</td>
<td>-0.012 0.022</td>
<td>0.101 0.176</td>
</tr>
<tr>
<td>Male</td>
<td>-0.012 0.021</td>
<td>-0.013 0.022</td>
<td>-0.012 0.022</td>
<td>0.101 0.176</td>
</tr>
<tr>
<td>LA agriculture</td>
<td>0.095 -0.168</td>
<td>0.101 -0.176</td>
<td>0.105 -0.188</td>
<td>0.105 -0.188</td>
</tr>
<tr>
<td>LA unemployment</td>
<td>1.075 -1.191</td>
<td>0.829 -1.447</td>
<td>1.210 -2.163</td>
<td>0.802 -1.400</td>
</tr>
<tr>
<td>County effects</td>
<td>2.69 = 91.61, P = .0054</td>
<td>2.69 = 111.70, P = .0001</td>
<td>2.69 = 101.30, P = .0007</td>
<td>2.69 = 110.03, P = .0000</td>
</tr>
<tr>
<td>Exogenity test</td>
<td>- -</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Predicted group prob.</td>
<td>0.137 0.428</td>
<td>0.137 0.428</td>
<td>0.137 0.428</td>
<td>0.137 0.428</td>
</tr>
<tr>
<td>Log likelihood</td>
<td>-4581.57</td>
<td>-4015.58</td>
<td>-4747.91</td>
<td>-4039.08</td>
</tr>
<tr>
<td>Pseudo R²</td>
<td>0.034 0.153</td>
<td>0.020 0.148</td>
<td>0.020 0.148</td>
<td>0.020 0.148</td>
</tr>
<tr>
<td>Sample size</td>
<td>4749</td>
<td>4749</td>
<td>4749</td>
<td>4749</td>
</tr>
</tbody>
</table>

Three category ordered probit estimates.

Predicted neighbourhood education uses the proportion of Local Authority tenants in the cohort members' age-13 Ward as instrument.

Inclusion of parental education gives marginal effects of -0.35+/+0.60 in Column (1) (t = 7.50); inclusion of parental interest dummies in Column (1) gives -0.26+/+0.46 (t = 5.892).

School quality is proportion of own sex, age 15 studying for GCEs, instrumented by school type and pupil-teacher ratio to remove catchment area effects on school quality (but not peer group effects or correlation between pupil performance due to selection by schools or by parents). Early attainments measure tests sensitivity to selection on ability to high performing schools, by schools or parents – without early attainment control estimates in Column (4) are -392, .666 for school quality parameter, -326, .554 for neighbourhood parameter.

Instruments have p-value < 0.01% in prediction equations.

- 88 -
2.12 Appendix E: Method for Figure 2-1 to Figure 2-3

The model is

\[ h_i^c = g(h_i^n) + \tilde{\beta}'f_i + \omega_i \]

An estimate of \( \tilde{\beta} \) is obtained as:

\[ \hat{\beta} = \sum_i \left( \tilde{f}_i \tilde{f}_i' \right)^{-1} \cdot \sum_i \tilde{f}_i h_i^c \]

where

\[ \tilde{f}_i = f_i - m_b(f_i | h_i^n) \]

\[ \tilde{h}_i^c = h_i^c - m_b(h_i^c | h_i^n) \]

and \( m_b(\cdot) \) is an estimate of the conditional mean obtained by kernel regression. An estimate of the function

\[ g(h_i^n) \]

is then obtained by a second stage kernel regression of \( h_i^c - \hat{\beta}'f_i \) on \( h_i^n \), at preset sequence of points \( \{h_i^n\} \).

2.13 Appendix F

The estimator is the Nadaraya-Watson estimator with a bivariate Gaussian kernel, estimating the conditional mean \( E[\tilde{h}_i^c | \tilde{h}_i^n, \tilde{h}_i^p] \) on a 40 × 40 matrix of regressor grid points. Since we do not have outcome, parental and neighbourhood measures on the same metric, the regression uses scores constructed from the raw variables based on their rank in the empirical distribution. The neighbourhood educational score is the rank of the neighbourhood in the distribution of Ward-proportions with high qualifications. We generate the parental educational score by ranking mother's and father's years in education. Where there are ties, rank is within educational categories according to social class in 1974. Mother's and father's educational scores combine to make the family ranking. The outcome educational score is based on ten ordered qualifications categories, with ranking within categories by time in education. The scores are all normalised so that they give the ranking on a scale of zero to one.
2.14 Appendix G

Any discussion of neighbourhood effects from adult outcomes presupposes that neighbourhoods differ in their composition, in terms of the characteristics of residents. Figure G-1 shows to what extent the educational composition of residents varies across Census wards in Britain, and charts the changes that have taken place between the 1971 to 1991 Census. Both panels show kernel density estimates of the distribution of the proportion of qualified residents across Census wards. The proportions are computed as the proportion of qualified residents aged 18 and over for 1971 and 1991. Panel 1 compares the 1971 and 1981 distributions. Panel 2 compares the 1981 and 1991 distributions.

A couple of points should be borne in mind when interpreting the diagrams. Firstly they are based on the Census numbers taken from the 10% Census samples, which are subject to sampling error, so the observed variance may overstate the true variance across wards. Secondly the 1971 and later Census years are not strictly comparable, because of major changes in the Census geography. Variation in the level of disaggregation will change the observed variance, even if there were no underlying changes in the degree of segregation. In the extreme case of observations on individuals, the distribution would be bi-modal with peaks at 0 and 1. Wards in 1981 and 1991 were, on average, larger than in 1971, so the estimated density function will appear compressed relative to the density based on 1971 wards.

Looking at the figures, there has been a flattening and rightward shift in the distribution of qualifications across wards over time. The density estimate for the proportion with degrees and similar qualifications in 1971 is sharply peaked relative to the curve in 1981. In 1981, the distribution of degrees across wards was closer to the distribution of A-levels and higher qualifications in 1971 – evidence of the general upgrading of qualifications that has taken place over the last decades.

Standard measures of inequality suggest that this observed spread in the density does not reflect increasing inter-Ward inequality attributable to increasing segregation along educational lines. In 1971 the mean proportion of residents with degrees and similar qualifications was only 6.4% with a standard deviation of 5.4%. By 1981, the proportion with these qualifications had increased to 10.7%, with a standard deviation of 6.6%. By 1991, the proportion was up to 14.3%, with a standard deviation of 8.2%. Using the coefficient of variation as a measure of inter-Ward inequality, this implies a decrease from 0.84 to 0.62 to 0.57 over the

---

24 No figure for the number of residents age 18 and over is available for the 10% sample in 1981. The denominator used here is the number of residents in the 10% sample, weighted by the proportion of residents aged 18 and over, taken from the 100% sample.
years covered by the Census. An alternative measure, which may be preferable in this case where the
distribution is highly skewed, is the 90th/10th percentile ratio. This tells a similar story, with a decrease from
7.4 in 1971 to 5.79 in 1981 to 5.1 in 1991. Even if we disqualify the 1971 Census because of the different
Census geography, there is no evidence of increasing inequality in the ten years between 1981 and 1991.

In fact, the change in the distribution can be attributed solely to the general growth in the proportion of
qualified individuals from one cohort to the next. If we re-scale the numbers by dividing by the mean
proportion in each year, so that the means are normalised to 1, we find that the empirical densities for 1981
and 1991 are almost identical. The 1971 line is also fairly close – see Figure G-2. A constant multiplicative
increase in the qualified proportion in each Ward widens the distribution because wards with few qualified
people show a smaller absolute increase than wards with many qualified people. By contrast, if there is
increasing segregation, then wards at the lower end of the distribution will show smaller proportional
increases than those at the mean, because individuals with low qualifications migrate to these wards and more
educated people leave. Wards at the top end would show larger proportional increases than average as highly
educated individuals became concentrated in these wards over time.

There is no evidence that increased segregation is a significant phenomenon in Britain, based on this Ward
level Census data on qualifications. But segregation could increase over the years within wards, whilst
leaving the distribution across wards unchanged. Unfortunately, the output of the qualifications variables
from the Census at the more disaggregated Enumeration District level is of little use due to the high sampling
errors from the 10% sample and cannot be used to check this. It is also plausible that convergence between
broader geographical groups may have obscured increases in inter-Ward inequality. To test this, Figure G-3
shows the kernel densities for 1981 and 1991, based on the deviation of the Ward proportion with
qualifications from the mean in the district to which the Ward belongs (there are around 500 Census districts
in Great Britain). The top Panel shows a widening of the distribution across wards, within districts. If,
however, we scale the data to allow for the change in the national mean proportion of qualified residents – see
Panel 2- we find little visual evidence of any underlying change across wards.25

---

25 If we do the same using characteristics which are measured accurately at ED level – residential
overcrowding or unemployment rates – there is still no evidence of increased spatial inequalities.
The kernel densities for the ED proportions in unemployment and persons per room in 1981 and
1991, or the deviations of these from Ward means, are almost exactly overlying.
Figure G-2: Changes in the spatial distribution of education

a) 1971 and 1981


b) 1981 and 1991

Distribution of proportion with A-levels, degrees, diplomas and professional qualifications shown for 1971.
Figure G-2: Changes in the spatial distribution of education, mean adjusted

a) 1971 and 1981

b) 1981 and 1991


Distribution of proportion with A-levels, degrees, diplomas and professional qualifications shown for 1971.
Figure G-3: Changes in the within-district spatial distribution of qualified residents

a) 1981 and 1991

b) 1981 and 1991, mean adjusted


Distribution of proportion with A-levels, degrees, diplomas and professional qualifications shown for 1971.
3 Geography, Resources and Primary School Performance

3.1 Introduction

Primary schooling plays a pivotal role in the generation of human capital and the determination of life chances. We can see that parents believe this by the premiums paid for homes that are close to good state schools, or the fees for private preparatory and pre-preparatory schooling – around £6500 per year on average in England in 2001. Primary age achievements feature strongly in empirical accounts of the determination of individual attainments and other outcomes. The school effectiveness literature has acknowledged the importance of primary schools on achievement in secondary school, but yet the primary phase remains relatively under-researched. A particular feature of primary schools in England is that they serve very localised communities, and many schools can be found within relatively narrow geographical areas, at least in the urban environment. This close link between communities and primary schools, and the close proximity of urban schools, means that their role in mediating community influences to individual attainments may be especially strong. On the one hand, we can expect the distribution of school success to follow closely the distribution neighbourhood disadvantage at the micro-geographic level – even more so than for secondary schools which serve far wider catchment areas. But also, the close proximity of primary schools

Figure provided by the Independent Schools Council Information Service


Goldstein and Sammons (1994) attribute up to three times as much of the variation in age-16 exam performance to junior school of origin than to the secondary school attended
and non-exclusivity of local catchment areas lends itself well to interactions between schools and between children who are neighbours but attend different schools. These interaction effects, if they exist, may encourage the diffusion of educational technologies and community-based advantages across geographical space. Like neighbourhood externalities in general, these kinds of interaction effects can imply increasing returns to neighbourhood-targeted policy initiatives. At a more abstract level, they hint at externalities that may motivate city formation and social interaction in general.

The main focus of this Chapter is on estimation of the relationship of location, local interactions and community to primary school performance. Any analysis of these issues would, however, be incomplete without consideration of the role of fundamental school-level ‘economic’ inputs – teacher/pupil ratios, expenditures – and the importance of these factors relative to ‘community’ inputs. All of these are central concerns for educational policy, both in terms of the efficacy of extra resources in promoting student achievement, and in terms of the targeting of these resources across geographical and socio-economic space. The substantial body of related literature, spread across educational research, economics, sociology, developmental psychology, urban geography and other disciplines, is evidence of the weight attached to these issues. In Britain, much of this research has been limited to relatively small-scale case studies, cross-sectional analyses, or has been confined to secondary school achievement as measured by pass rates on national examinations at school leaving age. The availability of primary performance tables since the mid 1990s, based on tests in the national curriculum in England, combined with a wealth of school-level information from the Department of Education and Skills (DfES), now provides a basis for investigation of primary schools at the national scale. By linking this data, through Postcodes and national grid references, to

29 Six years are now available from 1996 to 2001, but the accounting rules were changed in 2000, meaning that the results after 1999 are not necessarily consistent with earlier years
Census data from 1991 and to more recent property price, incomes and unemployment data, we can obtain accurate measures of the characteristics of the neighbourhoods in which schools are located. The proximity-based rules by which admissions are usually prioritised ensure that community characteristics are good proxies for characteristics of children at the schools, and of the schools' geographical and community context.

Estimation of the causal influence of school resource inputs is notoriously difficult. On the one hand, educational policy that allocates more resources to failing schools or disadvantaged pupils can lead to inference of negative input-output relationships. On the other hand, we might falsely infer positive effects, or measure upward biased positive effects for a number of reasons. There may be more, better quality teachers in high-performing schools if recruitment and retention is easier. More able or advantaged children might find their way into schools with more or better teachers due to the exercise of parental preferences over class size and teacher quality. If school inputs are linked to local taxation and property prices then pupils with more home-based resources will find themselves in better-resourced schools. Instrumental variables or control function approaches to overcoming these sources of estimation bias are unlikely to succeed, since there are virtually no plausible instruments. Instead, the approach adopted here is to assume that any policy or selection decision is based on persistent, unobserved school characteristics originating prior to the sample, so that fixed-effects estimation, or estimation conditional on prior performance is appropriate.

This paper steps back from investigation of the influence of the specifics of the school environment on individual attainment, and instead uses school level data to assess the impact of school location on average attainment. What we do not do here is address the detail of factors in school effectiveness or the differential impact of these factors on pupils with different needs. These issues have preoccupied school effectiveness research, where the focus is often on potential (rather than average) achievements and on the heterogeneous impact of organizational, leadership and instructional structures. Our
school level data precludes such investigation. In any case, individual level data offers no advantages unless explanatory variables are child-specific – personal learning time, and own class size for example – and this introduces a whole new level of endogeneity bias which needs to be overcome.

The structure of the Chapter is as follows. We start below with a brief survey of the related literature. In Section 3.3 we set out the empirical framework we use for estimating the impact of local conditions and school resources on performance. Section 3.4 outlines our data sources and the procedure for matching area characteristics to schools. Next we get on to the presentation and discussion of our empirical results in Section 3.5, starting with maps to illustrate the distribution of school performance across England, and descriptive indicators of the spatial associations in the data. We then move on to analyse the determinants of primary school attainments using the production function model of Section 3.3. Section 3.6 concludes with a summary, and an assessment of whether more school resources can overcome the adverse impact of area disadvantage highlighted in the earlier results.

3.2 A review of related work

The debate over the relative importance of school inputs and the socio-economic background of pupils and the community context of the school has remained largely unresolved since the findings of the Coleman report (Coleman and others (1966)) that school composition mattered far more than school resource inputs. These findings motivated over three-decades of school effectiveness research, which has, in the UK at least, focussed largely on the secondary phase. Even then, evidence on the impact of neighbourhoods and intake characteristics – beyond the impact of pupil free school meal entitlement – is limited. To the author’s knowledge, no study in the UK has looked in detail at the relationship between neighbourhood characteristics and primary school performance.
Even the basic relationship between teaching inputs and pupil attainments is relatively unexplored in the primary phase, although the literature on secondary school resources and outcomes, and on secondary school effectiveness and improvement is vast — see, for example, Burtless (1996) for the US, and White and Barber (1997) and Vignoles, et al. (2000) for UK surveys. Key recent works on the effect of school resources on secondary exam results in the UK, using national, pupil level data, are all based on the National Child Development Survey (Dustmann, et al. (1998), Dearden, et al. (2002), Feinstein and Symons (1999) and Dolton and Vignoles (1999)). Pupil teacher ratios have insignificant effects in most of these studies, and expenditure per pupil (at LEA level) is insignificant in those studies that include it. A recent database of national pupil A/AS-level results provides evidence of differences between institution types in overall attainment at upper-secondary level and in value-added performance relative to GCSE attainment (Yang and Woodhouse (2001)), but specific resource inputs are not tested. The authors find that, at most, 22% of the variance in A-level scores is attributable to establishments.

Only a few studies address the issue of community context empirically, even at secondary level. MacCallum and Redhead (1999) map 1991 Census data to residential addresses of just under 2000 pupils in 12 High Schools in the London Borough of Ealing in 1998. In their report summary, the authors report only simple correlation coefficients with no multivariate analysis. The proportion of higher-educated residents in the residential Enumeration District has the strongest correlation (0.215, p-value<0.0001) with individual Key Stage 3 attainments, and the proportion of ethnic minorities is most strongly associated with GCSE point score. Unsurprisingly, the correlation coefficients are much larger using school level aggregates (since the variance of unobserved determinants of attainments is reduced), up to 0.944 for the association of higher-educated neighbourhoods with Key Stage 3 maths scores for white pupils.
Lupton (2001) finds that 69% of the secondary schools placed in the Office for Standards in Education’s (OFSTED) lowest quality Special Measures category were in the top-ten percent of deprived wards ranked according to the Department of Environment, Transport and Regions’ (DETR) 2001 Index of Multiple Deprivation. Only 11 percent are in the least disadvantaged 40% of wards. Bradley and Taylor (1998) report large and highly significant negative effects from school-level free-school meal entitlement, but surprisingly weak effects from ethnic minorities and pupils with special educational needs on secondary GCSE performance. Using a Data Envelopment Analysis approach, Bradley, et al. (1999) find that more efficient secondary schools in terms of exam performance and attendance rates are in Local Authority districts with high proportions of professional and managerial residents, although also in areas with high unemployment rates. Close proximity to other non-selective schools also increases efficiency, a feature that the authors attribute to competitive effects of the quasi-market in secondary education, although they do not consider other potential forms of spatial interaction.

Using Census data and 1994 GCSE pass rates aggregated to Local Education Authority (LEA) level, Gordon (1996) argues that the proportion of non-employed lone parents is the most significant factor in generating spatial disparities in overall and higher-level (Grade A-C) pass rates, and that this is mediated entirely via unauthorised absence rates. Significant factors in higher-grade success rates, conditional on unauthorised absences, are the LEA proportion of non-manual workers (positive), unskilled (negative), non-earners (negative), ethnic group (positive for Indian, negative for Afro-Caribbean) and residential overcrowding (negative).

Variations in the average achievements of pupils by ethnic groups are important in overall inter-LEA performance inequalities. Nevertheless, each of the six ethnic groups studied in Gillbom and Mirza (2000) is the highest attaining group in at least one LEA, and the relative success by ethnic group has changed over the 1990s, and varies with
phase of education. Indian pupils have overtaken white pupils in GCSE results since 1991, but the position of Black pupils has not improved and black pupil attainments appear to decline relative to the mean with each stage of assessment during compulsory schooling. The same report highlights the persistent differences between social classes, within ethnic groups, but also distinct ethnic disadvantage within social class category for Blacks, Pakistani and Bangladeshi pupils.

Most reports on the influence of area disadvantage on school performance highlight the relationship of results to free school meal entitlement or receipt (for example OFSTED (2000)), since this is collected at school level in administrative data. The downward trend in performance as the proportion of the school intake on free-school meals increases is a striking empirical regularity at primary and secondary level. Although this result is unsurprising, the mechanisms that drive it are not well understood, and some schools perform better than others even with similarly disadvantage intakes. However, no single organisational recipe guarantees success (OFSTED (2000), Lupton (2001)).

In one of the few studies directed specifically at primary schools, Mancebon and Molineri (2000) estimate Key Stage 2 production efficiency using Data Envelopment Analysis on a sample of 176 primary schools in Hampshire, Southampton and Portsmouth. Only free-school meal entitlement appears as an input, and English and Science results as outputs (they reject other inputs in their model selection procedure – including special needs, pupil teacher ratios, gender mix and expenditure per pupil). Regression of the logits of the efficiency scores on a number of other explanatory variables indicates that Church of England schools are substantially more efficient than others, and that parental interest has a role to play – schools for which more parents responded to OFSTED surveys do better, as do those schools of which parents expressed a favourable opinion. The case for an effect of parental opinion on school efficiency, rather than vice-versa, is not made clear.
3.3 The primary school production function

The base-line strategy in this Chapter is to estimate an aggregate school-level production function that describes the probability of a pupil, drawn at random from the school, reaching the target grade in final year tests. We assume that this depends on teaching inputs to the school, local geographical interaction effects, on the characteristics of the pupil, on his or her family background and on unobserved school effects. Individual pupil characteristics are unobserved at school level, as are family background characteristics. But, using geographical data on neighbourhood composition, and information on the spatial location of the school, we can estimate the probability that a child’s household will have characteristics \( x_i^h \). Our general empirical representation of the primary school production function is:

\[
\Pr_{s,t} = \alpha_i + \bar{p}_j \beta + x_i^h \gamma + u_{s,t} + \eta f_j + \epsilon_{s,t}
\]  \hspace{1cm} (3-1)

where \( \Pr_{s,t} \) is a measure of success rates in school \( s \), \( \alpha_i \) is a general time effect, \( \bar{p}_j \) is the average measure of success in the nearest \( J \) schools in area \( j \). For simplicity, \( \bar{p}_j \) is specified a constant over time (for the number of periods in the sample), implying that average neighbourhood school performance does not change in the short run. The term \( u_{s,t} \) captures persistent unobserved school-specific factors, such as teaching quality, and we assume this follows an AR(1) process. The stochastic component \( \nu_{s,t} \) implies innovations to performance that persist over time (a change in teaching staff quality, for example):

\[
u_{s,t} = \rho \nu_{s,t-1} + \nu_{s,t}
\]  \hspace{1cm} (3-2)

We incorporate unchanging neighbourhood and catchment area influences by \( f_j \), an unobserved local area effect. The term \( \epsilon_{s,t} \) is a random school-year-specific disturbance – a cohort effect perhaps, or just noise – and this does not persist over time.
The vector $\mathbf{x}_{s,t}$ incorporates our observations or estimates of the expected characteristics of a child at the school at time $t$, and his or her family background. Parameter $\beta$ captures local interaction effects, $\gamma$ is a vector of parameters measuring the impact of school-specific inputs and the family characteristics of pupils, $\rho$ parameterises the persistence of performance over time, conditional on other observed characteristics and $\eta$ the degree to which neighbouring schools share the same area effects. Setting $p_{s,t} = \Phi^{-1}(\pi_{s,t})$, where $\pi_{s,t}$ is the proportion of pupils achieving the target grade, means that the production function can be estimated as a grouped probit$^{30}$.

The serially correlated components of the error term in (3-1) means that ordinary least squares estimates are inconsistent if the regressors are correlated with past school performance. Neighbouring school performance may be correlated with the composite error term in a number of ways. Firstly, school-to-school feedback mechanisms imply that anything unobserved in the determination of school performance $p_{s,t}$ will also influence neighbours performance $\bar{p}_{j,t}$. More importantly, estimates of $\beta$ will pick up the effects of unobserved components of neighbourhood composition, which are, in part, shared by school $s$ and its nearest neighbours. Regardless of the source of endogeneity, instruments for $\bar{p}_{j,t}$ are available, in that $\bar{\mathbf{x}}_j$, the mean of the characteristics of the school’s nearest neighbours, is correlated with $\bar{p}_{j,t}$ but not with $f_s$, nor with other unobservables (conditional on the characteristics of school $s$).

This identification of structural dependence of school performance on neighbouring school performance – endogenous neighbourhood effects, to use the terminology of (Mansi (1993)) – rests on the assumption that there are no exogenous or

$^{30}$ By weighted least squares using weights $\left\{ \frac{1}{\sqrt{\Phi^{-1}(\pi_{s,t})}} \right\}^{0.5}$
contextual neighbourhood effects so school performance is not affected by the observed determinants of school performance in neighbouring schools. The distinction is perhaps not so important in the current application since we wish only to identify the interdependence of schools, as distinct from their mutual dependence on outside influences. To make this clearer, we do not necessarily wish to rule out the impact on school performance of teaching techniques in a neighbouring school, whether or not this spillover is mediated via test pass rates in the neighbouring school.

Serious problems arise once we try to get consistent estimates of the parameters on most types of school and neighbourhood inputs, for which there are no plausible instruments. Local government targeting of resources on the basis of needs, and parental choice of school and residential location, means that many of the factors that we expect to influence school performance are determined by prior performance, and hence correlated with the term $u_{s,t-1}$. Classes may be smaller in poor performing schools that have low demand for places, or schools in difficulty may receive more teachers and resources. Resources are allocated to Local Education Authorities according to educational needs, and a similar redistribution occurs to some extent within LEAs. Property prices and incomes will be higher close to schools that perform well due to parental competition for places (see Gibbons and Machin (2001)). We can, however, apply a Cochrane-Orcutt transformation to the model to remove the serially correlated error term, and estimate the model conditional on prior performance. Because the performance of pupils taking tests in year $t$ is not specifically related to school characteristics in year $t$, we treat most school inputs and any time varying local characteristics as fixed in the short run at $x_s$, measured by the average over the sample periods. The assumption here is that $x_s$ is determined by a pre-sample value of $u_s$. Our model is now:

$$p_{s,t} = \rho p_{s,t-1} + \alpha_t - \rho \alpha_{t-1} + \bar{p}_j \beta \sigma + x'_t \gamma \sigma + f_{s,t} \sigma + \nu_{s,t} + \epsilon_{s,t} - \rho e_{s,t-1}$$  \hspace{1cm} (3-3)
in which $\sigma = (1 - \rho)$. The presence of the lagged disturbance term and the inclusion of the lagged performance term as a regressor imply that OLS estimation of (3-3) will be biased. IV estimation using $p_{s,t-2}$ as an instrument for $p_{s,t-1}$ gives consistent estimates.

The reasoning is similar to Anderson and Hsiao (1981). In their paper however, the lagged error term appears as a result of first-differencing a model with a lagged dependent variable so as to remove time-invariant fixed effects.

3.3.1 Family background or neighbourhood interaction effects

A non-zero $\beta$ in model (3-3) indicates area-based interaction effects between schools, between pupils or parents of children at different schools in close proximity. The kinds of interaction we have in mind are educational technology spillovers between schools, or learning process-based peer group effects amongst pupils associated with schools in close proximity. The mix of social and non-socially housed tenants in most primary schools suggests an alternative way of getting at community or neighbourhood effects in primary school performance. Owner-occupier and private tenants' demand for property and local amenities means that property prices and the physical characteristics of owner-occupied or private rented property are determinants of the mean local incomes of owner-occupiers and private tenants – and hence education and other related family resources of these groups. These factors do not, however, influence local social tenant incomes or resources, whose allocation to homes is unrelated to the process which sorts private market tenancy groups into high and low income neighbourhoods. Essentially, property prices are a proxy for wealth and other resources in the home-owner community, so relate directly to the probability that the child of a home-owner attains target grades in primary school age tests. Property prices will also change across geographical space with home-owners' perceptions of the neighbourhood and community, both its amenity value and its value as an input to the production of human capital. But property prices have no direct bearing on the incomes or other family resources of social tenants.
We can write the probability of a child reaching target achievement in primary school as:

\[ p_s = p^c + \left( p^0 - p^c \right) \theta^o_s + \omega_s \]

(3-4)

where \( p^o \) (\( p^c \)) is the probability that a non-social tenant child (social tenant child) reaches the target grade, and \( \theta^o_s \) is the proportion of non-social tenants in school \( s \). If we assume that the deviation of the expected attainment of a non-socially housed child from that of a social tenant's child is linearly related to local property prices (or a transformation of these prices) \( z_s \) then (3-4) becomes

\[ p_s = p^c + (\lambda_1 + \lambda_2 z_s + \nu_s) \theta^o_s + \omega_s \]

(3-5)

If \( z_s \) is exogenous, then (3-5) may be estimated by weighted-least-squares (allowing for the heteroscedastic error term). A test for general effects from \( z_s \), not specific to home-owners or private tenants, is to estimate

\[ p_s = p^c + \lambda_0 z_s + \lambda_1 \theta^o_s + \lambda_2 z_s \theta^o_s + x_s \beta + \omega_s \]

(3-6)

and test for \( \lambda_0 = 0 \). Semiparametric estimation of the regression surface of school performance on \( \theta^o_s \) and \( z_s \), conditional on \( x_s \), can also provide insights because we can observe the response of school performance to \( z_s \) at different points in the within-sample distribution of \( \theta^o_s \).

There are some obvious objections. Firstly, all interaction-based neighbourhood effects could operate through peer groups in the school itself. In such a case, there will be no general effects which are uncorrelated with the proportion of non-social tenancy children in the school, since \( \theta^o_s \) equal to zero means that the mediating peer group is non-existent. Secondly, social tenants who are more motivated towards their child's upbringing may self-select into accommodation in neighbourhoods with higher property prices or larger owner-occupier properties, though this is unlikely given the restricted
availability of council housing. Lastly, the incomes and other resources of neighbouring
social tenants may be objects of preference in owner-occupiers’ and private tenants’
housing demands, leading to correlation between property prices (and hence owner-
occupier incomes) and social tenant resources, even conditional on the proportion of
social tenants. We can test this last assumption using data on incomes of social and non-
social tenants living in close proximity. It is also worth noting that selection by owner-
occupiers on school performance itself cannot be factor driving a price-performance
relationship at the zero owner-occupier school admission rates which correspond to the
main effect of property prices $\lambda_0$.

3.4 The data, and matching methods

Our core dataset, provided by the Department of Education and Skills (DfES),
contains basic school characteristics and performance based on age-11 Key Stage 2
assessment tests. These tests are common to all schools, and assess progress through the
National Curriculum. Age-11 pupils are expected to reach Level 4 in these tests, and the
performance figures give the proportion in each school reaching this grade. Importantly
for the spatial emphasis in the current work, the data set includes location identifiers – the
region, parliamentary constituency, Postcode and National Grid Reference of the school
premises. Our central aim is to match available local area data to these school locations,
and to infer the relationships between performance of nearby schools, and between school
performance and local area characteristics. Our local area data comes from a number of
sources. Firstly, the 1991 Census provides detailed information at a highly disaggregated
level (down to a few hundred households in each Enumeration District). A commercial

\[31\] Property prices are certainly influenced by the proportion in social housing though the issue here
is whether property prices are correlated with social tenant resources, conditional on the
proportion.
data set of Postcode sector mean incomes for 1996 and 1999 provides more up-to date, but more aggregated, income measurements and household numbers. A Postcode sector typically contains 2500-3000 households. At the same level of aggregation, we have annual mean property prices from the Government Land Registry from 1995 on, plus measures of unemployment benefit claimants per household derived from NOMIS Postcode sector claimant counts for each year.

Key Stage Two assessment tests are taken in the spring, in what is normally a child's final year at primary school – generally at age 10-11. Children taking the tests in 1996 would have been aged 5-6 on the night of the Census in 1991. Children taking the tests in 1999 will have been 2-3 years old. Given this, the data from the Census in 1991 should be interpreted as characterising the neighbourhood community around the time of the children's entry into school. As central-tendency measures of the community in which pupils ordinate, this will be noisy to the extent that: 1) the characteristics of the school intake differ from those of the residents in the immediate vicinity, or 2) the classes sitting the tests from 1996 to 1999 are composed of children who have moved in to the area, and exclude those who have moved out, since the Census. Census-based area measures do, however, have a distinct advantage over information on school-year-level economic and social characteristics in that they are more plausibly pre-determined. School composition will change in response to parental selection on performance in previous years. Community composition in 1991 is sensitive only to school performance prior to 1991, and then only to the extent that movements in families with pre-school age children changes the mean composition of the neighbourhood.

More recent age data at Postcode sector level imputed by another marketing company, Experian, was found to offer no advantages over the Census equivalent. The correlation between the Experian and Census data is near unity.
To maximise the precision in measurement of the school’s catchment area characteristics, we apply a \( K \)-nearest neighbour approach to matching the Census data, property prices and local incomes to schools locations. In practice this involves assembling all the community data as means or proportions at Census Enumeration District (ED) level. We then estimate a school centred, school-age cohort weighted average of the data in the nearest \( K \) neighbourhoods around each school, at a set of locations defined by a matrix of school Postcode, 10 metre grid references. By way of illustration, Figure 3-1 shows the catchment area that we would obtain by matching eight EDs to one school in Hackney, East London.

Figure 3-1: Illustrative pseudo-catchment area matching on Census Enumeration Districts for a school in Hackney, London

![Scale: Circle diameter = 0.8km](image)

To refine the data matching still further, we adapt \( K \) according to the number of children in the school year, and a first-pass estimation of the number of school-entry-age
children in the nearest $K_0$ Enumeration Districts, where $K_0$ is a guess at the typical number of Enumeration Districts comprising a school catchment area. Details of the method are presented in Appendix A.

Our dataset thus has two distinct levels at which we measure intake and community characteristics. Most are derived in this way from the Census and other spatial datasets, and measure community characteristics in the imputed catchment area. We also have limited intake characteristics at school level – proportions of pupils entitled to free school meals, proportions with special educational needs and with statements of special needs and proportions of non-white ethnic groups, plus school type and age range. The strong, graphical relationship between free-school meals entitlement and school performance at both primary and secondary levels is well known – see for example OFSTED (2000). Free-school meal entitlement is often used as a benchmark for comparison of schools that differ in terms of their intake disadvantage.

### 3.5 Empirical results

#### 3.5.1 Clustering and the spatial distribution of primary performance

We start with a visual assessment of the distribution of primary school performance in England. The geographical distribution and spatial association of performance in primary schools give direct clues to the spatial processes underlying attainments. If schools are more often than not located near other schools with similar pupil achievements, then we might suspect a role for geographical factors – institutional differences between areas, differences in school intake, or endogenous effects based on interactions between local schools. Looking at some maps also reveals some interesting features about city versus rural school performance. Figure 3-2 illustrates the geographical spread across England at County level in 1999. Counties correspond to Local Education Authorities (LEAs) in most non-metropolitan areas.
Figure 3-2: a) Quintiles of Key Stage 2 scores, by English County, 1999

- 82.0 to 76.7%
- 76.7%
- 71.7%
- 71.6 to 63.1%

b) Quintiles of KS2 scores, adjusted for free-school meals 1999

- +12.0 to +1.8%
- +1.8%
- to
- -2.9%
- -2.9 to -6.9%
Panel a) shows the quintiles of County mean Key Stage 2 performance scores (weighted by the number of pupils tested). The top fifth of counties are concentrated in the south east and east of England – Surrey, Berkshire, Hertfordshire, Hampshire and Cambridgeshire – in Yorkshire in the north east, and Cheshire and Merseyside in the north west. The worst counties in the south east form an arc around the Thames estuary: Greater London, Essex, and Kent. The West Midlands conurbation and several counties in northern-central England are also in the bottom one-fifth.

These differences could reflect any number of underlying factors. A prime suspect must be the disadvantage of the school intake. Indeed, once we adjust for the proportion eligible for free-school meals\(^{33}\), we see a quite different pattern. Predominantly urban counties increase their ranking: Greater London and Tyne and Wear are now in the top quintile. A notable feature is that schools in a swathe across middle England that were in the top half of the ranking, fall to the bottom once we allow for the fact that more pupils there are from higher-income backgrounds. What is also striking, particularly in the case of London, is that an urban environment is *not* in itself a necessary condition for poor average school performance at the County level. A high proportion of low-income pupils can be a sufficient condition. We should also note that rural areas can perform well: the Isle of Wight, Cumbria and Cornwall are all in the top ranking of adjusted scores.

Counties present too broad a level of aggregation to tell us much about spatial clustering of performance, and mask deep local inequalities in performance. Figure 3-3 presents a more disaggregated picture, with quintiles of performance aggregated to Parliamentary Constituency level\(^{34}\).

\(^{33}\) Residuals from a regression of KS2 scores on a polynomial in free-school-meal entitlement

\(^{34}\) Although this has no natural correspondence with schooling, it is the only area classification for which it is straightforward to aggregate school performance and draw map boundaries
Figure 3-3: a) Quintiles of Key Stage 2 scores, by Parliamentary Constituency

- 88.8 to 79.6%
- 79.4%
- to
- 67.7%
- 67.6 to 52.8%

b) Quintiles of KS2 scores, adjusted for free-school meal entitlement

- +17.1 to +3.5%
- +3.5%
- to
- -3.5%
- -3.5 to -11.5%
The picture using raw data in panel a) is of good performance in the higher land area semi-rural Constituencies, spreading north from the south coast and running up the middle to west of England. As in Figure 3-2a), we see high-performing areas spreading right across the north. The urban constituencies – particularly in Greater London, the West Midlands – and in pockets in the southern Pennines, perform poorly in the unadjusted data. But the effect of adjusting for school intake incomes is startling: in panel b). The broad mass of high performing Constituencies that we saw in central England becomes a central swathe of poor performance, and the schools that are effective despite the income disadvantages of their intake turn out to be clustered in and around Greater London, around Birmingham in the centre, Newcastle, Durham and Middlesborough in the north east, in the north western Manchester-Merseyside conurbation and its hinterland, and in the central and southern urban areas of the Pennines. Schools in cities do better, once we adjust for income disadvantage. This, of course, says nothing about which city characteristics are performance enhancing – this is analysed further in 3.5.5.

We back up our impressions of clustering based on visual inspection using statistical measures of spatial association. First we use the Moran’s I statistic (see Appendix B), which indicates the fraction of the overall variance in the performance scores that is attributable to spatial clusters of locations with similar school performance. Computing this statistic for Figure 3-3a, we find low, but significant global spatial association in raw Key Stage 2 performance ($I = 0.039, z = 8.727$). Doing the same for the numbers in Figure 3-3b, we find an increase in spatial association ($I = 0.110, z = 18.000$). What we observe here is that spatial association in school productivity is masked in the raw school performance indicators by the impact of poverty in urban areas. We can verify which areas contribute most to this spatial concentration in performance by plotting the z-scores of the Local Moran’s I statistics at for each constituency (see Anselin (1995) for details). We do this in Figure 3-4.
Figure 3-4: a) Clusters of below-median primary school performance (free-school meal adjusted, Local Moran I p-values)

- p-value <0.05
- <0.05 with Bonferroni
- <0.01 with Bonferroni
- <0.001 with Bonferroni

b) Clusters of above-median primary school performance (free-school meal adjusted, Local Moran I p-values)

- p-value <0.05
- <0.05 with Bonferroni
- <0.01 with Bonferroni
- <0.001 with Bonferroni
A high positive z-score for a given constituency implies that there are clusters of similar performing schools in the neighbourhood of that constituency. To clarify the analysis, we split the Local Moran I's into two categories: a) those indicating clusters of bad performance, with positive z-scores and below-median free-school-meal adjusted primary school performance areas and b) those indicating clusters of good performance, with positive z-scores and above-median free-school-meal-adjusted primary school performance. The views for the whole of England are in Figure 3-4. Panel a) identifies two performance hot-spots – one in the North West, and one in the London area. Panel b) highlights areas of concentrated below average productivity in a range of constituencies across middle-England. The shading indicates the p-values of the Local Moran I statistics, with Bonferroni corrections as indicated on the key.

Figure 3-5 shows details of the London area. Clearly, much of inner London is characterised by clusters of neighbourhoods with effective schools relative to the economic disadvantage of their intake – including Constituencies like Bethnal-Green-and-Bow which ranks in the bottom 10% on the basis of raw performance scores, or Brent-East which falls in the bottom quartile.

Visual examination of global spatial clustering of performance at school level and on a national scale is infeasible. Again we can turn to statistical evidence based on summary measures of spatial association. The rank-adjacency statistic provides a non-parametric alternative to the usual Moran’s I indicator of spatial association (Ekawaru and Walter (2001), Walter (1994)). The statistic is based on the sum of the rank differences between each school and its closest associate in geographical space. The greater the similarity between the ranks of schools in close proximity, the lower the statistic, indicating correlation between scores in neighbouring schools.

---

35 Both indicators of spatial association are described in Appendix B
Table 3-1 presents rank-adjacency statistics for mean 1996-1999 Key Stage 2 scores, calculated separately for each Government Office Region. Significance tests are based on a normal approximation. Statistics with z-scores below $-1.64$ are significant at the 5% level in this one-sided test. Although the table presents results for weights which link nearest school-pairs, the results were broadly similar under different weighting schemes (e.g. inverse distance squared, nearest ten schools).

The summary statistics show more spatial association of performance within regions than we would expect if the distribution were random. This is hardly a surprising result if we believe that the background of the pupils drives performance in schools, since the spatial distribution will reflect the underlying distribution of parental resources.
Spatial correlation can also arise from differences across space in resource inputs, LEA policy, staff quality and other institutional factors that affect pupil success. The lower panels of Table 3-1 tests for the importance of these factors by regression adjusting the performance ranks to take out differences between a) local education authority areas b) pupil-teacher ratios, school size, school type (Voluntary, Foundation or Community), average class size and age-range c) the proportion eligible for free-school meals.

Although adjusting for free-school meal entitlement or Local Education Authority differences generally reduces the z-scores\(^{36}\), we reject the null of no spatial clustering of performance in all cases.

### Table 3-1: Rank adjacency statistics, Key Stage 2 performance measures, 1996-1999 mean

<table>
<thead>
<tr>
<th>Region</th>
<th>Unadjusted</th>
<th>LEA adjusted</th>
<th>PTR/type adjusted</th>
<th>Free meals adjusted</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N</td>
<td>D</td>
<td>z-score</td>
<td>D</td>
</tr>
<tr>
<td>NW</td>
<td>2271</td>
<td>564.925</td>
<td>-17.544</td>
<td>600.415</td>
</tr>
</tbody>
</table>

\( ^{36} \) This is in part due to the use of the estimated regression residual.
3.5.2 Associations with catchment area incomes and unemployment

Figure 3-2 and Figure 3-3 illustrated the importance of low-income on the geographical distribution of pupil attainment. We examine this link further, with some fairly descriptive regressions summarising how the income-performance relationship is distributed over time and geographical space. We restrict our attention to local incomes and unemployment, for which we have some time-series variation.

To motivate the analysis, Figure 3-6 illustrates the performance-income relationship, conditional on Local Education Authority, special needs, school type (community, voluntary aided, foundation or voluntary controlled), age-range and rural-urban location. Mean school performance over 1996-1999 rises rapidly as local incomes increase at the bottom and centre of the distribution, but there are diminishing returns to neighbourhood improvements. The lower panel illustrates the relationship between unemployed claimants per household and performance in subsequent years, unconditional on incomes. Note, these are associations rather than causal relationships.

3.5.3 Regional variation

Table 3-2 separates out and parameterises these relationships, by region and year, using a probit model. As before, additional controls are school type and age indicators, special needs measures, urban-rural and local education authority dummy variables. In reading this table, we should interpret income and unemployment as indices of local conditions, since we have done nothing here to separate the effects of these characteristics from most other aspects of family background. Performance rises with catchment area incomes, and the slope of the relationship at the mean is similar in all regions.

---

37 This regression line is estimated by semi-parametric regression to calculate linear coefficients on the control variables (see for example Robinson (1988)), followed by kernel regression of the residuals on incomes.
Figure 3-6: Primary school performance and catchment area characteristics

Incomes:

Unemployment:

Figure shows kernel regression of mean proportion achieving level 4 in Key Stage 2 tests on mean catchment area incomes or proportion unemployed. Additional controls are school-type, age-range, rural-urban, and LEA dummies. Thin lines show pointwise 95% confidence intervals for kernel regression line.

Looking at 1996 in Table 3-2, we could restrict the slope to be equal across regions at 0.428 and would not reject equality at the five percent level (p-value = 0.654). A one percent relative increase in incomes is linked to a 0.43 percentage point increase in
primary school achievement. As we move across the table from 1996 to 1999, the income effect across regions diminishes to 0.261 (p-value for test of equality = 0.198). This is a corollary of the mean improvement in Key Stage 2 achievement rates in all regions, since the improvement is a result of catch-up by schools at the bottom of the distribution of performance in 1996.

Table 3-2: School sensitivity to neighbourhood conditions by region, 1996-1999

<table>
<thead>
<tr>
<th>Region</th>
<th>1996</th>
<th>1997</th>
<th>1998</th>
<th>1999</th>
<th>Sample</th>
</tr>
</thead>
<tbody>
<tr>
<td>North East</td>
<td>Log income 0.379 (6.887)</td>
<td>0.308 (14.356)</td>
<td>0.313 (8.947)</td>
<td>0.226 (6.887)</td>
<td>601</td>
</tr>
<tr>
<td></td>
<td>Unemployed -0.773 (-2.020)</td>
<td>-0.866 (-5.399)</td>
<td>-0.877 (-2.894)</td>
<td>-0.907 (-2.020)</td>
<td></td>
</tr>
<tr>
<td>North West</td>
<td>Log income 0.430 (14.340)</td>
<td>0.366 (13.602)</td>
<td>0.310 (10.687)</td>
<td>0.245 (14.540)</td>
<td>2053</td>
</tr>
<tr>
<td></td>
<td>Unemployed -0.580 (-3.850)</td>
<td>-0.452 (-1.864)</td>
<td>-0.638 (-3.031)</td>
<td>-0.479 (-3.139)</td>
<td></td>
</tr>
<tr>
<td>Yorks and</td>
<td>Log income 0.435 (9.768)</td>
<td>0.318 (6.117)</td>
<td>0.358 (7.976)</td>
<td>0.328 (6.700)</td>
<td>1187</td>
</tr>
<tr>
<td>Humberside</td>
<td>Unemployed -0.545 (-3.140)</td>
<td>-0.575 (-2.373)</td>
<td>-0.342 (-1.579)</td>
<td>-0.189 (-1.405)</td>
<td></td>
</tr>
<tr>
<td>East Midlands</td>
<td>Log income 0.487 (8.127)</td>
<td>0.414 (7.708)</td>
<td>0.457 (9.032)</td>
<td>0.308 (6.779)</td>
<td>932</td>
</tr>
<tr>
<td></td>
<td>Unemployed -0.239 (-0.847)</td>
<td>-0.055 (-0.164)</td>
<td>-0.174 (-0.461)</td>
<td>-0.453 (-1.405)</td>
<td></td>
</tr>
<tr>
<td>West Midlands</td>
<td>Log income 0.490 (12.233)</td>
<td>0.405 (8.108)</td>
<td>0.423 (12.597)</td>
<td>0.340 (10.572)</td>
<td>1255</td>
</tr>
<tr>
<td></td>
<td>Unemployed -0.647 (-3.584)</td>
<td>-0.475 (-1.974)</td>
<td>-0.112 (-0.404)</td>
<td>-0.129 (0.807)</td>
<td></td>
</tr>
<tr>
<td>Eastern</td>
<td>Log income 0.326 (3.720)</td>
<td>0.306 (3.685)</td>
<td>0.300 (5.943)</td>
<td>0.234 (5.835)</td>
<td>1105</td>
</tr>
<tr>
<td></td>
<td>Unemployed -0.916 (-1.865)</td>
<td>-0.774 (-1.621)</td>
<td>-0.843 (-2.103)</td>
<td>-0.268 (0.848)</td>
<td></td>
</tr>
<tr>
<td>London</td>
<td>Log income 0.394 (8.280)</td>
<td>0.251 (5.452)</td>
<td>0.239 (4.057)</td>
<td>0.247 (5.643)</td>
<td>1403</td>
</tr>
<tr>
<td></td>
<td>Unemployed -0.551 (-2.911)</td>
<td>-0.672 (-2.658)</td>
<td>-0.530 (-2.610)</td>
<td>-0.368 (-1.769)</td>
<td></td>
</tr>
<tr>
<td>South East</td>
<td>Log income 0.363 (4.382)</td>
<td>0.361 (5.531)</td>
<td>0.300 (8.181)</td>
<td>0.227 (5.565)</td>
<td>1573</td>
</tr>
<tr>
<td></td>
<td>Unemployed -0.382 (-1.448)</td>
<td>0.052 (0.188)</td>
<td>0.022 (0.062)</td>
<td>0.003 (0.012)</td>
<td></td>
</tr>
<tr>
<td>South West</td>
<td>Log income 0.405 (7.183)</td>
<td>0.374 (8.439)</td>
<td>0.378 (7.243)</td>
<td>0.266 (5.441)</td>
<td>1080</td>
</tr>
<tr>
<td></td>
<td>Unemployed -0.393 (-2.142)</td>
<td>-0.171 (-0.819)</td>
<td>-0.074 (-0.245)</td>
<td>-0.327 (-1.279)</td>
<td></td>
</tr>
</tbody>
</table>

Marginal effects and t-statistics reported. Regressions include LEA dummies, school-type dummies, age-group dummies, rural-urban indicators, proportion in school with statements of special educational need, other special educational needs and non-white ethnic groups
Table 3-3 below shows the national figures for 1996-2001. Some of this growth is probably due to experience in meeting the demands of the tests, rather than real improvements in broader educational attainments.

Table 3-3: Key Stage 2 Performance in English Primary Schools, 1996-2001

<table>
<thead>
<tr>
<th>Year</th>
<th>Mean</th>
<th>s.d.</th>
<th>25th pctl</th>
<th>75th pctl</th>
<th>Schools</th>
</tr>
</thead>
<tbody>
<tr>
<td>1996</td>
<td>0.587</td>
<td>0.182</td>
<td>0.460</td>
<td>0.723</td>
<td>12534</td>
</tr>
<tr>
<td>1997</td>
<td>0.677</td>
<td>0.174</td>
<td>0.543</td>
<td>0.787</td>
<td>12752</td>
</tr>
<tr>
<td>1998</td>
<td>0.652</td>
<td>0.169</td>
<td>0.540</td>
<td>0.777</td>
<td>12984</td>
</tr>
<tr>
<td>1999</td>
<td>0.735</td>
<td>0.150</td>
<td>0.640</td>
<td>0.850</td>
<td>13200</td>
</tr>
<tr>
<td>2000</td>
<td>0.781</td>
<td>0.132</td>
<td>0.703</td>
<td>0.880</td>
<td>12793</td>
</tr>
<tr>
<td>2001</td>
<td>0.787</td>
<td>0.131</td>
<td>0.710</td>
<td>0.887</td>
<td>13507</td>
</tr>
</tbody>
</table>

This pattern of results shows considerable success, on average, in efforts to improve the reported performance of poor-performing schools in low-income neighbourhoods in the late 1990s. There is a lot more variation in the sensitivity to increases in catchment area unemployment rates, with coefficients of various signs, and many of them insignificant especially in more recent years. This instability indicates that the structural effect of unemployment is probably quite weak. What is clear is that the relationship observed in the lower panel of Figure 3-6 is largely attributable to household incomes, or income-related heterogeneity other than unemployment.

This all suggests that we can reasonably restrict our attention to a common effect of catchment area incomes on school performance across regions. The underlying distribution of incomes across the country influences the distribution of school performance, but the returns to income in terms of performance do not vary across geographical space. A stylised fact for the whole of England in 1999 is that pupils in neighbourhoods with a ten percent income advantage were 2.5 percentage points more likely to achieve target grades at Key Stage 2.
3.5.4 **Indices of latent neighbourhood status**

We have looked at the association between performance and incomes. We want to go further and incorporate other neighbourhood and intake characteristics. First, it is useful to restrict attention to indices of neighbourhood conditions, rather than a full vector of components. Characteristics of residents in school catchment areas tend to be highly mutually correlated, especially within local areas, since much of the variation is determined by sorting of agent types across space according to the amenities and residential characteristics if the neighbourhoods. For example, the correlation between property prices and incomes is 0.87, and between incomes and unemployment is 0.65. Exogenous variation – due, for example, to labour demand shocks across space – is not a major factor in generating variation within local authorities in neighbourhood unemployment rates or incomes. Income constraints in the demand for neighbourhood quality, plus the location of social housing are much stronger influences on the spatial distribution of characteristics.

So it makes sense to collapse these variables into a smaller number of indices of neighbourhood status by some dimensionality reduction technique. We do this using the principal factor method (see Bartholomew and Knott (1999)). In practice, one factor (or latent variable) dominates the covariance structure of the neighbourhood variables, and only this factor is retained. A theoretical justification of this procedure is that there exists an unobserved, scalar latent neighbourhood ‘quality’ variable that we can recover by factor analysis of the characteristics of residents. This ‘quality’ variable may simply reflect neighbourhood wealth, but since neighbourhood wealth will follow the distribution of neighbourhood quality (if residents are optimising and this neighbourhood quality is a normal good) we cannot identify its precise nature. In what follows, we shall refer simply to *neighbourhood status*. The advantage of this reduction in the data is that it offers us a simplified, broad-brush picture of how neighbourhood relates to pupil achievement.
Table 3-4 repeats the cross-regional analysis using this index of neighbourhood status derived by factor analysis of incomes, property prices, unemployment rates and the proportion at the school who are eligible for free school meals\textsuperscript{38}. The table reports the effect of a one-standard deviation increase in the index.

Table 3-4: School sensitivity to catchment area status index by region, 1996-1999

<table>
<thead>
<tr>
<th>Region</th>
<th>1996</th>
<th>1997</th>
<th>1998</th>
<th>1999</th>
<th>Sample</th>
</tr>
</thead>
<tbody>
<tr>
<td>North East</td>
<td>0.117</td>
<td>0.108</td>
<td>0.105</td>
<td>0.091</td>
<td>601</td>
</tr>
<tr>
<td></td>
<td>(18.182)</td>
<td>(31.555)</td>
<td>(12.516)</td>
<td>(15.318)</td>
<td></td>
</tr>
<tr>
<td>North West</td>
<td>0.111</td>
<td>0.098</td>
<td>0.094</td>
<td>0.078</td>
<td>2053</td>
</tr>
<tr>
<td></td>
<td>(17.103)</td>
<td>(26.483)</td>
<td>(17.624)</td>
<td>(14.223)</td>
<td></td>
</tr>
<tr>
<td>Yorks and Humberside</td>
<td>0.131</td>
<td>0.113</td>
<td>0.109</td>
<td>0.106</td>
<td>1187</td>
</tr>
<tr>
<td></td>
<td>(22.988)</td>
<td>(18.398)</td>
<td>(18.487)</td>
<td>(15.744)</td>
<td></td>
</tr>
<tr>
<td>East Midlands</td>
<td>0.138</td>
<td>0.125</td>
<td>0.137</td>
<td>0.114</td>
<td>932</td>
</tr>
<tr>
<td></td>
<td>(13.331)</td>
<td>(7.995)</td>
<td>(10.505)</td>
<td>(9.764)</td>
<td></td>
</tr>
<tr>
<td>West Midlands</td>
<td>0.133</td>
<td>0.117</td>
<td>0.106</td>
<td>0.096</td>
<td>1255</td>
</tr>
<tr>
<td></td>
<td>(11.359)</td>
<td>(8.867)</td>
<td>(8.555)</td>
<td>(10.695)</td>
<td></td>
</tr>
<tr>
<td>Eastern</td>
<td>0.118</td>
<td>0.124</td>
<td>0.128</td>
<td>0.102</td>
<td>1105</td>
</tr>
<tr>
<td>London</td>
<td>0.107</td>
<td>0.100</td>
<td>0.100</td>
<td>0.104</td>
<td>1403</td>
</tr>
<tr>
<td>South East</td>
<td>0.129</td>
<td>0.128</td>
<td>0.114</td>
<td>0.096</td>
<td>1573</td>
</tr>
<tr>
<td></td>
<td>(14.619)</td>
<td>(7.985)</td>
<td>(9.641)</td>
<td>(11.663)</td>
<td></td>
</tr>
<tr>
<td>South West</td>
<td>0.135</td>
<td>0.104</td>
<td>0.117</td>
<td>0.094</td>
<td>1080</td>
</tr>
<tr>
<td></td>
<td>(23.447)</td>
<td>(10.906)</td>
<td>(10.237)</td>
<td>(11.914)</td>
<td></td>
</tr>
</tbody>
</table>

Reported parameter is response to unit change (one standard deviation) in standardised neighbourhood index.

Index is first factor in factor analysis of log mean incomes, unemployment rate, log mean property prices, quadratic in proportion eligible for free school meals. Proportion of variance explained by this factor is 85-90%.

Regressions include LEA dummies, school-type dummies, age-group dummies, rural-urban indicators, proportion in school with statements of special educational need, other special educational needs and non-white ethnic groups.

The index uses the only time varying measures of intake and catchment area conditions we have available, except for the special educational needs and ethnic

\textsuperscript{38} The factor loadings and Eigenvalues are given in Appendix C. In practice, there is no ambiguity in the choice of factors for the index. Only the first factor in this analysis has factor loadings that correspond to the intuition that incomes and property prices increase economic status, whilst unemployment rates and free-school meal eligibility decrease it. Between 85-90 percent of the information contained in incomes, unemployment and property prices at the neighbourhood level, plus free school meal eligibility at school level, can be captured in this single linear combination.
indicators which are retained as additional regressors. This adjusts for linguistic and learning disadvantages of school pupils, which are not necessarily due to economic disadvantages. The pattern is similar to that in Table 3-2, with declining sensitivity to catchment area status in most regions from 1996 to 1999. Only in London is there no obvious change over time (the marginal effects are equal across periods, \( p \)-value = 0.987), though the coefficients are not significantly different in the statistical sense for the East or West Midlands (\( p \)-values 0.154 and 0.583 respectively). Any differences between regions in the relationship between intake disadvantage and performance are relatively small, although they are significant in 1997 and 1999. Schools in neighbourhoods that were one-standard deviation above the average neighbourhood could expect a 12.5 percentage point advantage in 1996, falling to around 9.5 percentage points in 1999. This average relationship is pretty much in line with what we see by visual inspection of Figure 3-6, with a ten percentage point change for a shift from mean incomes (£19500 in 1996) to one standard deviation above (£23900).

3.5.5 Disentangling the contributions of local characteristics

So far we have looked at the impact of local incomes and unemployment on school performance, and the relationship to a time-varying index of intake and local conditions across regions. We now dispense with concerns about regional differences, test for the importance of selection effects on the observed catchment area-performance relationship, then look at the idiosyncratic contributions of neighbourhood characteristics. Table 3-5 shows estimates of the influence of neighbourhood characteristics using pooled data for 1998 and 1999 for all regions. Again, we will start with a characterisation of the neighbourhood-performance relationship using a quadratic in a single factor derived by factor analysis of a number of catchment area characteristics and free-school meal entitlement. Special educational needs and ethnicity are retained as controls, along with school type, age range, rural-urban, year and LEA dummies.
Table 3-5: Catchment area status and primary school performance, 1998 & 1999

<table>
<thead>
<tr>
<th></th>
<th>OLS (1)</th>
<th>OLS (2)</th>
<th>IV (3)</th>
<th>IV (4)</th>
<th>Means</th>
</tr>
</thead>
<tbody>
<tr>
<td>Auto correlation ρ</td>
<td>-</td>
<td>-</td>
<td>0.811</td>
<td>0.830</td>
<td>(52.281)</td>
</tr>
<tr>
<td>Index of neighbourhood status</td>
<td>0.037</td>
<td>-</td>
<td>0.037</td>
<td>-</td>
<td>0.000</td>
</tr>
<tr>
<td></td>
<td>(13.004)</td>
<td>(4.136)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Index squared</td>
<td>0.012</td>
<td>-</td>
<td>0.018</td>
<td>-</td>
<td>1.000</td>
</tr>
<tr>
<td></td>
<td>(8.090)</td>
<td>(4.186)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Index of neighbourhood status with free-school meals</td>
<td>-</td>
<td>0.109</td>
<td>-</td>
<td>0.102</td>
<td>(41.259)</td>
</tr>
<tr>
<td>Index squared</td>
<td>-</td>
<td>0.021</td>
<td>-</td>
<td>0.024</td>
<td>(14.981)</td>
</tr>
<tr>
<td>Pupils eligible for free school meals</td>
<td>-0.870</td>
<td>-</td>
<td>-0.744</td>
<td>-</td>
<td>1.96</td>
</tr>
<tr>
<td></td>
<td>(-27.547)</td>
<td>(-7.660)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pupils eligible for free-school meals squared</td>
<td>0.623</td>
<td>-</td>
<td>0.431</td>
<td>-</td>
<td>0.069</td>
</tr>
<tr>
<td></td>
<td>(13.343)</td>
<td>(3.361)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pupils with statements</td>
<td>-0.752</td>
<td>-0.956</td>
<td>-1.87</td>
<td>-1.669</td>
<td>0.017</td>
</tr>
<tr>
<td></td>
<td>(-11.042)</td>
<td>(-13.371)</td>
<td>(-5.236)</td>
<td>(-5.250)</td>
<td></td>
</tr>
<tr>
<td>Others with special needs</td>
<td>-0.236</td>
<td>-0.402</td>
<td>-0.228</td>
<td>-0.355</td>
<td>0.194</td>
</tr>
<tr>
<td></td>
<td>(-15.084)</td>
<td>(-24.274)</td>
<td>(-4.046)</td>
<td>(-5.348)</td>
<td></td>
</tr>
<tr>
<td>Non-white ethnic group pupils</td>
<td>-0.037</td>
<td>-0.059</td>
<td>-0.057</td>
<td>-0.078</td>
<td>0.087</td>
</tr>
<tr>
<td></td>
<td>(-4.414)</td>
<td>(-5.247)</td>
<td>(-0.125)</td>
<td>(-2.823)</td>
<td></td>
</tr>
<tr>
<td>R2</td>
<td>0.592</td>
<td>0.567</td>
<td>0.600</td>
<td>0.593</td>
<td></td>
</tr>
<tr>
<td>Hausman test p-value</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.170</td>
</tr>
</tbody>
</table>

Sample size: N=11189 × T=2

All models include LEA (145), rural-urban (5), school type(3), age-range (4) dummies and full-time equivalent pupils numbers, 1999 dummy

Instrument in Column (4) is second time lag of dependent variable. All parameters (except ρ) are marginal effects at the mean. Lagged dependent variable parameters are multiplied by (1-ρ)^1

Estimates of catchment area/free school meals index parameters without controls for special needs and ethnicity are Index: 0.134 (8.826) Index^2: 0.028 (4.991) in lagged dependent variable model

For definition of catchment area indices see the Appendices

Column (1) of Table 3-5 excludes free-school meal entitlement from the catchment area index, whilst Column (2) includes it^39. All neighbourhood characteristics are potentially endogenous in a school production function. This issue is highlighted by the results in Gibbons and Machin (2001) that show substantial property price premia close to schools with exogenously better performance. Columns (3) and (4) repeat the analysis, but with instrumented lagged school performance as a control for persistent components

^39 Details of the factor analysis are in Appendix C
of school performance, which might induce residential selection and inconsistent estimates.

The impression from this Table is, perhaps unsurprisingly, of school performance increasing with catchment area status and decreasing with school intake income disadvantage. The effects of catchment area conditional on school intake are quite small, though increasing with neighbourhood status. A one-standard deviation increase from the mean increases performance by nearly four percentage points; another one-standard deviation increases it by a further five percentage points. Obviously if we were to remove the intake characteristics from the regression, we would get bigger effects. Instead, what we do in Column (2) is include free-school meal entitlement in the neighbourhood status index, and retain special needs and ethnic group as controls. Here we find that a one standard deviation neighbourhood improvement from the mean increases performance by just over ten percentage points, rising to thirteen percentage points for a further standard deviation shift. This is much like the result we saw in the cross-regional analysis in Table 3-4, suggesting that all the additional neighbourhood variables add very little information in terms catchment area effects on performance. The models in Columns 3 and 4 control for persistent unobserved school components, but we find little change in the estimated parameters. The Hausman tests of parameter equivalence do not reject the exogeneity of neighbourhood composition, as measured by this composite index. So far we have looked at the relationship between school performance and a neighbourhood index. Table 3-6 separates out the impact of the components of neighbourhood status.

40 Without the special needs controls, the effect of a one standard deviation increase in catchment area status rises to just under fourteen percent.

41 Tests on various components of the index entering individually in the regression confirm that the time-varying catchment area components – incomes, property prices and unemployment – are potentially endogenous, whereas the Census derived variables are not.
Table 3-6: Catchment area attributes and primary school performance, 1998 & 1999

<table>
<thead>
<tr>
<th></th>
<th>OLS</th>
<th>OLS</th>
<th>IV</th>
<th>IV</th>
<th>Means</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(1)</td>
<td>(2)</td>
<td>(3)</td>
<td>(4)</td>
<td></td>
</tr>
<tr>
<td>Autocorrelation $\hat{\rho}$</td>
<td></td>
<td></td>
<td>0.876**</td>
<td>0.798**</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(71.534)</td>
<td>(48.385)</td>
<td></td>
</tr>
<tr>
<td>Catchment area</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Log mean local household incomes</td>
<td>0.092**</td>
<td>0.045**</td>
<td>-5.0e-03</td>
<td>-7.8e-03</td>
<td>9.924</td>
</tr>
<tr>
<td>Log mean local property prices in previous year</td>
<td>(4.802)</td>
<td>(3.111)</td>
<td>(-0.064)</td>
<td>(-0.169)</td>
<td></td>
</tr>
<tr>
<td>Local unemployment claimants per household</td>
<td>0.016</td>
<td>0.013</td>
<td>3.1e-03</td>
<td>7.0e-03</td>
<td>11.159</td>
</tr>
<tr>
<td></td>
<td>(1.884)</td>
<td>(1.882)</td>
<td>(0.075)</td>
<td>(0.290)</td>
<td></td>
</tr>
<tr>
<td>Aged 8-12 not in social housing in 1991</td>
<td>-0.555**</td>
<td>0.130</td>
<td>-0.488</td>
<td>0.224</td>
<td>0.055</td>
</tr>
<tr>
<td></td>
<td>(-5.929)</td>
<td>(1.885)</td>
<td>(-1.316)</td>
<td>(1.012)</td>
<td></td>
</tr>
<tr>
<td>Higher educated over-18s in 1991</td>
<td>0.106**</td>
<td>0.023*</td>
<td>0.112*</td>
<td>0.027</td>
<td>0.720</td>
</tr>
<tr>
<td></td>
<td>(7.945)</td>
<td>(2.431)</td>
<td>(2.466)</td>
<td>(0.979)</td>
<td></td>
</tr>
<tr>
<td>Average age of population in 1991</td>
<td>0.323**</td>
<td>0.277**</td>
<td>0.275*</td>
<td>0.244**</td>
<td>0.119</td>
</tr>
<tr>
<td></td>
<td>(10.596)</td>
<td>(11.539)</td>
<td>(2.343)</td>
<td>(3.446)</td>
<td></td>
</tr>
<tr>
<td>Economically active men 25-34 in 1991</td>
<td>0.126*</td>
<td>8.1e-03</td>
<td>-0.043</td>
<td>-0.098</td>
<td>0.958</td>
</tr>
<tr>
<td></td>
<td>(2.185)</td>
<td>(0.200)</td>
<td>(-0.151)</td>
<td>(-0.562)</td>
<td></td>
</tr>
<tr>
<td>Economically active women 25-34 in 1991</td>
<td>0.019</td>
<td>-0.029</td>
<td>-0.057</td>
<td>-0.069</td>
<td>0.650</td>
</tr>
<tr>
<td></td>
<td>(0.725)</td>
<td>(-1.404)</td>
<td>(-0.659)</td>
<td>(-1.299)</td>
<td></td>
</tr>
<tr>
<td>Lone parents households in 1991</td>
<td>0.075</td>
<td>0.197**</td>
<td>0.214</td>
<td>0.234</td>
<td>0.045</td>
</tr>
<tr>
<td></td>
<td>(0.784)</td>
<td>(2.890)</td>
<td>(0.635)</td>
<td>(1.186)</td>
<td></td>
</tr>
<tr>
<td>Proportion of households with dependent children</td>
<td>0.252**</td>
<td>0.106**</td>
<td>0.197</td>
<td>0.077</td>
<td>0.323</td>
</tr>
<tr>
<td></td>
<td>(6.089)</td>
<td>(3.536)</td>
<td>(1.243)</td>
<td>(0.782)</td>
<td></td>
</tr>
<tr>
<td>Long term sick rate in 1991</td>
<td>-0.704**</td>
<td>0.061</td>
<td>-1.220**</td>
<td>-0.191</td>
<td>0.126</td>
</tr>
<tr>
<td></td>
<td>(-7.474)</td>
<td>(0.910)</td>
<td>(-3.508)</td>
<td>(-0.917)</td>
<td></td>
</tr>
<tr>
<td>One-year migrants in 1991</td>
<td>0.150**</td>
<td>0.068</td>
<td>-0.029</td>
<td>-0.040</td>
<td>0.065</td>
</tr>
<tr>
<td></td>
<td>(2.696)</td>
<td>(1.526)</td>
<td>(-0.125)</td>
<td>(-0.282)</td>
<td></td>
</tr>
<tr>
<td>Average rooms in owner-occupier hhs in 1991</td>
<td>0.015**</td>
<td>9.4e-03**</td>
<td>0.037*</td>
<td>0.021*</td>
<td>5.392</td>
</tr>
<tr>
<td></td>
<td>(3.230)</td>
<td>(2.720)</td>
<td>(2.224)</td>
<td>(2.126)</td>
<td></td>
</tr>
<tr>
<td>Households (1000s) per km² in 1991</td>
<td>9.1e-04</td>
<td>2.3e-03*</td>
<td>1.7e-03</td>
<td>2.8e-03</td>
<td>2.337</td>
</tr>
<tr>
<td></td>
<td>(0.706)</td>
<td>(2.332)</td>
<td>(0.360)</td>
<td>(1.052)</td>
<td></td>
</tr>
<tr>
<td>Agricultural employment in 1991</td>
<td>-0.193**</td>
<td>-0.075</td>
<td>-0.211</td>
<td>-0.086</td>
<td>0.015</td>
</tr>
<tr>
<td></td>
<td>(-3.035)</td>
<td>(-1.425)</td>
<td>(-1.089)</td>
<td>(-0.710)</td>
<td></td>
</tr>
<tr>
<td>School level:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pupils eligible for free school meals</td>
<td>-</td>
<td>-0.919**</td>
<td>-</td>
<td>-0.956**</td>
<td>0.196</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(31.790)</td>
<td></td>
<td>(-8.334)</td>
<td></td>
</tr>
<tr>
<td>Pupils eligible for free-school meals squared</td>
<td>0.694**</td>
<td></td>
<td>0.772**</td>
<td></td>
<td>0.069</td>
</tr>
<tr>
<td></td>
<td>(16.111)</td>
<td></td>
<td>(5.286)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pupils with statements</td>
<td>-</td>
<td>-0.776**</td>
<td>-</td>
<td>-1.154**</td>
<td>0.017</td>
</tr>
<tr>
<td></td>
<td>(-11.811)</td>
<td></td>
<td>(-5.542)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Others with special needs</td>
<td>-</td>
<td>-0.228**</td>
<td>-</td>
<td>-0.206**</td>
<td>0.194</td>
</tr>
<tr>
<td></td>
<td>(-16.194)</td>
<td></td>
<td>(-3.905)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Non-white ethnic group pupils</td>
<td>-</td>
<td>-0.055**</td>
<td>-</td>
<td>-0.075**</td>
<td>0.087</td>
</tr>
<tr>
<td></td>
<td>(-5.726)</td>
<td></td>
<td>(-2.824)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(R^2)</td>
<td>0.498</td>
<td>0.602</td>
<td>0.575</td>
<td>0.605</td>
<td></td>
</tr>
</tbody>
</table>

Sample size: N=11189 x T=2; Coefficients * significant at the 5% level, ** significant at 1% level
All models include LEA, rural-urban, school type, age-range dummies and full-time equivalent pupils numbers; Instrument in Column (4) is second time lag of dependant variable. All parameters (except $\rho$) are marginal effects at the mean.
All these covariates are moderately to highly correlated and are measured with an unknown degree of error, so we should take care before placing too strong a causal interpretation on their separate contributions. Column (1) is straightforward grouped probit on the battery of catchment area characteristics. Column (2) conditions on the intake characteristics available at school level. Columns 4-6 repeat this sequence but are estimates conditional on prior performance.

Most of the neighbourhood attributes have the signs we would expect if we interpret them as measures of household economic resources. With the Census-derived neighbourhood attributes in the regression, the impact of local incomes is reduced to around one-third of what we estimated in Table 3-2. The regressions show no statistically significant effects from local incomes, property values or unemployment rates in the transformed model with lagged performance. The only highly significant (at least at the 5% level) catchment area characteristics in Column (3) are the proportion highly-qualified, average age, owner-occupier property size, the proportion of non-socially housed children and long-term sick rates. The impact of the last two characteristics is absorbed by the proportion eligible for free school meals or with special educational needs in Column (4). The most significant neighbourhood factor in all specifications is the education of local residents. We cannot identify the precise causal mechanisms, but it looks like wealth-related factors are the major neighbourhood determinants of pupils' success in reaching target attainments in school – population age, education, and property size (and entitlement to free meals at school-intake level).

The statistically insignificant or positive relationship between lone-parents and performance is interesting. Gordon (1996) finds strong negative associations between non-employed lone-parents and GCSE secondary performance, and interprets these in terms of family structure effects. Here, looking at primary schools, the raw correlation between lone parents and school performance is certainly strongly negative. But, once we include other catchment area and school intake controls we find schools in areas with
more lone parent households doing better than other areas. In fact, we need only control for the school proportion on free-school meals and the local proportion highly qualified to get this result. On its own, the lone-parent proportion is a powerful proxy for low income as measured by free-school meal entitlement, since many lone parents are benefit claimants. This suggests that being a lone parent may not, by itself, damage a child's prospects in primary school - though having poorly educated, benefit-claiming parents does. Truancy, which mediates the effects of non-employed lone parents in Gordon's study, is probably less of an issue at primary school. We find no significant effect from lone parents in the IV, lagged performance models, but the coefficient remains positive.

3.5.6 Local interactions and neighbourhood externalities

One reading of the evidence in Table 3-5 and Table 3-6 - that catchment area characteristics matter over and above school intake characteristics - is that there are spillover effects from the local community or from peers from other schools. This idea is reinforced if we restrict attention to the predicted catchment area proportion eligible for free school meals: this has an impact over and above free-school meal eligibility in the school, with a ten percent reduction in the proportion of residents in poverty increasing school success rates by 1.7 percent. Of course, an equally plausible interpretation is that catchment area characteristics simply capture performance-related attributes of the school intake that are uncorrelated with free-school meal entitlement, educational needs and ethnic group. A more convincing argument is that if geographical interactions matter in the performance of primary schools, we should find that schools that perform well are clustered together and schools that perform badly are clustered together, even taking into account the type of neighbourhood. We found this to be the case using descriptive

\footnote{If we restrict attention to variables at neighbourhood level, then age, qualifications, long-term sick rates and social housing drive the lone-parent coefficient to near-zero insignificance.}
statistics in Section 3.5.1. Estimation of the model in (3-1) with a spatially lagged dependent variable allows us to test more rigorously for these spatial interaction effects.\(^{43}\)

Table 3-7 shows the results of the spatial auto-regressions, for various spatial lags.

<table>
<thead>
<tr>
<th></th>
<th>J=1</th>
<th>J=3</th>
<th>J=6</th>
<th>J=9</th>
<th>J=18</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Spatial lags not instrumented</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean performance in nearest J schools primary schools</td>
<td>0.016</td>
<td>0.086</td>
<td>0.084</td>
<td>0.099</td>
<td>0.059</td>
</tr>
<tr>
<td></td>
<td>(2.234)</td>
<td>(4.480)</td>
<td>(5.747)</td>
<td>(6.212)</td>
<td>(5.068)</td>
</tr>
<tr>
<td>Mean distance to nearest J schools in kilometres</td>
<td>-2.3e-03</td>
<td>-2.3e-03</td>
<td>-2.6e-03</td>
<td>-2.3e-03</td>
<td>-1.3e-03</td>
</tr>
<tr>
<td></td>
<td>(-2.334)</td>
<td>(-2.732)</td>
<td>(-3.540)</td>
<td>(-3.535)</td>
<td>(-2.504)</td>
</tr>
<tr>
<td>Households per km(^2) (1000s)</td>
<td>1.8e-03</td>
<td>1.9e-03</td>
<td>1.7e-03</td>
<td>1.7e-03</td>
<td>1.8e-03</td>
</tr>
<tr>
<td></td>
<td>(1.987)</td>
<td>(1.988)</td>
<td>(1.771)</td>
<td>(1.840)</td>
<td>(1.900)</td>
</tr>
<tr>
<td>Rural-urban location, p-value</td>
<td>0.279</td>
<td>0.263</td>
<td>0.181</td>
<td>0.214</td>
<td>0.315</td>
</tr>
<tr>
<td></td>
<td>0.590</td>
<td>0.591</td>
<td>0.591</td>
<td>0.592</td>
<td>0.591</td>
</tr>
<tr>
<td><strong>Spatial lags instrumented</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean performance in nearest J schools primary schools</td>
<td>5.5e-03</td>
<td>0.076</td>
<td>0.073</td>
<td>0.086</td>
<td>0.057</td>
</tr>
<tr>
<td></td>
<td>(0.430)</td>
<td>(3.091)</td>
<td>(3.491)</td>
<td>(3.816)</td>
<td>(3.000)</td>
</tr>
<tr>
<td>Mean distance to nearest J schools in kilometres</td>
<td>-2.3e-03</td>
<td>-2.3e-03</td>
<td>-2.5e-03</td>
<td>-2.2e-03</td>
<td>-1.2e-03</td>
</tr>
<tr>
<td></td>
<td>(-2.365)</td>
<td>(-2.699)</td>
<td>(-3.427)</td>
<td>(-3.352)</td>
<td>(-2.439)</td>
</tr>
<tr>
<td>Households per km(^2) (1000s)</td>
<td>1.8e-03</td>
<td>1.9e-03</td>
<td>1.7e-03</td>
<td>1.7e-03</td>
<td>1.8e-03</td>
</tr>
<tr>
<td></td>
<td>(1.904)</td>
<td>(1.972)</td>
<td>(1.768)</td>
<td>(1.835)</td>
<td>1.902</td>
</tr>
<tr>
<td>Rural-urban location, p-value</td>
<td>0.287</td>
<td>0.267</td>
<td>0.194</td>
<td>0.231</td>
<td>0.323</td>
</tr>
<tr>
<td></td>
<td>0.590</td>
<td>0.591</td>
<td>0.591</td>
<td>0.591</td>
<td>0.591</td>
</tr>
</tbody>
</table>

Sample size: \(N=11189 \times T=2\)

All models include LEA dummies, proportions: eligible for free school meals, for free school meals squared, with statements, with other special educational needs, non-white ethnic groups, non-socially housed, the number of FTE pupils, neighbourhood wealth index and index squared, school type, age-range, Household density and school distance are always jointly significant at the 1% level at least. Instruments are spatial lags (mean in nearest \(J\) schools) of school type, age-range, non-socially housed pupils and neighbourhood index (means of 98-99 if time varying)

\(^{43}\) The potential correlation between school characteristics and past performance is ignored and we use only Census and school intake characteristics as regressors in the vector \(X_{s,t}\), which in any case tested as exogenous.
In the top panel, Column (1), we find a very low, though significant correlation of school performance with that of the nearest school. Performance also decreases with the distance to the nearest school. There is no reason to believe that all interaction effects will relate to the nearest school, which may well be noisy measure of the relevant performance cluster. In Columns 2-5 the number of schools in the comparison group is increased from three to eighteen. The estimated correlation parameter rises to a maximum of nearly 0.1 when we consider the nearest nine schools, then falls away as the size of the group is increased. The effect of school dispersion is quite stable up to the nearest nine schools: an extra kilometre on average to the nearest schools reduces performance by 0.23 to 0.26 percentage points. This is not a very big effect, but is significant. The school distance effect, and the positive relationship with household density, is consistent with the maps we drew in Section 3.5.1 – schools in cities with high-density housing and with a high density of schools perform better than others. This is not attributable to the general rural/urban distinctions captured by a set of five dummy variables. The coefficients on these are not jointly or individually significant.

These will be inconsistent estimates of the interdependence of local school performance if there are unobserved area effects and because of feedback from own performance to performance schools in the comparison group. In the lower panel of Table 3-7 we instrument the spatially lagged school performance with the average of the school-type, age-range, and neighbourhood characteristics relating to the nearest schools. There is very little difference between these IV results and the OLS results above. The small, but significant correlation between the performance of local schools suggests a role for

44 Only one in three schools has this many neighbouring schools within a 2km radius, but three out of four schools have this many within a 4km radius.

45 These are jointly significant, although household density is only significant at around the 10% level.
geographically localised spillover effects between schools, either through interactions between neighbouring pupils, or through inter-school technological spillovers. Changing school inputs to generate a ten percentage point improvement in the probability of target attainment by pupils generates an additional one percentage point improvement through feedback effects.

As noted, performance falls as mean distance to the nearest schools increases. This reinforces the case for spatial interaction effects, arising either through mutual imitation of teaching technologies, or through neighbourhood based peer-group effects. Higher school concentration can mean greater pooling of ideas, knowledge, and expectations amongst both pupils and teachers. Increased pooling in an environment of technological and human capital spillovers can only exert a positive influence on achievements, since it increases the range of options available and options with an expected negative impact need never be exercised. School density has been interpreted in other work as an indicator of the intensity of local competition between schools Bradley, et al. (1999). The indication here is that this localisation improves pupil achievement, though this could equally be through shared-technology and neighbourhood interactions as through ‘competitive’ effects.

3.5.7 Property effects on non-property owners

In Section 3.3.1, we considered how interaction between pupils and their neighbours, through role model effects, or expectations, might be detected through property-price effects on social tenants. If the residential location decisions of social tenants are unrelated to income-based residential sorting processes in the private housing sector, then social tenant family resources will be uncorrelated with the family resources of local owner-occupiers. In this case, there is no relationship between property prices and school performance that is not mediated through the proportion of owner-occupier children in the school catchment area. We first test this assumption using data on
location, family incomes and tenancy group from the Survey of English Housing, 1996-1999 and confirm that incomes of social tenants living close to richer owner-occupiers are not significantly higher than the incomes of social tenants living in poor residential areas. Details and results are in Appendix D.

Given this, we should not find a property price-school performance relationship that is independent of the proportion of owner-occupiers in the area. Nevertheless, Figure 3-7, shows that property prices are related to school performance throughout the distribution of school catchment area tenancy group composition.

**Figure 3-7: Relationship between primary school performance, local property prices and social-tenant children, 1996-1999**

Figure shows kernel regression surface of mean key stage 2 level 4 achievement, 1996-1999, on log-mean local property prices and local proportion of non-socially housed 8-12 year olds. Controls are Local Authority dummies, school type, age-range and rural/urban category dummies. Kernel: Gaussian, bandwidth matrix = 0.25*variance-covariance matrix of regressors. Cholesky decomposition of bandwidth matrix = \{ 0.098, 0.017 \ 0, 0.057 \}

N = 11571, Overall R² = 0.535
The figure shows the kernel regression surface of mean Key Stage 2 performance (from 1996-1999) on log-mean property prices (in 1995) and the proportion of primary school age children in the school catchment area. The edge of the surface running along the nearest house-price axis corresponds to the hypothetical case of primary schools where 100% of pupils are social tenants. In high house price areas, Key Stage 2 pass rates of social tenants are 25 percentage points higher than in similar schools in low-price areas. And this relationship cannot be generated by reverse causality – the influence of school performance on property prices – since home-owners and renters are will have no incentive to bid up property prices in areas where nearly all the children are in social housing. Even at the top of the distribution of property prices, schools with low proportions of non-social tenant children struggle to reach the mean in terms of school performance.

Note, however, that the relationship between performance and local property wealth is, as we would expect, much stronger when the proportion of social tenants is low: moving from the bottom to the top of the property price distribution leads to a 35 percentage point improvement in performance, from around 50% success rates to around 85% success rates at key stage 2.

Estimation of the parametric version of this surface in equation (3-6), yields an estimates of $\lambda_1 = 0.072$ ($t=4.758$), $\lambda_2 = 0.140$ ($t=8.093$)\(^{46}\): a ten percent relative increase in local property prices gives a baseline improvement of around 0.75 percentage points on the probability of success at Key Stage 2 – for all pupils, whether or not they are from home-owning households. The additional impact on non-socially housed tenants, those

\[^{46}\text{These are the marginal effects, not the coefficients.}\]

- 135 -
for whom property wealth has a direct impact through parental resources, is around 1.4 percentage points.\textsuperscript{47}

3.5.8 Do teaching and resource inputs matter?

What about the key economic resources – pupil teacher ratios and expenditures? Do these influence pupil success in primary schools? The key problem in inferring causal effects is the geographical distribution of base-line pupil advantages, the allocation of resources and parental selection based on unobserved, idiosyncratic school effectiveness. Our detailed mapping of area characteristics to school locations and our panel structure present an ideal opportunity to address these issues. The estimates in Table 3-8 do just that. The first Column provides weighted least squares estimates of the association of performance with teaching inputs measured at the school level, and Local Authority expenditures per pupil. The specification includes LEA and time dummies, so the expenditure effect is identified off non-general changes within LEAs over time. With controls for school type, intake and neighbourhood status, we find positive effects from total LEA expenditure per pupil, but generally negative associations with the ratio of teachers to pupils. The only exception is for teachers on the ‘other’ category, which is made up of small numbers of student, ‘licensed’, unqualified teachers and language

\textsuperscript{47} In principle, we could estimate this model using Instrumental Variables for property prices and the price-composition interaction. Property characteristics and interactions provide potential instruments. In practice, the results from this exercise implied that all property wealth effects on school performance were general effects across tenancy groups, with no home-owner-specific components. Whilst the idea that property wealth effects are purely neighbourhood spillover effects that benefit all members of the community is interesting, it seems unlikely, and the results suggest some misspecification of the IV model. Nevertheless, tests of the exogeneity of property size from comparison of the (adjusted) coefficients in the school production model with and without lagged performance confirm exogeneity. The adjusted coefficients are equal in both models (\textit{p-value}=0.883).
assistants\(^48\). School size has a separate negative impact, but this is very small in magnitude – a reduction of 0.5 percentage points for an extra one hundred pupils at the school.

Table 3-8: Resources, school structure and primary school performance, 1998 & 1999

<table>
<thead>
<tr>
<th></th>
<th>OLS (1)</th>
<th>Within (2)</th>
<th>OLS (3)</th>
<th>IV (4)</th>
<th>Means</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>(\hat{\rho})</strong></td>
<td>-</td>
<td>-</td>
<td>0.436</td>
<td>0.828</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>(63.176)</td>
<td>(58.302)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Qualified teachers/100 pupils</td>
<td>-0.094</td>
<td>8.3e-03</td>
<td>-3.6e-03</td>
<td>0.026</td>
<td>4.219</td>
</tr>
<tr>
<td>(fte)</td>
<td>(-3.066)</td>
<td>(2.827)</td>
<td>(-0.952)</td>
<td>(2.356)</td>
<td></td>
</tr>
<tr>
<td>Other teachers/100 pupils</td>
<td>0.021</td>
<td>3.0e-03</td>
<td>0.036</td>
<td>0.115</td>
<td>0.036</td>
</tr>
<tr>
<td>(fte)</td>
<td>(2.154)</td>
<td>(0.478)</td>
<td>(2.986)</td>
<td>(3.452)</td>
<td></td>
</tr>
<tr>
<td>Support teachers/100 pupils</td>
<td>-6.2e-03</td>
<td>-4.9e-04</td>
<td>-5.8e-03</td>
<td>-4.1e-03</td>
<td>0.780</td>
</tr>
<tr>
<td>(fte)</td>
<td>(-2.776)</td>
<td>(-0.224)</td>
<td>(-2.106)</td>
<td>(-0.499)</td>
<td></td>
</tr>
<tr>
<td>Administration staff/100</td>
<td>-0.015</td>
<td>1.3e-03</td>
<td>-0.013</td>
<td>-4.7e-03</td>
<td>0.423</td>
</tr>
<tr>
<td>pupils (fte)</td>
<td>(-1.893)</td>
<td>(0.172)</td>
<td>(-1.381)</td>
<td>(-0.174)</td>
<td></td>
</tr>
<tr>
<td>Log total LEA expenditure per</td>
<td>0.036</td>
<td>0.041</td>
<td>0.039</td>
<td>0.042</td>
<td>7.473</td>
</tr>
<tr>
<td>primary school pupil</td>
<td>(2.773)</td>
<td>(3.212)</td>
<td>(2.293)</td>
<td>(1.980)</td>
<td></td>
</tr>
<tr>
<td>Full time equivalent pupils</td>
<td>-5.3e-03</td>
<td>2.0e-4</td>
<td>-5.3e-03</td>
<td>5.4e-03</td>
<td>2.783</td>
</tr>
<tr>
<td>(100s)</td>
<td>(-4.844)</td>
<td>(-4.392)</td>
<td>(-3.850)</td>
<td>(-1.368)</td>
<td></td>
</tr>
<tr>
<td>Foundation</td>
<td>0.036</td>
<td>-</td>
<td>0.038</td>
<td>0.049</td>
<td>0.023</td>
</tr>
<tr>
<td></td>
<td>(6.404)</td>
<td></td>
<td>(5.160)</td>
<td>(2.210)</td>
<td></td>
</tr>
<tr>
<td>Voluntary Aided</td>
<td>0.047</td>
<td>-</td>
<td>0.048</td>
<td>0.052</td>
<td>0.233</td>
</tr>
<tr>
<td></td>
<td>(16.845)</td>
<td></td>
<td>(13.534)</td>
<td>(4.899)</td>
<td></td>
</tr>
<tr>
<td>Voluntary Controlled</td>
<td>0.015</td>
<td>-</td>
<td>0.012</td>
<td>-0.003</td>
<td>0.112</td>
</tr>
<tr>
<td></td>
<td>(4.663)</td>
<td></td>
<td>(3.048)</td>
<td>(-0.222)</td>
<td></td>
</tr>
<tr>
<td>Age range</td>
<td>0.0000</td>
<td>-</td>
<td>0.0000</td>
<td>0.1894</td>
<td>-</td>
</tr>
<tr>
<td>Year = 1999</td>
<td>0.087</td>
<td>0.101</td>
<td>0.091</td>
<td>0.093</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>(70.705)</td>
<td>(72.591)</td>
<td>(57.947)</td>
<td>(46.983)</td>
<td></td>
</tr>
<tr>
<td>(R^2)</td>
<td>0.577</td>
<td>0.665</td>
<td>0.594</td>
<td>-</td>
<td></td>
</tr>
</tbody>
</table>

Sample size N=1189 x T=2

All models include proportions: eligible for free school meals, with special educational needs, with statements of SEN, non-white, plus an index of neighbourhood status. All except Column(2) include LEA dummies, rural-urban dummies,

Instrument in Column (4) is second time lag of dependant variable. All parameters (except \(p\)) are marginal effects at the mean.

Expenditure elasticity alone is: Column (1): 0.037 (t=2.740), Column (2): 0.042 (t=3.320), column3 : 0.042 (t=2.068), Column (4): 0.043 (t= 1.878)

---

\(^48\) Further investigations suggested that is the student and licensed teacher types that drive the positive coefficient on this variable.
The decrease in performance as we increase the number of teachers relative to pupils exemplifies the endogeneity issue— we cannot identify the causal effect of teachers on performance in the cross-section without knowledge of resource allocation policy and the parental selection process. Even controlling for observable school intake differences is insufficient to ensure exogeneity of the teaching inputs. However, taking deviations from the school averages over time— see Column (2)— we find that increases in qualified teacher-pupil ratios within schools over time are linked to increases in pupil attainments, but the coefficients on other teacher and staff types are not significant. The weakness of teacher effects is unsurprising, since it is well known that the within-group transformation reduces the signal to noise ratio and increases attenuation due to measurement error. We also lose the estimates on non-varying school attributes. A one percent increase in LEA expenditure increases the probability of Key Stage 2 target attainment by just over four percentage points.

Columns (3) and (4) implement the model of equation (3-3). By way of demonstration, in Column (3), lagged pupil achievement is un-instrumented. We should expect the estimated serial correlation parameter to be downward biased by the presence of unobserved components of lagged performance in the error term. The marginal effects of other variables are broadly similar to those in Column (1), though the coefficient on qualified teachers is now effectively zero. The final Column is the preferred specification, in which we condition on prior performance and instrument with performance in the year prior to this, as we did in the models in Table 3-6 and Table 3-7.

Instrumenting lagged performance almost doubles the inter-period correlation coefficient, from 0.436 to 0.828. As we saw before, school performance, even conditional on observable school and area characteristics, is highly persistent. The performance of a school which is ten percentage points above the average for a school of its type and intake can be expected to fall to five percentage points above similar schools after four years, and to one percentage point above similar schools after 12 years. This long decay makes
clear the importance of persistent unobserved school-level characteristics. Important unobserved differences in intake quality are, however, unlikely given the controls we have available here⁴⁹. We must infer that technologies at the school – technologies such as teaching practice, organisational structure, team cohesion, and leadership style – really can make a sustained difference to pupil achievement.

Taking out the components of teaching and resource inputs which are correlated with prior school performance allows us to make meaningful inferences about the causal effect of teacher and pupil numbers on pupil attainment. The ratio of both qualified and unqualified teachers to pupils increases the probability of pupil success. One extra qualified teacher in one hundred pupils at primary school improves the probability of target attainment at Key Stage 2 by 2.6 percentage points. Additional non-qualified teachers appear to have an even bigger impact, though the stock of ‘other’ teachers is less than 1% of the stock of qualified teachers, so the proportional impact of changes to the stock of qualified teachers is large by comparison. Administration staff and support teachers – support teachers for minority ethnic groups and the small percentage with statements of special education needs – do not have a significant impact on the performance of average pupils.

Primary schools do not appear to have an optimal size, unlike that found by Bradley and Taylor (1998) for secondary schools. In all the specifications, smaller is marginally better – a 0.05 percentage point improvement for ten less pupils – but the effect is statistically insignificant in the IV, lagged dependent variable specification. At the primary level, there are none of the economies of scale that might arise from the increased scope for teacher specialisation at secondary level.

⁴⁹ Replacing the neighbourhood index with the full set of neighbourhood attributes makes no substantive difference to the results
It also worth noting that institutional factors – whether the school is a Community school, (funded and administered by the LEA), or whether it is Voluntary Aided, Voluntary Controlled or Foundation school (all mostly church schools) – have a significant impact on performance. Voluntary Aided and Foundation schools (which make have their own admissions procedures) offer a four to five percentage point advantage over Community schools. Pupils in Voluntary controlled schools do slightly better than those in Community schools, though this advantage could be entirely attributable to parental selection effects since it is lost once we condition on past performance.

3.6 Summary and conclusions

We have sought to explain geographical inequalities in primary school performance through direct effects from catchment area status and through local interactions between schools and pupils. We have also used our information on catchment areas and a school production function model with lagged dependent variable to investigate the dependence of intake and school resources on prior performance, and to re-evaluate the impact of teaching inputs and expenditures.

As we might have expected at the outset, there are strong relationships between pupil attainments and the characteristics of the catchment area in which the school is located. Primary age pupils are more likely to reach target grades at age 11 in schools in higher income neighbourhoods, and the magnitude of this association is constant across regions. Geographical differences in performance are attributable to differences in underlying characteristics, not the sensitivity of performance to these characteristics. In 1999, a ten percent increase in average incomes was associated with 2.5-percentage point improvement in Key Stage 2, Level 4 attainment rates. However, it is by no means clear that this is a causal relationship. After we allow for neighbourhood composition and residential selection on prior school performance, we find no evidence that average
incomes influence school performance, although the well-known relationship between performance and pupil free school meal entitlement persists. Most of the relationship between performance and catchment area characteristics could be attributed to differences in the underlying attributes of the people who live there, rather than the incomes these people receive. An alternative view is that the important attributes — education, age and property size — pick up unobserved permanent income components, which are what really matters in terms of pupil achievements. We cannot differentiate between these two hypotheses here.

We get a clearer descriptive picture if we consider a single index of neighbourhood economic status that combines education, income and employment and demographic information. We can interpret this as a measure of unobserved latent neighbourhood quality or wealth. Pupils in schools that are one standard deviation above the average in terms of this index (including free-school meal entitlement) are around twenty percentage points more likely to achieve Level 4 in Key Stage 2 than those at one standard deviation below the mean, and the marginal returns to neighbourhood status are increasing. This is conditional on school ethnic composition and special educational needs, and allows for institutional differences, general differences between LEAs, and potential selection of residents on prior school performance. Differences in general catchment area status are evidently important determinants of success in schools.

Summary statistics of spatial correlation in school performance reveal that good primary schools tend to be located near other good primary schools and bad schools near other bad schools. This is what we would expect if school performance is in any part determined by the characteristics of pupils and their families, and if these characteristics exhibit spatial autocorrelation. What is interesting here, is that this clustering remains, even after controlling for the characteristics of schools and catchment area residents. What we have shown is that primary school performance is not only related to the spatial distribution of area characteristics, but that it actually depends on the performance of
local schools. This spatial dependence could be a neighbourhood human capital spillover, operating through social interaction of pupils from neighbouring schools, or a knowledge spillover in terms of teaching technologies. There is scope to separate out these effects in future work using new Pupil-Level Annual Census (PLASC) data from the DfES, which holds information on pupils' residential postcodes alongside their school and personal characteristics.

Schools also perform better if they are near other schools, and in areas of higher household density. This, and inspection of maps of residual performance (after adjustment for pupil disadvantages), implies that city schools are more effective than others. This clustering of school effectiveness in cities hints at a role for spillover effects in human capital production in the urban environment: these areas do \textit{better} than other areas with similar levels of intake poverty. The larger pool of employment opportunities in the city may also attract better teachers. Evidence of an underlying sensitivity of school performance to property prices, regardless of the proportion of home-owner/private tenant pupils, reinforces the impression of neighbourhood spillovers: home-owner/private tenant wealth influences the outcomes of social tenants, either through social interactions of children or by enhancing life expectations.

Naïve inference using the English primary school data would lead us to conclude that increasing the number of teachers in schools would worsen pupil attainments, even conditional on school intake measures related to poverty and educational needs. Removing the variation in school performance that is attributable to community characteristics weakens these associations but is insufficient to change their sign. This is an unexpected result, since such negative associations are often attributed to area-based resource allocation. However, we find significant positive effects on teacher inputs once we properly condition on lagged school performance, or work with differences over time. Local education authority level expenditure per pupil has weakly positive effects on performance in all specifications.
Despite these measured positive effects of observable economic resources, their influence is dominated the influence of community and location. What is the relative contribution of teaching resources and school intake to school performance? Clearly, any answer to this depends on what characteristics we define as resource inputs and which we assume are neighbourhood or intake characteristics. What’s more, the result will depend on the number of characteristics and the accuracy in measurement of these variables. Nevertheless, it seems worthwhile trying to make some statement on this issue, based on the characteristics that are significant in the preferred specification (Column (4) in Table 3-8) and that we can reasonably categorize as either resource or neighbourhood inputs.

The model partial sum of squares attributable to observable intake factors – proportion on free-school meals, proportion non-white, the neighbourhood index and rural-urban indicators is 57.74. For observable resource-related factors – teacher-pupil ratios, LEA expenditure and school type – the partial sum of squares is 10.86. These figures only relate to variation in the data that is unrelated to prior school performance. By this calculation, intake and neighbourhood currently explain around than five times as much of variance in primary school performance than do these basic observable resource inputs and school types. This result is hardly changed if we include school age range and size in the school resource set. Of course, there may be plenty of unobserved characteristics of either category, but the indication here is that the background of pupils and the neighbourhood context of the school dominates in explaining the variation in school performance.

Another way of looking at this is in terms of anticipated impacts of changes in the input factors, using the marginal effects in Table 3-8. On this basis, a one standard deviation increase in the number of qualified teachers per 100 pupils (0.486 in 98/99) would increase average school performance by around 1.3 percentage points. This implies an extra 18,000 qualified teachers!. By comparison, a one-quarter standard deviation decrease in the proportion of pupils eligible for free school meals (4% of the school
population or about 150,000\textsuperscript{50}) could increase school performance by 1.8 percentage points, since the marginal effect (unreported in the table) is \(-0.461\). Assuming a qualified primary school teacher’s starting salary of around £20,000 per year\textsuperscript{51}, the cost of the first change would be around £360 million. This is equivalent to £2400 per year for each family of the 150,000 children we needed to get off free-school meals to provide a comparable performance change. In terms of LEA expenditure, an additional 10\% (£180) on current expenditure per pupil would cost around £684 million, providing an performance improvement of only 0.4 percentage points on average.

The reasons for the link between pupil attainments and location are not identified in this work. It is almost tautological to say that some sort of disadvantage of pupil family background is the underlying factor. Still, we do not know whether this operates through abilities, a poor home environment for learning, bad peer group influences, or through the impact of deprivation on the organisational effectiveness of schools. Lupton (2001), for example stresses that poor performance in disadvantaged areas may be a result of the diversion of resources to tackling behavioural and attendance problems, strain on staff-management relations and staff recruitment difficulties.

Area disadvantage is not a sufficient condition for poor pupil attainments, but ineffective school organisation in an environment of pupil disadvantage can be. Although the impact of observed resource inputs is small, there is still a substantial amount of school level variation in performance left unexplained. It is likely that much of this is systematically related to unobserved school-level processes – the leadership qualities of

\textsuperscript{50} The number of primary school children in 1999 was around 3.8 million

\textsuperscript{51} Range of starting salaries: £16,038 - £24,843, dependent on qualifications, age on entry and prior relevant experience (Salary data collected April 2001). Range of salaries at age 40: £24, 843 - £35,648 (classroom teacher) (Salary data collected April 2001). Source CSU/AGCAS Career Services Unit/Association of Graduate Careers Advisory Services
the head and teaching practice in the school, for example – but our national school-level data is fairly silent on these issues. School effectiveness studies attribute between 8-18% of variation in pupil attainments to idiosyncratic school level factors at secondary level (see, for example White and Barber (1997) p.85). The strength of the relationship we observe between voluntary-aided status and primary performance – even conditional on past performance and school intake – indicates a strong role for organisational differences and school ethos, unless these schools really are selecting pupils on the basis of attainment-enhancing characteristics that we just do not observe. The performance advantage suggests these, primarily religious, schools do employ more effective technologies than Community schools. Some of these differences may be attributable to unobserved aspects of the composition of their intake – Coleman (1988), for example, suggests a role of social capital effects in Catholic schools in the US – but this seems unlikely given the robustness of institutional differences to conditioning on past performance. More generally, idiosyncratic unobserved school performance is highly persistent from year to year – even conditional on a wide range of community and intake measures – implying a strong role for systematic school level factors.

52 This effect at secondary level is, of course, the justification for the proposed expansion of the church-school sector proposed in the Government’s White Paper on education (Department of Education and Skills (2000)) p45). The widely discussed implications of this for equity of educational provision across ethnic groups need not be repeated here.
3.7 Appendix A: k-nearest neighbour pseudo catchment area matching

Formally, each school with Postcode grid reference \( c_{1s}, c_{2s} \) is assigned catchment area characteristics estimated by:

\[
\bar{x}_s = \sum_k q_k x_k \quad \text{where} \quad q_k = \begin{cases} 
\frac{P_k}{P_K} & \text{if } k \in M \\
0 & \text{otherwise}
\end{cases}
\]

\( P_k \) is the number of 0-4 year olds in the 1991 Census Enumeration District, \( P_K = \sum_{k \in M} P_k \) and the set \( M \) is defined by:

\[
M = \{ k : \left[ (c_{1k} - c_{1s})^2 + (c_{2k} - c_{2s})^2 \right] \leq \left[ (c_{1K} - c_{1s})^2 + (c_{2K} - c_{2s})^2 \right] \}
\]

\( K \) is the index of the \( K^{th} \) Enumeration District in the spatial data set, where observations are ranked according to their inverse distance from the school location:

\[
d^{-1} = \left[ (c_{1k} - c_{1s})^2 + (c_{2k} - c_{2s})^2 \right]^{-0.5}
\]

The number of nearest neighbours is taken as

\[
K = \frac{P_0}{5P_s}
\]

where \( P_s \) is the number of pupils in the year-6 (age-11) cohort in 1999, and \( P_0 \) is the number of 0-4 year olds (5 cohorts) in the nearest 8 enumeration districts in 1991. Similar techniques applied to the school performance scores give us estimates of the mean performance in the nearest \( J \) schools, for a number of choices of \( J \). A cluster of seven primary schools\(^{53} \) is probably at the top end of what most people would consider as a school neighbourhood, and that a family in a typical neighbourhood might consider for their child. This is also the median number of schools within a 1km radius.

\(^{53} \) And this corresponds to a first-order grouping on an equilateral triangular lattice
3.8 Appendix B: Measures of spatial association

3.8.1 Moran’s I

Moran’s I for a variable $x$ is
\[ I = \frac{\sum_i \sum_j w_{ij} x_i x_j}{\sum_i x_i^2} \]
where $w_{ij}$ is a weight that depends on the proximity of observation $i$ and $j$.

The Local Moran’s I statistic is $I_i = \sum_j w_{ij} x_i x_j$

3.8.2 Rank adjacency statistic

In the current application, this is defined as:
\[ D = \frac{\sum_{r=1}^{n} \sum_{s=1}^{n} w_{rs} |y_r - y_s|}{\sum_{r} \sum_{s} w_{rs}} \]

where $w_{ij} = I\{d_{ij} < d_{ik} \land i \neq j\}$ and $I\{\}$ is the indicator function and $y_r$ is the rank of school $r$ in terms of mean Key Stage 2 pass rates between 1996 and 1999.
3.9 Appendix C

Table 3-9: Factor loadings and scoring coefficients for neighbourhood status

<table>
<thead>
<tr>
<th>Variable</th>
<th>Pooled years in Table 4-4 Loading</th>
<th>Coefficient</th>
<th>Table 4-5, Col. (1) and (3) Loading</th>
<th>Coefficient</th>
<th>Table 4-5, Col. (2) and (4) Loading</th>
<th>Coefficient</th>
</tr>
</thead>
<tbody>
<tr>
<td>Free school meals</td>
<td>-0.680</td>
<td>-0.179</td>
<td>-</td>
<td>-0.877</td>
<td>-0.207</td>
<td></td>
</tr>
<tr>
<td>Free-meals squared</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-0.815</td>
<td>-0.060</td>
</tr>
<tr>
<td>Log mean incomes</td>
<td>0.902</td>
<td>0.488</td>
<td>0.782</td>
<td>0.066</td>
<td>0.748</td>
<td>0.154</td>
</tr>
<tr>
<td>Log mean property prices</td>
<td>0.827</td>
<td>0.235</td>
<td>0.740</td>
<td>0.180</td>
<td>0.597</td>
<td>0.034</td>
</tr>
<tr>
<td>Unemployed claimants per hh</td>
<td>-0.713</td>
<td>-0.198</td>
<td>-0.766</td>
<td>-0.065</td>
<td>-0.781</td>
<td>-0.057</td>
</tr>
<tr>
<td>Aged 8-12 not social housed</td>
<td>-</td>
<td>0.780</td>
<td>0.110</td>
<td>0.788</td>
<td>0.082</td>
<td></td>
</tr>
<tr>
<td>Higher educated</td>
<td>-</td>
<td>0.669</td>
<td>0.041</td>
<td>0.635</td>
<td>0.029</td>
<td></td>
</tr>
<tr>
<td>Mean population age</td>
<td>-</td>
<td>0.405</td>
<td>0.148</td>
<td>0.406</td>
<td>0.113</td>
<td></td>
</tr>
<tr>
<td>Economically active males 25-34</td>
<td>-</td>
<td>0.595</td>
<td>0.034</td>
<td>0.610</td>
<td>0.028</td>
<td></td>
</tr>
<tr>
<td>Economically active males 35-54</td>
<td>-</td>
<td>0.861</td>
<td>0.130</td>
<td>0.855</td>
<td>0.105</td>
<td></td>
</tr>
<tr>
<td>Economically active females 25-34</td>
<td>-</td>
<td>0.684</td>
<td>0.088</td>
<td>0.671</td>
<td>0.071</td>
<td></td>
</tr>
<tr>
<td>Economically active females 35-54</td>
<td>-</td>
<td>0.734</td>
<td>0.087</td>
<td>0.735</td>
<td>0.068</td>
<td></td>
</tr>
<tr>
<td>Lone parents</td>
<td>-</td>
<td>-0.808</td>
<td>-0.133</td>
<td>-0.818</td>
<td>-0.108</td>
<td></td>
</tr>
<tr>
<td>Long term sick</td>
<td>-</td>
<td>-0.701</td>
<td>-0.180</td>
<td>-0.689</td>
<td>-0.141</td>
<td></td>
</tr>
<tr>
<td>One year migrants</td>
<td>-</td>
<td>-0.026</td>
<td>0.016</td>
<td>-0.037</td>
<td>0.010</td>
<td></td>
</tr>
<tr>
<td>Average numb. of rooms</td>
<td>-</td>
<td>0.567</td>
<td>0.051</td>
<td>0.568</td>
<td>0.040</td>
<td></td>
</tr>
<tr>
<td>Households per km^2</td>
<td>-</td>
<td>-0.451</td>
<td>-0.061</td>
<td>-0.484</td>
<td>-0.046</td>
<td></td>
</tr>
<tr>
<td>Agricultural employment</td>
<td>-</td>
<td>0.249</td>
<td>0.016</td>
<td>0.253</td>
<td>0.012</td>
<td></td>
</tr>
</tbody>
</table>

Table 3-10: Positive Eigenvalues of principal factors of catchment area status

<table>
<thead>
<tr>
<th>Factor</th>
<th>Pooled years in Table 4-4 Eigenvalue</th>
<th>Proportion of variance</th>
<th>Table 4-5, Col. (1) and (3) Eigenvalue</th>
<th>Proportion of variance</th>
<th>Table 4-5, Col. (2) and (4) Eigenvalue</th>
<th>Proportion of variance</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2.451</td>
<td>0.938</td>
<td>6.746</td>
<td>0.568</td>
<td>8.083</td>
<td>0.587</td>
</tr>
<tr>
<td>2</td>
<td>0.430</td>
<td>0.165</td>
<td>1.897</td>
<td>0.160</td>
<td>1.904</td>
<td>0.138</td>
</tr>
<tr>
<td>3</td>
<td>-</td>
<td>-</td>
<td>1.452</td>
<td>0.122</td>
<td>1.586</td>
<td>0.115</td>
</tr>
<tr>
<td>4</td>
<td>-</td>
<td>-</td>
<td>1.081</td>
<td>0.091</td>
<td>1.096</td>
<td>0.080</td>
</tr>
<tr>
<td>5</td>
<td>-</td>
<td>-</td>
<td>0.615</td>
<td>0.052</td>
<td>0.619</td>
<td>0.045</td>
</tr>
<tr>
<td>6</td>
<td>-</td>
<td>-</td>
<td>0.482</td>
<td>0.041</td>
<td>0.503</td>
<td>0.037</td>
</tr>
<tr>
<td>7</td>
<td>-</td>
<td>-</td>
<td>0.198</td>
<td>0.017</td>
<td>0.394</td>
<td>0.029</td>
</tr>
<tr>
<td>8</td>
<td>-</td>
<td>-</td>
<td>0.156</td>
<td>0.013</td>
<td>0.200</td>
<td>0.015</td>
</tr>
<tr>
<td>9</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>0.147</td>
<td>0.011</td>
</tr>
</tbody>
</table>

- 148 -
### 3.10 Appendix D

The top panel of Table 3-11 shows coefficients obtained by regressing Postcode sector mean household incomes of social tenants (council or housing association tenants) on the Postcode sector mean incomes of other tenancy groups, with regional or Local Authority area dummy variables. Data is from the Survey of English Housing, 1996-1999. The low coefficients and high standard errors confirm our expectation that incomes of the socially housed are locally uncorrelated across geographical space with the incomes of those in other tenancy groups. Incomes of social tenants living close to richer owner-occupiers are not significantly higher than the incomes of social tenants living in poor residential areas. The incomes of private tenants are, as we would expect, quite strongly associated with the incomes of owner-occupiers – see the lower panel. The distribution of homeowner incomes and private tenant incomes across space is determined by their demands for amenities and housing quality as reflected by prices and rents in the housing market.

**Table 3-11: Relationship between local tenancy group incomes**

<table>
<thead>
<tr>
<th></th>
<th>Regional effects</th>
<th>Local Authority effects</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Council tenants</td>
<td>HA tenants</td>
</tr>
<tr>
<td>Council tenants</td>
<td>-</td>
<td>-0.067</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.047)</td>
</tr>
<tr>
<td>Housing association</td>
<td>-0.018</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>(0.012)</td>
<td></td>
</tr>
<tr>
<td>Owner-occupiers</td>
<td>0.018</td>
<td>-0.035</td>
</tr>
<tr>
<td></td>
<td>(0.018)</td>
<td>(0.026)</td>
</tr>
<tr>
<td>Private renters</td>
<td>0.005</td>
<td>0.044</td>
</tr>
<tr>
<td></td>
<td>(0.011)</td>
<td>(0.027)</td>
</tr>
<tr>
<td>Private rental</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Owner-occupiers</td>
<td>0.360**</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.064)</td>
<td></td>
</tr>
</tbody>
</table>

Significant at *10% level, **0.1% level

Dependent variable is mean proportion in Postcode sector reaching Key Stage 2 Level 4 and above
Incomes are Postcode sector mean incomes
Standard errors adjusted for clustering on Postcode sectors (5687).
All models include Local Authority dummy variables
Models weighted by school size
Instrument is average rooms per household in Postcode sector, from 1991 Census

- 149 -
4 Paying for Neighbourhood Human Capital

4.1 Introduction

Much of the existing empirical work on inter-neighbour spillovers or 'neighbourhood effects' focuses on estimation of the impact of a child's neighbourhood on contemporaneous or subsequent outcomes – typically educational outcomes. The usual approach, is similar to that adopted in Chapter 2 of this Thesis. We find micro data on family and neighbourhood characteristics in childhood, and on outcomes for children for these families, and to apply regression techniques to estimate the effects of neighbourhood conditional on family characteristics. One drawback of this approach is that the important neighbourhood and family characteristics are often highly correlated due to spatial residential sorting attributable to preferences, land prices and housing costs. And we also risk underestimating the long-run impact of neighbourhoods if the characteristics of parents are in part attributable to historical neighbourhood-driven processes. Measurement of these effects of neighbourhood on human capital accumulation is critical for addressing issues of equality of opportunity and the distribution of education, earnings and work across geographical space. Nevertheless, by concentrating solely on these effects we risk ignoring other, potentially substantial, economic costs of neighbourhood deprivation. One example is the cost associated with higher local crime rates in areas where household permanent incomes and employment expectations are low.

Here we look at these issues in a different way. We side-step measurement of direct effects on individual outcomes by looking at the overall value owner-occupier residents place on good neighbourhoods, and use the existence of pockets of social
housing in residential communities in England and Wales as a source of exogenous community stratification. The empirical framework is a hedonic property price model of the type frequently used to value local amenities in the urban, environmental and housing economics literature. Here we apply it to estimate the implicit price of neighbourhood educational composition. We use educational composition as a measure of the local human capital stock, and an index of neighbourhood conditions. In a hedonic equilibrium, this implicit price amounts to a marginal valuation of the services provided by 'educationally rich' relative to 'educationally poor' neighbourhoods. These services could include neighbourhood-related inputs into the production of human capital in residents and their children, direct and indirect effects of local crime rates, and any other local consumption and production externalities.

A number of theoretical models propose community sorting equilibria based on household preferences over some measure of the stock of human capital in the neighbourhood (e.g. Benabou (1996), Nesheim (2002)) and where property prices reflect demands for neighbourhood human capital. However, no existing empirical studies have tried explicitly to estimate the value of neighbourhoods in this way. That is not to say that the impact of neighbourhood on property prices has not been investigated; numerous studies for the US include an estimate of the implicit prices of neighbourhood attributes in property value models (e.g. Kain and Quigley (1970), Schaffer (1972), Berry and Bednarz (1979), Grieson and White (1989), Dubin (1991), Galster, et al. (1999), Lee, et al. (1999), Munneke and Slawson (1999), Ding, et al. (2000)).

The basis of the estimation strategy in this Chapter is that the correlation between property prices in recent transactions and existing local educational composition – across neighbourhoods that are otherwise observationally similar – reflects incoming owner-occupier preferences over the educational composition of their neighbourhood. But clearly, any unobserved differences across neighbourhoods in their utility-bearing attributes – physical size and quality of housing, access to amenities for example –
generates observed differences in educational composition. This occurs because, with imperfect capital markets, wealthier residents purchase greater quantities of any local normal good, and because education is a strong predictor of wealth. This means that educational composition is endogenous in a model in which local property prices are partly determined by local characteristics, unless all utility-bearing local attributes are included in the regression.

To address this problem, we devote a lot of attention to identification of the causal impact of neighbourhood composition on prices, and our paper has a strong methodological emphasis. Our empirical approach exploits variation in the proportion of purpose-built social housing across neighbourhoods as an instrument for neighbourhood educational composition, and — as it turns out — more importantly, exploits variation between neighbourhoods that are closely spatially associated. These methods are generally applicable to the identification of endogenous variables in an empirical property value model.

The paper is structured as follows. Section 4.2 discusses the background to this work and sets it in context. Section 4.3 outlines the standard hedonic property value model in the current context. Section 4.4 describes the data. Section 4.5 explains the empirical methods used. Section 4.6 presents the results. Section 4.7 concludes with an assessment of the size of community returns in relation to mean private returns per household.

4.2 Context

Residents might value neighbourhood education levels or neighbourhood incomes because of the impact on a wide range of outcomes. Property crime rates may be lower, streets safer, the physical environment better maintained, gardens more pleasant, behaviour more orderly. More importantly for families, education levels in the community may matter because of spillovers in the production of human capital in
children. These spillovers could include direct effects from adults to kids through expectations, role models and skills transfers - classified as collective socialisation effects in the sociological literature. They could also include peer group effects that operate through interactions between kids of similar age in the street and at school (Jencks and Mayer (1990), Gephart (1997), Ellen and Turner (1997)). As we saw in Chapter 2 and Chapter 3 in this Thesis, these effects operate to increase the expected educational attainments of children with highly educated neighbours, relative to others.

Rather than specifying all these factors in detail, we can assume that mean education levels provide a sufficient statistic for the distribution of an unobserved composite neighbourhood good, which is the object of preference when considering beneficial community characteristics in the choice of residential location. Let us call the ranking of a neighbourhood on this scale its educational status. This type of composite commodity is closely related to social capital, but as conceived (Coleman (1988)), social capital describes social interactions and community organisational structures that are not exclusively linked to educational attainment of residents in the community.

Although neighbourhood deprivation is multi-faceted, education is a key factor. It is well established that educational attainment is one of the best single predictors of long run earnings and employment. Poor educational attainments obviously mean lower expected incomes for individuals and their families, but there are also high potential external costs, of the type mentioned at the beginning of this section. These external costs of neighbourhood deprivation in education mirror the external social benefits of education, which underpin the principle of public subsidy in educational provision. The Education Reform Act of 1870 which introduced compulsory, publicly funded schooling to Britain, was motivated by liberal conceptions of education's place in a civilised and educated democracy, rather than the need for vocational skills. Nevertheless, most of the empirical work in labour economics and the economics of education focuses only on the private returns to education in the narrowest sense – the increment to earnings from
additional time in education. Others have looked further at social returns conceived as
effects from human capital on production, which increase aggregate output.
Whilst these are interesting issues from the perspective of policy directed to improving
economic performance or addressing inequality, they say little about the value placed by
society on wider educational benefits.

The term *social returns* or *social benefits* is used in a variety of contexts and needs
clarification. In Behrman and Stacey (1997) the phrase is used to “refer to benefits of
education other than the enhancement of labor market productivity and earnings” (page
1). More commonly, the social returns of education are the sum of all private and
external, pecuniary and non-pecuniary benefits, net of any costs of provision. Social and
private returns differ if there are benefits to other members of society arising from an
individual’s school achievements, or participation in higher education. In practice, much
of the literature concentrates on effects on wages if there are *externalities in production*,
whereby the productivity of others is increased by association with more educated
workers – by more productive work relations, by direct transfers of knowledge, or where
an educated community stimulates technical innovation (Moretti (1998), Acemoglu and
Angrist (1999), Moretti (1999), Ciccone and Peri (2000), or see Blundell, et al. (1999) for
a survey).

However, alongside these external benefits that accrue to society through increased
aggregate production and growth we must consider a catalogue of benefits which are
welfare improving, but which may have little or no effect on wages, or output. Some of
these non-market benefits are private benefits – personal health and enjoyment of leisure
time are good examples. But most are public goods that are more or less geographically
localised: social cohesion, citizenship, crime reduction, improved public health. These
types of benefits are surveyed in various books and articles, for example Haveman and
Wolfe (1984), Behrman and Stacey (1997), Schuller, et al. (2001). These benefits, along
with any productive externalities in the formation of human capital, are perhaps better
referred to as the community benefits of education. These are the educational benefits addressed in this Chapter.

Our assumption that the stock of education in the community is a sufficient statistic for households’ evaluation of neighbourhood quality provides our basis for measuring the long-run, social, community-based returns to education. There are some obvious caveats, which we need to discuss here. A reasonable counter-claim is that education merely proxies other behaviours of individuals which are unobserved in the data – drug abuse, vandalism, criminal activity – which impose costs on others in the neighbourhood. We assume that these characteristics originate in lack of human capital: if these characteristics are innate or otherwise fixed prior to educational decisions, and an individual’s educational attainments are determined by these characteristics, then we cannot infer the social returns this way.

One possibility is that parental characteristics and social background generate initial conditions – psychological or economic – which inhibit an individual’s acquisition of education, or mean that any education acquired is valueless in the social context, even under a supportive policy regime. In this setting, educational policy will have relatively small effects on educational outcomes and will have few benefits in the short run. Nevertheless, there may be long run effects if even small improvements in the parents’ generation means a better setting for a child’s acquisition of education. If acquisition of education is mediated solely through genetic or other innate and unalterable characteristics, then we cannot interpret property price effects that originate in preferences for these characteristics or their benefits as monetary realisations of the social benefits of education. Property prices still reflect the perceived benefits of a neighbourhood ‘cleanup’, but the mechanisms for achieving this are not education-based. This view might, for example, find support amongst those who consider criminal behaviour as fundamentally innate, and that lower educational attainments amongst participants in crime is indicative of a preference for crime over legitimate activity. If the
distribution of property prices and education levels are related through fear of crime, or the costs of attacks on property, then the implicit price of educational status measures a transformation of the social benefits of crime-reduction policy.

Willingness to pay for higher-educated neighbours will also overstate the community benefits of education if households place value on their location in the distribution of neighbourhood status. If high-education/high-income households experience no direct costs from living amongst low-education/low-income households, but benefit solely from the status conferred by living in relatively wealthy neighbourhoods, then policies that increase educational attainments by compressing the distribution may inadvertently generate net social costs.

4.2.1 Precedents in the literature

Estimation of neighbourhood impacts on property prices has a long history in the US. Many studies of the factors affecting property values include some neighbourhood characteristics as covariates, though the response to neighbourhood human capital is never the main focus. Kain and Quigley (1970) estimate that prices of owner-occupied housing increase by 7.8%, and rents increase by $2.55 for each additional year of mean adult education in the Census tract, using a small sample in St. Louis. Berry and Bednarz (1979) found that a $1 increase in median Census tract incomes increases the value of single-family homes in Chicago by about $0.70. Both studies condition on a number of neighbourhood and property attributes, but otherwise ignore the potential endogeneity of community-related neighbourhood characteristics. More recently, Dubin (1991) approaches the issue in a different way and uncovers neighbourhood externality effects on prices by analysis of spatial autocorrelation in hedonic models. Hilber (2002) looks at the issue of neighbourhood risk in the context of investment models of home-ownership, and finds that higher variance in neighbourhood noise, crime and litter discourages home-ownership and lowers house prices.
A number of studies look specifically at the effect of social housing projects and other property development on local property prices. An early example is Schaffer (1972), who looks at the impact of housing construction for low income families under the US 1961 “Below Market Interest Rate” scheme using treatment and control sites in Los Angeles. He finds no significant difference between the price trends at the two sites, probably due to the fact that most of the new residents already lived locally. Ding, et al. (2000) are more concerned with the impact of local residential investment on property prices. Using data on Cleveland, Ohio, they find a $0.87 increase in property prices with each $1 of median Census tract income (in 1990) corresponding to an elasticity of 0.36 at the sample mean. Their estimates also imply an elasticity of −0.04 with respect to the Census tract proportion of African-Americans. They also report a negative, but insignificant effect from the proportion in poverty. Crime rates attract a strong negative coefficient, corresponding to an elasticity of −0.13.

Munneke and Slawson (1999) are interested in potential negative externalities from mobile home parks in one parish in Louisiana and estimate a two-step selectivity model to adjust for the endogeneity of mobile home park location. Location within 0.25 miles of a mobile home park in a residential area leads to a 5% decline in the value of a single family dwelling, relative to properties located between 0.25 and 0.5 miles radius. They offer no theory as to the cause of this externality, but perceptions of the behaviours of mobile home residents is presumably a key issue. Two other recent studies investigate the impact of social housing programs in the US. Lee, et al. (1999) consider the effect of public and assisted housing on property values in Philadelphia. The authors of the first paper find negative effects from proximity to public housing developments and other assisted housing schemes, but these effects largely disappear, or their sign is reversed once neighbourhood composition controls are included. They find no statistically significant effects from the physical type of development, which suggests that it is the characteristics of residents and not the physical structure of social housing that generates
the externality. Log property prices increase by 1.6% for each thousand dollars of
neighbourhood median incomes – an elasticity of 0.41. Galster, et al. (1999) look at the
impact on property prices of neighbours in receipt of “Section 8” certificates, which
entitle low-income households to a housing subsidy. They use a model with spatial fixed
neighbourhood effects to find heterogeneous impacts from assisted housing programs in
Baltimore, with adverse effects in lower price areas, but positive impacts from small-
scale programs in higher valued tracts. Interestingly, the authors conducted focus group
studies in four communities with distinct socio-economic compositions, to gauge
residents’ opinions of social housing developments. Some respondents expressed
sensitivity to the physical condition of rental accommodation, with a fear that assisted
housing brought physical decay and vandalism. Many groups expressed clear antipathy to
problem tenants, believing that those in socially assisted housing had different values and
standards than what the current residents desired for their neighbourhood. Many feared
that subsidised housing brought increased crime.

Few studies on the value of neighbourhoods exist for Britain – due to the lack of
data. One example is Cheshire and Sheppard (1995), who find positive amenity values in
Reading from local schools, the proportion of white collar workers and the proportion in
non-Afro-Caribbean ethnic groups. They estimate aggregate land values (over the
geographical space of their sample) of £43,430 attributable to schools, and £81,820
attributable to social and ethnic composition, but offer no estimate of the mean benefits
per household.

No existing studies propose a link between the willingness to pay for good
neighbourhoods and the measurement of community benefits of education. However,
there are a number of approaches to measuring other benefits beyond the traditional
private market returns on earnings. Since Lucas (1988), who discussed the potential role
of human capital externalities in economic growth, a strand of empirical research has
emerged which has tried to measure the impact of state, region or country average
education levels on wages, productivity and growth. A few examples will give the flavour of this research programme.

Weale (1992) uses private returns and international comparisons of growth rates and educational attainments to suggest that long run social returns incorporating spillover effects on growth rates could be two to three times the magnitude of the private returns. Jaffe (1989) looks at the social rate of return to university research in the form of state-specific spillovers into corporate patents, and finds positive effects with elasticities as high as 0.3 in some industries. Acemoglu and Angrist (1999) find strong effects on wages from state education levels, conditional on individual education using OLS estimates on US Census data, but these social returns become weak and insignificant once they instrument the educational variables with state compulsory school attendance laws and individual date of birth effects. Ciccone and Peri (2000) find negative effects from city education levels on individual returns using data from 173 US cities in 1970, 1980 and 1990. Using data on average wages in cities, they find insignificant, near-zero, effects on wages, but small positive effects on productivity of around 1%.

More directly related to the work in hand is Haveman and Wolfe (1984), who present a meta-analysis of earlier work to compute an approximate figure for the annual value of an additional year of schooling based on non-marketed effects on the production of children’s cognitive development, contraceptive use, efficient budget allocations, criminal apprehension and health. Their technique is based on obtaining the shadow price of the non-marketed input from the ratio of its marginal product to that of another, marketed input with a known price. Their figures suggest a value of social and private non-marketed benefits in the order of $5000 in 1975 – a value of a similar order to the annual value of a year of schooling in standard private rate-of-return estimates.
4.3 The hedonic model

We shall use a standard hedonic property value framework to assess the implicit price of neighbourhood educational status, which we take as a measure of the stock of human capital of existing residents in the neighbourhood. For a recent survey of the hedonic approach, see Sheppard (1999). Following the standard hedonic model, we specify household preferences with a utility function:

\[ U = U(c, x, y^*(x), q, l) \]  \hspace{1cm} (4-1)

where \( c \) is a numeraire composite consumption commodity, \( x \) is the measure of neighbourhood educational status. In the empirical work, we will measure this by the proportion of highly qualified residents in Postcode-sectors in England and Wales. The function \( y^*(x) \) is a human capital production function for children in the household, \( q \) is a vector of structural housing characteristics and \( l \) is a vector of locational characteristics.

House prices are determined as a function \( P_h(\cdot) \) of the same attributes, where the attributes are traded at a set of exogenous prices \( \theta \) fixed by demand and supply equilibrium at a broader geographical level. The household lifetime budget constraint is:

\[ y = c + P_h(x, q, l; \theta) \]  \hspace{1cm} (4-2)

Assuming the choice space is continuous so that households can purchase their optimum bundle the first order condition for \( x \) is:

\[ \frac{\partial U}{\partial x} + \frac{\partial U}{\partial c} \cdot \frac{\partial y^*}{\partial x} = \frac{\partial P_h}{\partial x} \]  \hspace{1cm} (4-3)

where the first term on the left captures the direct consumption benefits of \( x \), and the second term captures benefits working through children's increased human capital. This standard condition justifies the use of an estimated implicit price function \( P_h(\cdot) \) in the estimate of the marginal willingness to pay for local educational status. If consumers are heterogeneous in their marginal benefits from \( x \) then stratification into high and low \( x \)
communities can occur, and \( P_a(t) \) can be non-linear in \( x \). Without information on human capital of children, it is not possible to identify separate contributions of educational status to human capital formation and consumption value in the implicit price, only the sum of the marginal benefits. But, we can estimate the overall willingness to pay for marginal improvements in neighbourhood quality, given an appropriate specification of \( P_a \) and data on house prices, neighbourhood, housing and locational characteristics.

4.4 Description of the data

Our empirical methods are designed to suit the data in hand, so it will help to describe the data at the outset. British data on individual property transactions with local area identifiers is not readily available. Instead, we must use locally aggregated data available from the Government Land Registry. This covers most market value property transactions in England and Wales, aggregated to Postcode-sector level. In the UK, Postcodes contain up to seven alphanumeric characters, and contain four hierarchical components. The Postcode-sector, which omits the last two characters of the full Postcode, is the unit of observation in our house price data set and the unit adopted here as a neighbourhood identifier.

This Land Registry data is disaggregated by property type – detached, semi-detached, terraced, flat/maisonette - but this is the only information provided about the characteristics of properties included in the price data. It contains mean house prices and total sales volumes for each dwelling type in each Postcode-sector, where annual sales numbered more 3 or more. Properties under £10,000 and over £1,000,000 are excluded. This amounted to only 0.5% of all property sales in 1999. Sales at non-market value transactions are also excluded. This is an advantage in our application, because market prices will not be contaminated by discounted sales of council houses to tenants under the Right to Buy scheme introduced in the 1980s.
Micro-spatially aggregated data has an advantage over property level data in the current application. This is because we are only interested in the variation in prices attributable to mean neighbourhood characteristics. What we do need though are property prices and neighbourhood attributes at the same level of disaggregation, which rules out most of the more recent measures of neighbourhood status (e.g. Department of Environment (2000)). The obvious source is the 1991 Census Small Area Statistics. This provides a count of the number of over-18s with degrees, diplomas and other high qualifications, based on a 10% sub-sample of the Census population. The age of the Census data and the 10% sampling scheme are a drawback, but in compensation we have a straightforward interpretation of the relationship between property prices and the proportion of highly qualified adults, and the fact that re-aggregation to Postcode-sector level rates is fairly straightforward. In any case, this measure of educational status is quite highly correlated with the 2000 Indices of Local Deprivation. The 1991 Census also provides information on other neighbourhood characteristics and grid references for Enumeration District centroids. For the central results presented in this Chapter, we match the Land Registry property price data for 1995 to 1991 Census data, re-aggregated

---

54 The relevant question is number 19 in the Census form, which asks for details of post-compulsory-age educational qualifications of all persons over the age of 18 in the household on Census night (21-22 April 1991). The Census Small Area Statistics contain counts of persons with higher degrees, degrees, or diplomas, nursing or teaching qualifications, based on a random sample of 10% of the responses from each Census Enumeration District (a much smaller geographical unit than the Postcode-sector).

55 OLS regression shows that one-standard deviation in the spatial distribution of the Ward-proportion highly qualified in 1991 corresponds to 0.65 standard deviations in the Education Domain of the 2000 Index of Deprivation. Instrumenting the proportion highly-qualified using the Ward-proportion in social housing (to correct for measurement error in the 10% Census sample) suggests a one-for one relationship.
from Enumeration District level to Postcode-sector level using the Postcode-Enumeration District lookup tables available from the Census Dissemination Unit.

More national data on property values is available from the Survey of Mortgage Lenders, an annual 5% sample (around 25000) of mortgage transactions. This has the advantage of property level prices, dwelling characteristics and total household incomes, but the disadvantage of broader, Local Authority geographical identifiers. We will use this data in comparison estimates for our main results. A third data set provided by a property value company, Ekins, gives us a sample of property-level characteristics and prices for the London area only. Again, we use this to show that our results are robust to changes in methodology.

4.5 Empirical methods

4.5.1 Empirical problems in the measurement of neighbourhood value

Our objective is to obtain consistent estimates of the partial derivative of the hedonic price function with respect to neighbourhood educational status. In standard hedonic modelling, this means estimating a regression of property prices (or some transformation) on housing and neighbourhood characteristics. Housing supply must be treated as inelastic if the housing characteristics are to be exogenous in the property value model. We may find this assumption acceptable for physical characteristics of housing, at least in the short run. But it makes little sense to apply it to socio-economic characteristics of the neighbourhood – particularly those, like educational status, that are highly correlated with incomes or housing preferences. The hedonic model itself implies that the property prices and socio-economic composition are simultaneously determined. Incomes and preferences track the distribution of housing-related goods across geographical space. Estimates of the model parameters will be biased unless all relevant utility bearing attributes are included in the regression.
The usual strategy seems to be to include socio-economic characteristics as regressors – like mean or median incomes, or ethnic composition – and assume that using these to proxy unobserved neighbourhood or housing characteristics eliminates the bias on the variables of interest. This is explicit in at least one early study Ridker and Henning (1967). But this is a misguided approach. Consider a simple model of property prices with one observed neighbourhood factor $x_i$, an unobserved neighbourhood component $f_i$ (land price, for example) and other unobserved components $\varepsilon_i$.

$$p_i = \beta x_i + p f_i + \varepsilon_i$$  \hspace{1cm} (4-4)

Our interest is in obtaining consistent estimates of $\beta$. The researcher is unsure whether socio-economic characteristics are important in the property value model, so includes a measure of local incomes $z_i$ (or this could be some other measure of neighbourhood status) as a regressor, and proposes to test whether it belongs there using standard statistical tests.

Now, the geographical distribution of incomes tracks the geographical distribution of property prices, although there is independent variation. Let us write local incomes as function of the exogenous factors in the model

$$z_i = \rho x_i + f_i + \omega_i$$  \hspace{1cm} (4-5)

It is immediately obvious that $z_i$ is correlated with the unobserved factors in (4-4), so is unsuitable as a regressor. But the researcher estimates a regression of the following form:

$$p_i = \phi_1 x_i + \phi_2 z_i + \xi_i$$  \hspace{1cm} (4-6)

Substituting for $f_i$ in (4-4), we can see that what's being estimated here is the equation:

$$p_i = (\beta - \rho \gamma)x_i + \rho z_i + \varepsilon_i - \rho \omega_i$$  \hspace{1cm} (4-7)

Coefficient estimates in this model are biased in OLS regression because of the correlation between $z_i$ and the unobserved component $-\rho \omega_i$. What are the expected
values of the coefficients in this regression? It is easy to show that the expected value of
the regression coefficient on $z_i$ is

$$\text{plim} \left[ \varphi_z \right] = \rho \left( 1 - \frac{\sigma_w^2}{\sigma_z^2 + \sigma_w^2} \right)$$  \hspace{1cm} (4-8)

This is clearly not, in general, equal to zero, the structural impact on property prices.
Neither is it, in general, equal to $\rho$, which is what we might be hoping if we are proxying
the unobserved common components $f_i$. Now consider the coefficient on the structural
determinant of property prices and local incomes. Again its is fairly straightforward to
show that:

$$\text{plim} \left[ \phi \right] = \beta - \rho \beta + \frac{\hat{\lambda} \rho \sigma_w^2}{\sigma_z^2 + \hat{\lambda}^2 \sigma^2_i - 2 \hat{\lambda} \gamma \sigma_z^2}$$  \hspace{1cm} (4-9)

where $\hat{\lambda}$ is the regression coefficient obtained from a regression of $x_i$ on $z_i$.

Ordinary Least Squares estimation of a property value model with socio-economic
characteristics as explanatory variables provides inconsistent estimates of all the model
parameters. Since the fundamental purpose of this Chapter is to measure the implicit price
of neighbourhood educational status, we must find a way round this problem The next
sections describes our approach, which combines: 1) elimination of nuisance spatial
variation in prices in the housing market, using a semi-parametric approach to estimating
neighbourhood fixed effects; 2) Instrumental Variables estimation based on exogenous
sources of neighbourhood stratification.

4.5.2 Empirical model

We propose to estimate the implicit price of community educational status
assuming the following simple empirical specification. The log-price of a house of type
$r$ in neighbourhood $i$ is:

$$\ln P_r = \alpha + \beta x_i + \gamma' z_i + g(e_i) + h_r + u_r$$  \hspace{1cm} (4-10)
where $x$ is the proportion highly qualified, $z$ is a vector of *exogenous* housing and neighbourhood characteristics and $g(c_i)$ is an unknown function mapping spatial location $c_i$ to property prices. Dummy variables estimate the property type fixed effect, $h_i$.

Estimation of a full structural specification of the mapping of location $c$ to house prices requires data on local amenities, local housing characteristics, the proximity of neighbourhoods to transport services, local labour demand, environmental quality and other unknown local goods. The function $g(c_i)$ could then be replaced by a specific function of available covariates. In the absence of this data and any prior knowledge about exactly what should be included, we directly estimate the impact of location on property prices and all the other variables in the model. We allow for general effects of location on property prices using non-parametric estimates of the impact of location on our model variables – in effect abstracting from the unobserved area-specific effects on prices, $g(c_i)$.

What we do is compute spatially weighted means of the variables in our model at each observation in the data, in which the nearest observations receive the highest weights. These averages capture general, unobserved, area and amenity impacts on the housing market, *centred at the location of the unit of observation*. We then transform the data into deviations from these spatially weighted, means, and use the transformed variables in our regressions. But we still need to specify how rapidly the weights in our spatial averages decay as we move away in space from a given observation at location. We use a weighting function $m(\cdot | c_i, b)$, with a bandwidth $b$ that specifies how rapidly the weights decay with distance.\(^{56}\) Expressing the model in deviations from these estimated spatial averages, the regression model becomes:

\(^{56}\) Formally, the spatially weighted mean is defined by the Nadaraya-Watson estimator
\[
\ln P_{irt} - m(\ln P_{irt} \mid e, b) = \beta [x_{it} - m(x_{it} \mid e, b)] + \gamma [z_{irt} - m(z_{irt} \mid e, b)] + \omega_{irt}
\] (4-11)

Parameters $\beta, \gamma$ and their variance covariance matrix can then be estimated by applying Ordinary Least Squares to the transformed variables. This smooth spatial effects (SSE) estimator is an application of the semi-parametric partial linear model (see, for example Robinson (1988), Hardle (1990), Stock (1991)).

Bandwidth selection presents an additional problem. As the bandwidth $b$ tends towards zero, the estimator approaches an estimator with Postcode-sector fixed effects. Since we have no time-series variation in the Census education measure this is inappropriate. At the other extreme, an infinite bandwidth is equivalent to the OLS estimator, with the function $g(e_i)$ estimating a constant. Since there is enormous variation in Postcode-sector land areas and household density, a common bandwidth for all observations will lead to inconsistent estimates. The estimated price and regressor surfaces will be over-smoothed in areas of high household density and low land area Postcode-sectors, and under-smoothed in rural areas. To fix this, we weight the neighbourhood bandwidth using data on household density matched in from the 1991 Census. Our baseline results will use bandwidths corresponding to an average of 3400

\[
m(y \mid e_i, b) = \frac{\sum_i y \times k\{e - e_i\} B^{-1}(e - e_i)}{\sum_i k\{e - e_i\} B^{-1}(e - e_i)}
\]

where $B$ is a $2 \times 2$ bandwidth matrix (e.g. $b^2 \times I_2$) and $k\{\cdot\}$ is a multivariate kernel. For the Gaussian kernel this is $k\{\cdot\} = 2\pi^{-1} \exp\{-0.5\psi\}$. For details of multivariate kernels, see Silverman Silverman (1986)

\[57\] Fixing the number of households $n$ in a circular spatial group of radius $b$, gives us a bandwidth weighting rule dependent on housing density $h$: $b = \sqrt{\frac{n}{\pi h}}$
households, but comparisons are made with other bandwidth choices\(^{58}\). Sensitivity to bandwidth choice can be tested by a Hausman test for equivalence of parameters in alternative estimators (Greene (1997)). Too narrow a bandwidth gives a consistent but inefficient estimator; too wide a bandwidth results in inconsistent estimates.

4.5.3 Instrumental variables identification strategy

As discussed above, inclusion of any community characteristics on the right hand side of a hedonic model causes problems. The ability of households to move across space implies that household characteristics are almost certainly endogenous in a property price equation; variation in unobserved determinants of property prices drives the variation in characteristics of residents across geographical space. This is particularly true of educational composition, which is highly correlated with home-owner wealth.

The structure of the problem is common to all endogenous regressor models. The relationship of interest is (dropping unnecessary terms and subscripts from (4-10)):

\[
\ln P_i = \beta x_i + \gamma' z_i + g(e_i) + \rho v_i + \varepsilon_i
\]

(4-12)

Where \(v_i\) is a component of neighbourhood choice which is observed to property buyers, but unobserved to the researcher, and \(\varepsilon_i\) represents components of property price formation which are unobserved to both – optimisation errors, local estate agent activities for example. But neighbourhood status \(x\) is partly determined by migration of home-owners between neighbourhoods, because of selection on unobserved components in the determination of property prices, such as underlying land prices or structural differences.

We can write neighbourhood status as determined by these factors:

\[^{58}\text{Hardie (1990) p.187 presents bandwidth adjustment factors for comparing smoothers using different kernels. The 43\% downward adjustment is based on achieving similar smoothing to a uniform kernel with unit bandwidth, i.e. if we want equivalent smoothing to a uniform kernel of radius encompassing, on average, to two Postcode-sectors (6000 households) we need a Gaussian bandwidth of }0.57*6000 = 3420 \text{ households.}\]
\[ x_i = \lambda z_i + g(\epsilon_i) + \nu_i + \xi_i \]  

(4-13)

Hence, \( E[\rho \nu_i + \xi_{i\nu} | x_i] \neq 0 \) and regression estimates are biased.

Identification of the implicit price of a local amenity, which is not exogenous to other determinants of residential land and building values, requires one of two strategies. Firstly, the traditional approach would be to saturate the model with property descriptors and exogenous local characteristics. But, determinants of property prices that are left unobserved to the researcher must also be unobserved to property buyers, or considered irrelevant, if they are to be truly exogenous to own education and wealth.

Estimation on differences from local expected values (our SSE estimator) partly overcomes the identification problem by removing most of the variation attributable to local labour markets, local environmental goods, and transport services. This assumes that \( \nu_i \) is subsumed in \( g(\epsilon_i) \). A better approach is to combine this estimation strategy with Instrumental Variables for the local characteristic of interest, using some Columns of the vector \( z_i \). In the current context, we need local characteristics which are correlated with local educational composition or local incomes, but which affect the education and incomes of home buyers only through their influence on local education or income, valued as an amenity. Candidate instruments are the proportion and characteristics of households in social housing. Current proportions in social housing are determined largely by the stock of social housing, which in turn is historically determined by the location of council estates and other social housing developments. It is primarily a historical, policy-driven characteristic.

The identifying assumption is that the incomes of home-owners are locally uncorrelated with the proportion, incomes and education of tenants in social housing – except in so far as the presence of low-education tenants in social housing generates an externality that home buyers are willing to pay to avoid.
One alternative instrument is the location of higher education institutions, which generate highly educated academic enclaves, unrelated to any local amenity value (assuming demand for proximity to universities does not itself produce a price premium).

Incorporating these instruments into our smooth spatial effects estimator, we have:

\[
\hat{\beta}_{IV-SSE} = \left[ \tilde{X}' \tilde{Z}' \hat{Q} \tilde{Z} \tilde{X} \right]^{-1} \tilde{X}' \tilde{Z}' \hat{Q} \tilde{Z} \tilde{p}
\]  \hspace{1cm} (4-14)

where \( X \) is the regressor matrix, \( Z \) is the full instrument matrix and \( p \) the house-price vector. The tilde indicates deviations from the non-parametric estimates of the smoothed surface means in the smooth spatial effect models. The matrix \( \tilde{Z}' \hat{Q} \tilde{Z} \) is estimated using the Huber-White method, with clustering on Postcode-sectors to allow for the fact that we have multiple house type in each Postcode-sector. The variance covariance matrix is estimated by the inverse of the first term in square brackets.\(^{59}\)

### 4.6 Empirical Results

#### 4.6.1 Summary and assessment of the data

**4.6.1.1 Property price data**

Results in this Chapter are presented separately for three broad geographical regions of England and Wales. These regions correspond to grouped Standard Statistical Regions:

- **South East and East**: London, South East (rest), East Anglia
- **The North**: East Midlands, Yorkshire and Humberside, North, North West
- **South West and West**: West Midlands, South West, Wales

\(^{59}\) We compare with bootstrap standard error estimates in one case to assess the accuracy of the standard errors.
Table 4-1 summarises the main variables in our data set. The property price sample includes only those properties with recorded Postcodes. This sub-sample slightly under represents higher price properties in 1996 when compared with the full sample used by the Land Registry or the random 5% sample conducted by the Society of Mortgage Lenders. The Postcode-sector data under-represents higher priced detached housed and flats in all regions, because it under represents new high-end properties without Postcodes. Given that the difference between the means in the Postcode sample and the full sample is only around 5% this should not be a serious problem.

Table 4-1: Summary statistics for local education and property prices

<table>
<thead>
<tr>
<th></th>
<th>South East and East</th>
<th>The North</th>
<th>South West and West</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean/s.d</td>
<td>Min/Max</td>
<td>Mean/s.d</td>
</tr>
<tr>
<td>1995 sector mean price</td>
<td>80435</td>
<td>(50139)</td>
<td>49245</td>
</tr>
<tr>
<td></td>
<td>10916</td>
<td>776000</td>
<td>10714</td>
</tr>
<tr>
<td>Mean annual sector sales volume</td>
<td>131</td>
<td>(74)</td>
<td>92</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>584</td>
<td>3</td>
</tr>
<tr>
<td>Postcode-sectors</td>
<td>2965</td>
<td>-</td>
<td>2858</td>
</tr>
</tbody>
</table>

Means of property prices are means of Postcode-sector means, weighted by sales volumes
Means of qualifications and social housing weighted by households in 1991

4.6.1.2 Assessing the instruments

As discussed in 4.5.3, identification of the implicit price of neighbourhood education or neighbourhood incomes requires an Instrumental Variables approach. Our main instrument is the Postcode-sector proportion of households in social housing. Only 2.7% of social tenants claimed to have higher education qualifications in the 1991 Labour
Force Survey, compared to 17.5% in the private housing sector. Inclusion of the proportion of these social tenants in ethnic groups originating from the Indian subcontinent gives some overidentification, since average educational attainments are higher for these groups than for other social tenants (5.7% with higher education in 1991, according to LFS data). Both these instruments are highly significant in first stage regressions of educational composition on the exogenous variables and instruments, with an F-statistic of 240 for the South West and West, 392 for the South East and East, and 321 for the North of England. The within-area R²s are 0.32, 0.41 and 0.38 respectively.

Our identifying assumption is that education and incomes of home-buyers and social tenants are locally uncorrelated, except through the influence of the proportion of low-education/low-income social tenants on property prices. Obviously this will not be true over larger areas, in which case differences in labour market opportunities and earnings will affect home-owners and social tenants jointly. Estimation within highly localised geographical groups makes it unlikely that these instruments are uncorrelated with unobserved determinants of property values. Further checks using the 1994 to 1998 Survey of English Housing show that Postcode-sector mean incomes of neighbouring social tenants and property owners are uncorrelated within Local Authority areas. Even more convincing, our Survey of Mortgage Lenders data tells us that the incomes of households taking out new mortgages are uncorrelated with our main instrument (the local proportion in social housing) once we control for local educational status.

4.6.1.3 Changes between 1991 and 1995

Complete data on Postcode-sector property prices is only available from the Land Registry since 1995. The Census data on tenancy groups and local qualifications dates from 1991. Our estimates will be biased if there were substantial changes in area characteristics between 1991 and 1995. If the higher-educated proportion in each Postcode-sector in 1995 is just a multiple of the 1991 value, then we need to adjust the
coefficient on the 1991 proportion downwards. However, changes in the distribution of education across Postcode-sectors will attenuate the estimates, since it introduces a form of measurement error. Comparing 1991 and 2000 deprivation indices at Ward level, indicates that the attenuation due to distributional changes is only in the order of 8% – see Appendix A.

We should also be concerned if the proportion of households in social housing changed dramatically between 1991 and 1995. This would make our instrument a poor predictor of neighbourhood status at the time of the property transactions in 1995. Construction of social housing was at a low level in the early 1990s and any the changes in the local proportion in social housing would be due to sales to the private sector. In fact, the total stock of households in the social rented sector (Council and Registered Social Landlords) remained fairly stable over the 1990s falling from 4.7 million in 1990/1 to 4.6 million at the end of the decade (Department of Transport (2000)).

4.6.2 Implicit price of neighbourhood educational status

4.6.2.1 Postcode-sector data results

Table 4-2 presents the central estimates from the smoothed-spatial effects estimators, by the three broad geographic regions. In each region-specific panel of the table, Column (1) presents a basic OLS regression of log mean property prices in Postcode-sector-dwelling-type cells in 1995, on the proportion of highly qualified adults in the Postcode-sector in 1991, dwelling type dummies, mean rooms in owner-occupied housing, and the density of purpose built flats. The last two regressors are included as controls, to ensure that our social-housing-based instruments are not picking up aversion to high-density purpose built housing, or prejudice by the white majority against minority ethnic groups. Column (2) shows estimated parameters and standard errors from the smoothed-spatial effects estimator (SSE), with the same regressors. In Column (3), we apply the IV procedure. All the tables show estimated coefficients and robust standard
Just to illustrate our empirical method, Figure 4-1 shows the smoothed proportion with higher education qualifications in the London area, that is the function $m(x, | c, b)$ in (4-11). Estimation is based on deviations of the raw proportion higher-educated from this surface.

**Figure 4-1: Smoothed Postcode-sector proportion with high qualifications: London and surrounding area 1991**

Spatial bandwidth 1 km, Gaussian kernel, Eastings 50000 to 56000, Northings 15000 to 21000

---

60 Standard errors for the smooth spatial effect models were checked against bootstrap standard errors for a London sub-sample. For a coefficient of 1.398 the standard error was 0.313 or 0.327 when bootstrapped (100 repetitions).
Table 4-2: Property price response to neighbourhood educational composition: Postcode-sector data, by region, 1995

<table>
<thead>
<tr>
<th></th>
<th>South East and East</th>
<th>The North</th>
<th>The South West and West</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>OLS</td>
<td>OLS-SSE</td>
<td>IV-SSE</td>
</tr>
<tr>
<td>Proportion highly qualified in 1991</td>
<td>3.663</td>
<td>1.152</td>
<td>1.119</td>
</tr>
<tr>
<td>Density of purpose built flats (100s/km²)</td>
<td>9.0 e-3</td>
<td>0.2 e-3</td>
<td>0.2 e-3</td>
</tr>
<tr>
<td>Proportion Black, Indian, P’stan, B’desh</td>
<td>0.269</td>
<td>-0.973</td>
<td>-0.979</td>
</tr>
<tr>
<td>Mean rooms in owner-occupied housing</td>
<td>0.020</td>
<td>0.208</td>
<td>0.209</td>
</tr>
<tr>
<td>Detached</td>
<td>0.463</td>
<td>0.470</td>
<td>0.470</td>
</tr>
<tr>
<td>Flat/Maisonette</td>
<td>-0.537</td>
<td>-0.563</td>
<td>-0.563</td>
</tr>
<tr>
<td>Terraced</td>
<td>0.176</td>
<td>-0.175</td>
<td>-0.175</td>
</tr>
<tr>
<td>Constant</td>
<td>10.447</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Overall R²</td>
<td>0.725</td>
<td>0.2925</td>
<td>0.2925</td>
</tr>
<tr>
<td>P-value test of restrictions</td>
<td>-</td>
<td>-</td>
<td>0.840</td>
</tr>
</tbody>
</table>

Dependent variable is log of property price

Instruments in Columns 4 & 6 are Postcode-sector proportion in social housing and proportion of social housing tenants from Indian, Pakistani, and Bangladeshi ethnic groups

Sample size (sectors x property type): 9431 in South East and East, 8082 in The North, 6058 in the South West and West samples

Min, mean, max bandwidth: .24 km, 1.27 km, 8.75 km in South East and East; 28 km, 1.44 km, 22.90 km in The North; 0.40 km, 1.69 km, 12.57 km in the South West and West

The sample centroids are, by region: (52854, 18343), (41272, 41884), (34637, 20930)

- 175 -
To start our discussion of the results, let us focus on the South East and East. The standard OLS estimates in Columns 1 illustrate the partial correlations between property prices and the regressors in the South and East of England. But, these numbers cannot be interpreted as structural parameter estimates in a model of property price determination. Differences within the region in labour market returns to skills and employment opportunities will simultaneously determine property prices and educational composition. As soon as we estimate on deviations from local means using the semi-parametric SSE model, the coefficient on highly qualified residents falls dramatically, to 1.152 (0.143) in Column (2). Unsurprisingly, taking out the mean differences between localities removes biases in the estimate of the implicit price of introduced by local labour market driven property price and educational composition simultaneity.

The Instrumental Variables estimate in Column (4) is only slightly below this at 1.119 (0.209). This is a somewhat surprising result, because any unobserved differences between neighbouring Postcode-sectors in the mean physical characteristics of housing should generate variation in mean property prices, and, we might expect, variation in the mean education and incomes of purchasers of these properties. The similarity implies that this source of endogeneity is not a serious problem once we consider only highly localised variation in our SSE model. Measurement error in the educational status variable may be another factor that leads to higher IV estimates. The educational status variable is taken from the 10% sample of the Census, so the sampling variance is high. The variables used as instruments come from the 100% sample. One point to note is that, according to our discussion in Section 4.5.3, the ethnic composition of the neighbourhood could be endogenous: lower income groups select lower price neighbourhoods. We show the results with this variable included, but

---

61 Comparison of 10% and 100% sample unemployment rates, for example, suggests that 20% of the variance of the 10% sample is sampling noise
removing it makes almost no difference to our IV estimates. Excluding ethnic composition as a regressor increases the parameter of interest in the SSE model by around 10%, but in the IV-SSE model we find virtually no change. The insensitivity of the IV-SSE estimate to the inclusion of both ethnic group and flat density controls is a good indication of the validity of the instruments.

Moving across Table 4-2 to the estimates for the North of England regional group, we see that the point estimates fall by around 20% once we abstract from local area effects (Column (6) against Column (5)). Instrumenting the proportion higher-educated increases the coefficient estimates (Column (7)), an effect which can surely only be attributed to the sampling noise in the regressor. This difference between the SSEIV and SSE estimates of the key parameter is nevertheless insignificant in a Hausman test \( \chi^2(0) = 0.654 \). The final estimates of our key coefficient in the North are almost double those obtained for the South and East at 2.163 (0.128).

Results for the South West and West are much like those for the North. On this sample, however, we reject the null hypothesis that the overidentifying restrictions and model specification are correct (the estimated residuals are correlated with the instrument vector – see the \( \chi^2 \) test in the bottom row). The reason for this misspecification is unclear, but is probably linked to the highly rural geography of Wales and the South West peninsula. Given the similarity between the IV-SSE and SSE model estimates in the other regions, we can quite reasonably assume that the non-IV based estimates are acceptable here too.

Comparing the parameters across regions, an obvious point is that the response to percentage point changes is markedly different across regions. This is, of course, largely related to differences in mean education levels between regions, since valuations of neighbourhood education level are likely to be based on local expectations. Converting the coefficients into elasticities at the sample mean we get 0.211 (0.035) for the South and
South East, 0.276 (0.044) for the North, and 0.222 (0.041) for the West, SW and Wales. A minimum distance estimate of the elasticities is 0.237 (0.039), and we do not reject equality to the minimum distance parameter for all regions ($p$-value = 0.356). According to these estimates, a one percent relative improvement in educational status of an average neighbourhood – as measured by the proportion with higher education qualifications – is valued at around £150 in year-2000 prices.

4.6.2.2 Sensitivity to bandwidth choice

Since we have no prior information on the ‘best’ bandwidth to use to define the local area groups in the SSE and SSEIV models, we need to check how the parameters vary with bandwidth choice. We should be worried if the estimates change dramatically for small changes in bandwidth, as this would invalidate the claim that this uncovers the parameters of a model of property price determination operating at the household level. In principle, an optimal bandwidth could be chosen based on a loss function which makes a compromise between bias and efficiency – as the bandwidth increases, efficiency increases but at the risk of biased parameter estimates. A slightly more ad-hoc approach is to re-estimate the models at intervals above and below the 3400 household bandwidth used in the main tables. The results of this exercise are in Table 4-3.

In all regional groups, the SSE estimates (without instruments) decrease steadily as bandwidth is reduced from 5100 to 850 households. This is to be expected, as sampling error in the 10% Census sample leads to increasing attenuation in the estimated coefficients as we remove across-space variation.

---

62 The elasticities are even closer across regions if we constrain the elasticities to be constant within regions by estimating a double-log model: SE&E 0.214 (0.018); North 0.195 (0.013); SW&W 0.230 (0.019). The problem with this specification is that it implies near zero property prices in areas with near zero proportions with high qualification, and is inconsistent with the evidence presented later on the log-linearity of the property-price/proportion-highly-qualified regression line.
Table 4-3: Sensitivity of education parameter estimates to bandwidth choice

<table>
<thead>
<tr>
<th></th>
<th>South and East</th>
<th>North</th>
<th>West, SW and Wales</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>OLS</td>
<td>IV</td>
<td>OLS</td>
</tr>
<tr>
<td>850 households</td>
<td>0.687</td>
<td>1.082</td>
<td>1.574</td>
</tr>
<tr>
<td></td>
<td>(0.277)</td>
<td>(0.383)</td>
<td>(0.233)</td>
</tr>
<tr>
<td>Sargan test p-value</td>
<td>-</td>
<td>0.28</td>
<td>-</td>
</tr>
<tr>
<td>Mean bandwidth</td>
<td>0.64 km</td>
<td>0.72 km</td>
<td>0.85 km</td>
</tr>
<tr>
<td>1700 households</td>
<td>0.926</td>
<td>1.081</td>
<td>1.818</td>
</tr>
<tr>
<td></td>
<td>(0.214)</td>
<td>(0.280)</td>
<td>(0.218)</td>
</tr>
<tr>
<td>Sargan test p-value</td>
<td>-</td>
<td>0.79</td>
<td>-</td>
</tr>
<tr>
<td>Mean bandwidth</td>
<td>0.90 km</td>
<td>1.02 km</td>
<td>1.19 km</td>
</tr>
<tr>
<td>3400 households</td>
<td>1.152</td>
<td>1.119</td>
<td>1.924</td>
</tr>
<tr>
<td></td>
<td>(0.143)</td>
<td>(0.209)</td>
<td>(0.170)</td>
</tr>
<tr>
<td>Sargan test p-value</td>
<td>-</td>
<td>0.84</td>
<td>-</td>
</tr>
<tr>
<td>Mean bandwidth</td>
<td>1.27 km</td>
<td>1.44 km</td>
<td>1.69 km</td>
</tr>
<tr>
<td>5100 households</td>
<td>1.244</td>
<td>1.170</td>
<td>1.954</td>
</tr>
<tr>
<td></td>
<td>(0.119)</td>
<td>(0.184)</td>
<td>(0.144)</td>
</tr>
<tr>
<td>Sargan test p-value</td>
<td>-</td>
<td>0.619</td>
<td>-</td>
</tr>
<tr>
<td>Mean bandwidth</td>
<td>1.56 km</td>
<td>1.77 km</td>
<td>2.08 km</td>
</tr>
<tr>
<td>Hausman tests:</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1700 against 850</td>
<td>$\chi^2(1) = 0.00$</td>
<td>$\chi^2(1) = 0.075$</td>
<td>$\chi^2(1) = 0.74$</td>
</tr>
<tr>
<td>3400 against 1700</td>
<td>$\chi^2(1) = 0.34$</td>
<td>$\chi^2(1) = 0.36$</td>
<td>$\chi^2(1) = 5.1$</td>
</tr>
<tr>
<td>5100 against 3400</td>
<td>$\chi^2(1) = 0.26$</td>
<td>$\chi^2(1) = 0.37$</td>
<td>$\chi^2(1) = 3.4$</td>
</tr>
</tbody>
</table>

By contrast, for the South East and East, and North regional groups the SSEIV estimates are remarkably stable. The IV estimate for the South East and East increases by only 8.5%, and the estimate for the North increases by only 15% as we increase the bandwidth by a factor of 6. Hausman tests of the difference between pairs of estimates computed at different bandwidths all fail to reject the null of equality. For Wales, West and South West of England, the IV estimates are not stable across different bandwidth choices and the Sargan test statistics suggest a misspecification. The non-IV estimates are, however, relatively insensitive to changes in bandwidth around 3400 households, and are consistent with the elasticities calculated for the other regions, so we take these as the preferred estimates for this region.
4.6.2.3 Comparison with property level data

Some readers will feel uncomfortable with results based on micro-spatially aggregated data, without any controls for individual housing or owner-characteristics, despite the spatial identification strategy employed here. To test the robustness of the results, we really want to observe the incomes of home purchasers and estimate the models conditional on own household incomes. As discussed in Section 4.4, our second property price data set from the Survey of Mortgage Lenders (SML) has property prices and the household incomes on which the mortgage is based, but no neighbourhood identifiers. Nevertheless, we can use it to estimate the relationship between prices and broader Local Authority educational status, for which we have identifying codes. The estimates based on matched Census-SML data are tabulated in Table 4-4, for 1997 property data for England.

Table 4-4: Property price response to local education: SML data, by region, 1997

<table>
<thead>
<tr>
<th></th>
<th>SE and East</th>
<th>The North</th>
<th>SW and West</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>OLS</td>
<td>IV</td>
<td>OLS</td>
</tr>
<tr>
<td>Local Authority proportion with higher education</td>
<td>2.210 (0.194)</td>
<td>1.625 (0.514)</td>
<td>2.050 (0.050)</td>
</tr>
<tr>
<td>Own log-income</td>
<td>0.435 (0.017)</td>
<td>0.445 (0.018)</td>
<td>0.380 (0.013)</td>
</tr>
<tr>
<td>R²</td>
<td>0.713</td>
<td>0.711</td>
<td>0.723</td>
</tr>
<tr>
<td>Sample size</td>
<td>11085</td>
<td>7110</td>
<td>4367</td>
</tr>
<tr>
<td>Mean log-price</td>
<td>11.306</td>
<td>10.886</td>
<td>11.035</td>
</tr>
<tr>
<td>Mean price</td>
<td>£95577</td>
<td>£61369</td>
<td>£70638</td>
</tr>
</tbody>
</table>

Dependent variable is log of property price
All models include: main purchaser age, number of males, number of females, bungalow, detached, semi-detached, terraced, flat/maisonette (converted), flat/maisonette (purpose built), other dwelling type, built 1919-39, 1940-60, 1961-1980, after 1980, new, number of rooms, dwelling type x number of rooms, dwelling type x property age, County dummies
Minimum distance estimate of IV coefficient on educational status = 2.010 (0.352); test of equality across regions p-value = 0.630
Without own-incomes as control, IV coefficients on local educational status are:
SE&E: 1.681 (0.877), North: 2.353 (0.387), SW&W: 4.008 (1.205)
Minimum distance estimate = 2.384 (0.413); test of equality across regions p-value = 0.651
In Table 4-4, the instrument for the Local Authority proportion higher-educated is just the proportion in social housing. The regressions include a broad set of property type interactions and household characteristics, as listed in the table notes. Looking at the Table and comparing with Table 4-2 we see that the OLS estimates are somewhat higher for the South East and East and the West and South West regions, but similar for the North. Instrumenting local education brings the coefficient down in the South East and East, makes little difference in the North and increases it in the South West and West, but none of these coefficients is significantly different from the OLS estimates. If we work with elasticities at the mean and calculate the minimum distance estimate from the IV coefficients across all samples and regions we get an elasticity of 0.250, with equality across regions (p-value = 0.254). Overall, the coefficients estimated using property level data with own income, plus more property and household characteristics, are entirely consistent with the estimates of more localised human capital effects in the main tables.

A really important point to note from these results is that IV estimates are similar whether or not we include own-incomes, which is what we would expect of the instruments are working. This is another strong indication that the social housing instrument is exogenous to home-purchaser incomes. The minimum distance IV estimate of the impact of educational status unconditional on incomes is 2.384 (0.413), or 2.010 (0.352) conditional on own incomes. Without IV, the corresponding parameters are 3.058 (0.274) and 2.067 (0.051).

We make one further comparison, using our valuation property data for London. Again, this records information for individual properties, this time for 2001. The full sample exceeds 8000, but we restrict the sample to observations with non-missing. Unfortunately we do not have information on purchasers’ demographics, but we have a richer set of property characteristics and can match neighbourhood and local amenity data by exact Postcode. Because we have fairly precise grid references for each property, we can construct accurate measures of neighbourhood education and social housing by
averaging Enumeration-District-level Census data around the property grid-reference\textsuperscript{63}.

Table 4-5 shows the results of OLS and IV regressions using this data. As before we instrument with the neighbourhood proportion in social housing.

**Table 4-5: Results for London using property-level data, 2001**

<table>
<thead>
<tr>
<th></th>
<th>OLS</th>
<th>IV</th>
</tr>
</thead>
<tbody>
<tr>
<td>Neighbourhood proportion with higher qualifications in 1991\textsuperscript{1}</td>
<td>1.433</td>
<td>1.615</td>
</tr>
<tr>
<td></td>
<td>(0.201)</td>
<td>(0.446)</td>
</tr>
<tr>
<td>R\textsuperscript{2}</td>
<td>0.764</td>
<td>0.764</td>
</tr>
<tr>
<td>Sample size</td>
<td>3803</td>
<td>3803</td>
</tr>
</tbody>
</table>

Note: \textsuperscript{1} this is a spatial average of neighbouring Census Enumeration Districts (EDs), excluding the ED in which the property is located.

Instrument in Column (2) is the local spatial average proportion in social housing.

Additional neighbourhood regressors are: Enumeration District proportion in social housing, mean house size, population density, household density, purpose built flat density, distance to Soho and distance squared, distance to nearest underground station, distance to nearest park, distance to nearest town centre, distance to nearest river crossing, secondary school performance, primary school performance and Local Authority dummy variables.

Additional property characteristics are: 12 style dummies, year built, number of rooms, number of floors, floor area, garage. Month of sale

These numbers are slightly higher than we found earlier, but then we are using 2001 property data matched to 1991 qualifications data. There is no difference between the IV and OLS estimates in the statistical sense. What we achieved using the Postcode-sector data and our SSE strategy, we achieve here by including a much broader set of exogenous characteristics and more ad-hoc spatial controls.

4.6.3 Non-linearities in response

The model in (4-10) imposes a log-linear relationship between property prices and the proportion of highly qualified neighbours. Some empirical verification of this assumption is in order. Non-linearities in the implicit price function will have

---

\textsuperscript{63} The technique used is described in more detail in Chapter 6, where we make full use of this data set to estimate the impact of local crime rates on property prices.
implications for evaluation of the aggregate social benefits of an increase in educational attainments. Evidence of non-linearities may further enrich the empirical analysis by suggesting threshold effects in the spirit of 'contagion' theories of neighbourhood deprivation which have been popular in the sociological literature - see Crane (1991), Buck (2001), and McCulloch (2001) for examples. If these ideas are correct, then homeowners should be fairly indifferent to neighbourhood educational composition until education levels fall below some critical threshold.

In order to check for such non-linearities in a non-parametric fashion, we can generalise the procedure above, to estimate the linear parameters in the function:

\[ \ln P_{ir} = g(x_i, c_{1i}, c_{2i}, b_i) + \gamma w_{ir} + u_{ir} \]  

(4-15)

where \( w_i \) is the vector of characteristics assumed to enter linearly in the hedonic log-price function. We then compute the averaged kernel regression line from a regression of \( (\ln P_{ir} - \hat{\gamma} w_{ir}) \) on \( x \) at a sample of our coordinate grid points \( c_1, c_2 \). Details of the method are in Appendix C.

Figure 4-2 shows an example of this exercise for the London area. Results for other regions are similar. We present the kernel regression line based on log prices as the dependant variable, with bootstrap confidence intervals. The figure also shows the log of the kernel regression line obtained using prices untransformed. Either way, it looks like the semi-log parametric specification is acceptable. Apart from a few local irregularities, the relationship between the natural log of property prices and the generated neighbourhood proportion with higher education is linear. There are no threshold effects or discontinuities. Nothing indicates that people are willing to pay proportionally more as educational status increases, though of course, the absolute amount paid increases with each one percentage point shift in educational status.
4.6.4 Comparing local income and education effects

Can we determine whether income or education is more valued in the neighbourhood social environment? Education might be important if residents seek out the productive externalities in human capital formation; if income is more important, we might emphasise considerations such as lower crime rates and a well maintained physical environment. Good data on local incomes is hard to come by in the UK, but we have access to a commercial data set of Postcode-sector mean incomes surveyed by a marketing company (CACI Ltd.) in 1996 and 1999.

Disentangling income and education effects in micro data is notoriously difficult, and the problem is exacerbated here by the high correlation between incomes and education in spatially aggregated data. If we include neighbourhood incomes and
education together in property value regressions for pooled regions, and instrument with characteristics of tenants in social housing, we get coefficients of \(-0.017 (0.16)\) on log-incomes, and \(2.163 (0.592)\) on the proportion higher educated education. The impact of education dominates, despite the instruments being significant in first stage education and income equations.

Repeating the main analysis using log-incomes instead of the proportion higher-educated gives estimated elasticities of prices with respect to mean incomes of \(0.519 (0.054)\) for all regions – see Appendix B. What stands out from this result is that the coefficient on the proportion with higher education qualifications in the main tables is much higher than what we would expect if neighbourhood mean income was really the principle object of preference. Conventional estimates of the private returns to higher qualifications are around 0.25 to 0.3 for higher education qualifications (e.g. Blundell, et al. (1999)), so substituting the neighbourhood proportion higher-educated for log-mean incomes in our property value models should give us a coefficient of less than \(0.519 \times 0.3\). Our actual estimates are ten-times this figure! It seems probable on this evidence, that education is valued as a local commodity for other reasons than just its impact on average incomes. We conjecture that mean neighbourhood income merely acts as a noisy proxy for the underlying educational status of the area.

4.6.5 Robustness to unobserved neighbourhood heterogeneity

The models presented in the main tables use relatively few right hand side controls, and rely on the spatial effects and instruments to achieve identification. We can, of course, just include more community characteristics and observe what happens to the estimated coefficient on the proportion higher-educated. Table 4-6 presents results from this exercise, assuming we can treat the characteristics as exogenous. Row 1 shows the baseline model, comparable to the results in the main tables, although for ease of
computation we dispense with the smooth spatial effects and assume simple Postcode-district geographical fixed effects.

Table 4-6: Robustness to neighbourhood heterogeneity, pooled regions

<table>
<thead>
<tr>
<th>Additional controls</th>
<th>Educational status</th>
<th>Sample Size</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>( \beta )</td>
<td>Elasticity</td>
</tr>
<tr>
<td>1 Dwelling type, year, proportion non-white, average</td>
<td>1.852</td>
<td>0.270</td>
</tr>
<tr>
<td>rooms, purpose-built-flat density, Postcode-district</td>
<td>(0.063)</td>
<td>(0.009)</td>
</tr>
<tr>
<td>fixed effects</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2 Model A, plus proportion of social households in</td>
<td>2.213</td>
<td>0.324</td>
</tr>
<tr>
<td>1981: IV using 1991 proportion of social households</td>
<td>(0.226)</td>
<td>(0.033)</td>
</tr>
<tr>
<td>3 Model A, IV using location of higher education</td>
<td>1.530</td>
<td>0.223</td>
</tr>
<tr>
<td>institutions</td>
<td>(1.190)</td>
<td>(0.173)</td>
</tr>
<tr>
<td>4 Model A, plus mid-year unemployment per household</td>
<td>1.382</td>
<td>0.202</td>
</tr>
<tr>
<td>(NOMIS), household density per km(^2), proportions</td>
<td>(0.078)</td>
<td>(0.011)</td>
</tr>
<tr>
<td>in agricultural employment, one year migrants, lone</td>
<td></td>
<td></td>
</tr>
<tr>
<td>parents, long-term sick, age 65 plus</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5 Model C, plus primary school performance (1996),</td>
<td>1.378</td>
<td>0.200</td>
</tr>
<tr>
<td>proportion of pupils with statements and special</td>
<td>(0.079)</td>
<td>(0.012)</td>
</tr>
<tr>
<td>educational needs (Wales excluded)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Dependent variable is log of property price, 1996 property price data
School data taken from DfES 1996 Primary School Performance Tables
Unemployment counts from Nomis, all other controls from 1991 Census

Row 2 of Table 4-6 includes the proportion of households in social housing recorded in the 1981 Census as an additional regressor. The 1991 proportion in social housing instruments neighbourhood education. The vast majority of social housing was built prior to 1981, so, if the historical supply of social housing to low price areas is influencing our results, we would expect the coefficient on the proportion of social housing in 1981 to be significant, and for the coefficient on educational or income status to fall. In fact, the coefficient on the 1981 proportion in social housing is insignificant, and the IV estimates are not significantly different from those in Row 1. This procedure also indicates that our results are not related to historical effects from the proximity of social housing on the supply of dwellings for owner occupation. We can also check that it is not the conversion of houses to lower cost flats close to social housing that drives our results, by excluding property transactions on flats from the sample. Doing this increases
the point estimates by about 10%. So, we infer that the relationship between local education levels and property prices is not attributable to unobserved variation across neighbourhoods in the supply of housing quality.

An important point to note here is that deterioration in the quality of housing occurring as a result of a negative externality from low human capital neighbourhoods – for example as poorer home owners move into the area – does not result in upward-biased estimates of the cost of the educational externality. Instead, if the supply of housing measured in quality units falls back in response to falling demand, then the estimated implicit price, conditional on housing characteristics, will under-estimate the impact of the externality. For social valuation we want to measure the derivative of prices with respect to local human capital, including the effect on price resulting from property deterioration.

We do more to check if it is something particular about social tenants, or properties near social housing which generates the observed education-price relationship by using an alternative instrument – the location of higher education institutions. This is a weak instrument (t-statistic of 1.67 in the prediction equations), and only 1.5% of the sample sectors have higher education institutions located within them. Nevertheless, the point estimate of the effect estimated in Row C is almost identical to the national average effect implied by the main results. The coefficient is not, however, significant at acceptable levels. Still, the coefficient indicates that we are picking up education related effects in our main results, rather than pure prejudice against social tenants.

To gauge the extent of the importance of other unobserved neighbour heterogeneity, we introduce more neighbour attributes in the regressions in Rows 4 and 5. Controlling for these characteristics brings down the estimated impact of local education and incomes, but the elasticities are now almost identical to those obtained using the semi-parametric spatial fixed effect models in the main tables, so it does not look like our initial estimates are wrong.
We would really like to know to what extent the effects of local education that we measure might be related to aversion to local crime rates. For the US, Lochner and Moretti (2001) estimate that the external benefits of education associated with crime reduction alone are between 14% and 26% of the private returns to education. Unfortunately, good local crime rate data is not available in the UK. Still, we can construct a crude measure of neighbourhood crime at Postcode-district level using the 1992 British Crime Survey, an individual level victimisation survey that includes 575 Postcode identifiers. The Postcode-district mean of property crimes in the last year recorded per respondent attracts a negative and significant coefficient (−0.014, s.e = .007) when entered on its own in a property price regression. However, the coefficient becomes near-zero and insignificant (−0.004, s.e. 0.006) once we control for Postcode-district education levels. Again, we must conclude that local educational status is the more important factor. We look deeper into the impact of urban crime on property prices in Chapter 6.

4.6.6 Evidence for human capital externalities

One prediction from a model where neighbourhood educational status generates an externality in the production of children’s human capital is that the implicit price of improvements in neighbourhood educational composition must be increasing in family size (treating family size as exogenous). We should find that the implicit price of educational status is higher in neighbourhoods with more children per household, or with a higher proportion of households with children. Indeed, this is what we do find. The implicit price of neighbourhood educational status is 38 percentage points higher in above-median family size sectors than in below-median family size sectors. The difference is significant ($t = 2.66$). Similarly, the proportion of home-owners with children is also increasing in the proportion of highly qualified residents – once this is instrumented by the proportion of social tenants in the neighbourhood to fix the
endogeneity. A reduction on the proportion of social tenants equivalent to a one percentage point rise in the proportion of residents with diplomas and degrees is associated with a 0.14% rise in the proportion of home-owners with children (from a mean of 30%). Our interpretation of this is that home-owners with children are willing to bid more for marginal improvements in the educational status of neighbourhoods because of the impact on the educational attainments of their children.

4.6.7 Other evidence

We have looked for more direct evidence that residents prefer more educated neighbourhoods. Burrows and Rhodes (1998) combine Survey of English Housing and 1991 Census data to model the geographical distribution of neighbourhood dissatisfaction in terms of the percentage of households in each Ward who say they are very dissatisfied with their neighbourhood. Regressing this Ward-level indicator of neighbourhood dissatisfaction on the Ward proportion with high qualifications suggests a 0.01% (s.e. = 0.0047%) decrease in the proportion expressing dissatisfaction as the proportion of highly qualified residents increases by 1% – a elasticity of 0.016 at the mean. This result is unchanged if we include other key Census variables – the proportions professional, unskilled, unemployed, non-white, lacking housing amenities, in social housing, in agricultural employment, plus household density and average property size. The Ward proportion with high qualifications is the only statistically significant coefficient (at the 5% level) in this regression⁶⁴. Admittedly, the magnitude of the effect is small using this data, but educational composition seems to be one of the stronger candidates amongst local factors for a contributor to residents’ self-reported perceptions of satisfaction with their neighbourhood.

⁶⁴ It is not one of the characteristics used to model the dissatisfaction variable
4.7 Concluding discussion

These results demonstrate that neighbourhood property prices respond to positively to the presence of more educated neighbours. Households value residence in 'educationally rich' neighbourhoods. With imperfect capital markets, this implies some degree of endogenous stratification into communities according to education levels, since more educated and wealthier home-owners will purchase housing services in more highly educated communities (and for other sufficient conditions see Benabou (1996)). For estimation purposes, we have relied on the exogenous stratification of communities generated by the juxtaposition of purpose built socially housed and owner-occupier communities in England and Wales. The estimated elasticities in the average neighbourhood are stable across regions, at around 0.24 for the proportion higher-educated. We get similar elasticities on Local Authority educational status using property-level micro-data and conditioning on home-purchaser incomes. Interestingly, the relationship shows no sharp discontinuities throughout its range, and is well approximated by a semi-log functional form.

Using additional Census, school, unemployment and crime data, we can see that the sensitivity of prices to local educational status is undiminished once we abstract from other observable characteristics of individuals in the neighbourhood. We conclude that households place particular importance on the educational status of a neighbourhood in choosing a residential location. That households with more children seem prepared to pay more, highlights the potential importance of community spillovers in the production of human capital.

These results have direct relevance to the cost-benefit analysis of measures to improve the educational status of deprived neighbourhoods, as well as to educational policy in general. Unfortunately, without detailed information on neighbourhood educational attainments other than higher education qualifications, we can make no
assessment of the extent to which higher education matters over and above education in general. On the assumption that it is the educational status of communities that matters to households, we can make a tentative assessment of the long-run social, community-level benefits of education and compare this with the private returns. Let us focus on households headed by someone under the age of forty. Mean household annual earnings for these households was around £19000, and the mean property price was £65000 in 1995. A 10% relative increase in the proportion of adults with higher education qualifications in 1995 would have meant a 1.9% absolute increase in the proportion with higher education qualifications. Assuming the private returns to higher education qualifications are in the order of 25% to 30%, this improvement in education implies an increase in mean earnings of around 0.53 percent, or £100 on average household income.

From the estimates presented in this Chapter, this change in educational attainments would be valued at £1500 at the 1995 mean property price. Average mortgage interest rates in 1995 were around 7% and the mean loan period 22.5 years, so £1500 is equivalent to £130 per annum in mortgage payments. By this calculation, the money-metric value of the non-pecuniary benefits to society from an individual gaining a higher qualification is higher than the mean return in terms of increased earnings. Of course, it is probably true that communities with more residents with higher education qualifications will have more education generally (A-Levels, for example) so the direct comparison between the increment to mean incomes and the increment to mean property prices from

---

65 Sources of the figures that follow are variously: Family Resources Survey 1995/6, Survey of English Housing 1995, Survey of Mortgage Lenders 1995. Age 40 is the 75th percentile in the age distribution of those taking out mortages in the survey of mortgage lenders. Returns to education control for gender, ability and family background – calculations from National Child Development Survey and 1970 British Cohort Survey, but see also Blundell, et al. (2000), or Harkness and Machin (1999). Mortgage interest rates in 1995 were 6.15% according to the 5% Survey of Mortgage Lenders, though the figure given in Building and Construction Statistics is 7.83% per annum.
a change in educational composition is inaccurate. Even adjusting for this, we find that
the increment to mean incomes from an increase in the proportion of educated residents
and the average value placed on that increase by others in the community are of a similar
order of magnitude\textsuperscript{66}.

If we believe that households value community educational status purely as an
input into children’s human capital accumulation, and that parents can transfer income
directly to children, it follows that the average household, which has one child, expects a
10% relative improvement in quality of the average neighbourhood to increase a child’s
expected household income by around 0.7%\textsuperscript{67}. We can take this is an upper bound to the
average impact of neighbourhood quality on a child’s future household income. If all
improvements in income are linked to better individual educational attainments, and
returns to education are expected to remain unchanged, then parents expect this 10%
relative change in neighbourhood status to improve their child’s chances of gaining
higher education qualifications by a similar proportional amount\textsuperscript{68}. This unit elasticity is
substantially higher than the child outcome-neighbourhood educational elasticities
estimated in the neighbourhood effects literature, which are in the order of 0.05-0.2 (e.g.
Kremer (1997), Aaronson (1998) or Gibbons (2001)). Clearly, not all the expected

\textsuperscript{66} We can redo the calculation assuming that the numbers in all post-school qualification groups in
a community vary in proportion to the number higher educated.

\textsuperscript{67} The mean number of children per owner-occupier household for owner-occupiers headed by
someone under 40 in the Survey of English Housing in 1995 is one. The calculation assumes that
expected child’s household income is the same as current mean household income, so we just
divide the value of the benefits (£130) by household income.

\textsuperscript{68} Because the change in the proportion of people with higher qualifications necessary to increase
earnings by 0.7% is roughly 2.3%, assuming the return to higher qualifications is around 0.3. The
current proportion with higher qualifications is tending towards 23%.
benefits of a better neighbourhood relate to better earnings-related outcomes for children, or else the existing literature severely underestimates the magnitude of these influences.

These back-of-an-envelope calculations are, of course, approximate. Still, the message comes across that incoming residents are prepared to pay to live in communities with higher stocks of human capital. It seems that the aggregate, non-earnings related community benefits per household are of a similar order the aggregate private returns per household as measured by the increment to earnings from higher education qualifications. It should be borne in mind also that the social benefits measured here are only those that accrue locally, so will not include spillovers in production, in workplace relations, in technological innovation and in other areas where action is at a broader geographical level. Given the size of these effects measured here, these community benefits warrant further analysis. Focussing on the private returns to education seriously understates the value of education to society, and any policy decisions based on these returns alone may result in sub-optimal provision of educational services.
4.8 Appendix A: Changes in educational composition

Figure 4-3 documents the relevant national changes. It shows that the proportion of social tenants with diplomas, degrees and other higher education qualifications increased from 2.7% in 1991 to 3.6% in 1994/5 - a growth of 33% - whilst the proportion of non-social tenancy groups with higher education rose from 17.5% to 22.4%. Overall there was a 28% growth in the proportion of those with high qualifications between 1991 and 1994/95.

Figure 4-3: Proportion with high qualifications (degrees, diplomas and teaching qualifications) 91-99

What are the implications of this change for estimates of the response of 1995 property prices to the spatial distribution of education levels as measured in 1991? If all neighbourhoods experienced 28% increase in the proportion with diplomas and degrees, then conversion of parameters measuring the response to 1991 education levels to measurements of the sensitivity to contemporaneous education levels is straightforward. The coefficients must be adjusted downwards by 28%, and the elasticities will be unchanged. However, changes in the distribution of education across neighbourhoods mean that estimated coefficients and elasticities will be downward biased relative to the parameters on contemporaneous education levels (the classical measurement error problem).
Unfortunately there is no information on changes in the ranking and distribution of educational composition between 1991 and 1995, but we can get a feel for the scale of the problem by comparing 1991 Ward level indices of deprivation and the DETR Indices of Deprivation 2000 (Department of Environment (2000)). Various composite indices are available from 1991 Census data – the Carstairs, Townsend, DoE, Jarman – though none is directly comparable to any of the 2000 indices. Nevertheless, the 2000 indices can explain up to 66% of the variance in the 1991 indices.

We assume we can write the local proportion of highly qualified persons in 1991 as a multiple of true local proportion in year \( t \), plus an uncorrelated error term that grows with time.

\[
z_{o,i} = \alpha_t z_{t,i} + \epsilon \cdot t
\]

A regression estimate of \( \beta \) in the model

\[
y_{t,i} = \beta \cdot z_{t,i} + \nu_t
\]

where \( z_{t,i} \) is proxied by \( z_{o,i} \) will give an inconsistent estimate of \( \beta \):

\[
\text{plim} \hat{\beta} = \frac{\beta}{\alpha_t} \left( 1 - \frac{\sigma_v^2}{\sigma_z^2} \right)
\]

The 2nd term inside the brackets could be estimated as the residual sum of squares from a regression of the 1991 measure on the contemporaneous measure, divided by the total sum of squares (or 1-\( R^2 \)). Assuming the relationship between local educational composition in each period is the same as the relationship between deprivation indices in 1991 and 2000, we calculate that \( \sigma_v^2 / \sigma_z^2 \) is 0.34%. The attenuation on regression coefficients resulting from our use of 1991 data as a proxy for 1995 neighbourhood composition is then about 8.5% (=1-25x0.0034). General growth in the proportion of qualified residents leads to an upward bias of around 28%. Under these assumptions, we need to adjust \( \hat{\beta} \) downwards by around 20% if we wish to interpret is as the contemporaneous of \( y_{t,i} \) to \( z_{t,i} \), whilst the elasticities should be increased by over 8%. However, we should bear in mind that a high proportion of property completions in 1995 were initiated in 1994, and may be based on the decisions made some years earlier.
### 4.9 Appendix B: Neighbourhood incomes

#### Table 4-7: Property price response to neighbourhood incomes, 1996 and 1999

<table>
<thead>
<tr>
<th></th>
<th>SE&amp;E</th>
<th>IV-SSE</th>
<th>The North</th>
<th>SSE</th>
<th>IV-SSE</th>
<th>SW&amp;W</th>
<th>IV-SSE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Log mean Postcode-sector incomes</td>
<td>0.465</td>
<td>0.487</td>
<td>0.742</td>
<td>0.531</td>
<td>0.603</td>
<td>0.556</td>
<td></td>
</tr>
<tr>
<td>Density of purpose built flats (100s/km²)</td>
<td>0.2 e⁻³</td>
<td>0.2 e⁻³</td>
<td>1.8 e⁻³</td>
<td>0.0 e⁻³</td>
<td>5.6 e⁻³</td>
<td>6.4 e⁻³</td>
<td></td>
</tr>
<tr>
<td>Proportion Black, Indian, P'stan, B'desh</td>
<td>-1.174</td>
<td>-1.178</td>
<td>-0.659</td>
<td>-0.858</td>
<td>-0.817</td>
<td>-0.813</td>
<td></td>
</tr>
<tr>
<td>Mean rooms in owner-occupied housing</td>
<td>0.172</td>
<td>0.167</td>
<td>0.096</td>
<td>0.152</td>
<td>0.148</td>
<td>0.169</td>
<td></td>
</tr>
<tr>
<td>Detached</td>
<td>0.473</td>
<td>0.473</td>
<td>0.506</td>
<td>0.507</td>
<td>0.473</td>
<td>0.473</td>
<td></td>
</tr>
<tr>
<td>Flat/Maisonette</td>
<td>-0.586</td>
<td>-0.585</td>
<td>-0.405</td>
<td>-0.404</td>
<td>-0.474</td>
<td>-0.474</td>
<td></td>
</tr>
<tr>
<td>Terraced</td>
<td>-0.186</td>
<td>-0.186</td>
<td>-0.272</td>
<td>-0.272</td>
<td>-0.198</td>
<td>-0.198</td>
<td></td>
</tr>
<tr>
<td>Overall R²</td>
<td>0.940</td>
<td>0.940</td>
<td>0.898</td>
<td>0.898</td>
<td>0.913</td>
<td>0.913</td>
<td></td>
</tr>
<tr>
<td>Within R²</td>
<td>0.853</td>
<td>0.853</td>
<td>0.820</td>
<td>0.812</td>
<td>0.839</td>
<td>0.840</td>
<td></td>
</tr>
<tr>
<td>P-value test of restrictions</td>
<td></td>
<td>0.336</td>
<td></td>
<td>0.152</td>
<td></td>
<td>0.207</td>
<td></td>
</tr>
<tr>
<td>Sample size</td>
<td>19445</td>
<td>16241</td>
<td></td>
<td>12486</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean log-incomes</td>
<td>3.164</td>
<td>2.941</td>
<td></td>
<td>2.946</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean log-price</td>
<td>11.40</td>
<td>10.85</td>
<td></td>
<td>11.00</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean price</td>
<td>£108797</td>
<td>£58334</td>
<td></td>
<td>£67453</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean bandwidth</td>
<td>1.30 km</td>
<td>1.50 km</td>
<td></td>
<td>1.80 km</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Dependent variable is log of Postcode-sector mean property-type price

Instruments in Columns 4 & 6 are Postcode-sector proportion in social housing and proportion of social housing tenants from Indian, Pakistani, and Bangladeshi ethnic groups
4.10 Appendix C: Estimating non-linearities

In order to check for non-linearities in the response of log prices to educational status, we must generalise our smooth spatial effects approach, to estimate:

\[
\ln P_{ir} = g(x_i, c_{i1}, c_{i2}, b_i) + \gamma \tilde{W}_{ir} + u_{ir} \tag{4-19}
\]

Visual representation of the relationship between \( \ln P \) and \( x_i \) is infeasible in the general case. Instead, a function \( h(x) \) can be estimated as the average relationship between expected prices and \( x_i \), averaging over the distribution of spatial co-ordinates, i.e.

\[
h(x_i, b) = \int \int g(x_i, c_{i1}, c_{i2}, b) f(c_{i1}, c_{i2}) dc_{i1} dc_{i2} \tag{4-20}
\]

where \( C \) is the support of \( (c_{i1}, c_{i2}) \) in the sample, and \( f(c_{i1}, c_{i2}) \) is their joint density. The computational procedure is as follows:

1) Estimate the linear coefficients \( \gamma \) in the model (4-21) below, replacing \( m(\cdot) \) with a 3-regressor kernel regression estimates with observation-dependent bandwidths \( b_i \) for \( c_{i1}, c_{i2} \), and a fixed bandwidth for \( x \):

\[
[\ln P_{ir} - m(\ln P_{ir} | e, x, b)] = \gamma' (\tilde{W}_{ir} - m(\tilde{W}_{ir} | e, x, b)) + \sigma_{ir} \tag{4-21}
\]

2) Estimate \( \hat{h}_m(x) \) by kernel regression of \( [\ln P_{ir} - \tilde{y}\tilde{W}_{ir}] \) on \( x, c_{i1}, c_{i2} \) at a number of grid points \( \hat{x}_g \), all at a fixed co-ordinate-pair \( (\hat{c}_{im}, \hat{c}_{jm}) \). This co-ordinate pair is drawn at random from the sample.

3) Re-calculate \( \hat{h}_m(x) \) at M different co-ordinate pairs drawn at random from the sample.

4) Calculate \( \hat{h}(x) = \frac{1}{M} \sum_{m=1}^{M} \hat{h}_m(x) \), that is the average of the M kernel regressions over the sub-sample of randomly chosen, within-sample, spatial locations at which the estimated joint density \( f(x, c_{i1}, c_{i2}) \neq 0 \).

Since in this application, educational composition \( x \) is treated as endogenous, it is replaced by the generated regressor:

\[
\hat{x}_i = \tilde{x}(z_{i1}, c_{i1}, c_{i2}, b_i) + \tilde{y}\tilde{W}_{ir} + \hat{h}_e \tag{4-22}
\]
where $Z$ is a suitable instrument, $\hat{x}(z, c_u, c_{2i}, b_i)$ is estimated by kernel regression, and the other parameters are estimated using the partial linear model described above. If $\hat{h}(\hat{x})$ is to be a good representation of $h(x)$, we require that the instrument is continuous, is a strong predictor of $X$, and that $\hat{x}$ has a similar support to $X$. As discussed in Section 4.5.3, the proportion of social tenants is the main instrument, and this satisfies these requirements. Rilstone (1996) discusses the use of generated regressors in non-parametric estimators, and their asymptotic properties.
5 Valuing Primary Schools

5.1 Introduction

Severe inequalities in the measured performance of English primary schools across geographical space have parents clamouring to get their children into the best schools. There is a lot of anecdotal evidence to suggest that parents are prepared to move house to try to secure admission to a good school, and that they are often prepared to pay a high premium on property prices. Stories of soaring house prices close to good schools are commonplace. We have heard stories from Local Education Authority staff of complaints and appeals by families failing to gain admission to a school of their choice, despite having moved house specifically for that purpose. One anonymous interviewee spoke of an expectant mother calling for advice on which streets she should consider moving to in anticipation of her unborn child’s primary education. Another family, known to one of the authors, recently sold a three bedroom Victorian terrace in north London for a much smaller semi-detached house just over a mile away. This move cost them around £140000. For what net gain? A 35% increase in the proportion of children at the local primary school reaching the target level in age-11 assessment tests. For sure, these moves may buy more than just better schools – good schools are typically in neighbourhoods that are better in other ways: lower crime rates, quieter neighbours, cleaner streets, better local amenities. But some component of any premium paid for a re-location from a bad-school neighbourhood to a good-school neighbourhood may well be attributable to the price of an improvement in school quality.

This phenomenon is by now widely recognised in the US, and many attempts have been made to quantify it. The US literature is rich in efforts to evaluate school performance, or school characteristics, through house price models. But school
characteristics are, in part, determined by neighbourhood socio-economic composition and hence by house prices. Also, in the US, school funding is determined by the local property tax base. Recent work uses a variety of strategies to try to correct for biases in property value models induced by this endogeneity. Researchers in the US compare properties in similar neighbourhoods on either side of school district boundaries (Bogart and Cromwell (1997), Black (1999)), use repeat sales when school district boundaries are re-drawn (Bogart and Cromwell (2000)), or include an extensive set of neighbourhood characteristics as controls (Downes and Zabel (1997)).

By contrast, the literature for the United Kingdom is very thin, despite being discussed a great deal in the media, and amongst politicians and parents. It is not feasible to construct data-intensive estimators of the type favoured in the US, because geocoded data on individual property transactions is scarce, and because school admissions districts are not rigidly enforced. Instead, in this paper we use strategies that compare average school performance and house prices in proximate neighbourhoods and employ exogenous, permanent school characteristics as instruments. Our kernel-based technique for removing spatial fixed effects is based on that in Chapter 4. We estimate the premium attracted by improvements in primary school quality in England, using property price data from the Land Registry joined to the Department of Education and Employment (DfEE)\(^{69}\) school performance tables. This sample gives us near-universal coverage of property transactions and school performance measures in England from 1996 to 1999.

This is the first research to value primary school performance in England. Others have looked at the value house buyers attach to secondary schools in England. Rosenthal (2000) finds rather low elasticities of house prices with respect to school performance,

---

\(^{69}\) Since the last government election DfEE has become DfES (the Department of Education and Skills).
comparable with the US work of Judd and Watts (1981). Cheshire and Sheppard (1995) estimate the value of location within specific school catchment areas in Reading and Darlington. Leech and Campos (2001) do the same for Coventry. Neither study relates value to school performance measures. None of these studies look at primary school performance. We think this may be more important generally: hedonic methods relying on spatial associations to link properties to schools may be poor for evaluating secondary school performance, except in special cases where catchment area boundaries are well defined and exclusive. Teenagers are a fairly mobile group and can travel long distances to school. Mobility between Local Education Authorities is high for children in secondary education. In contrast, primary age children typically attend schools which are within walking distance (at least in urban areas), and catchment areas can shrink down to just a few blocks for those in the highest demand.

Our interest in primary schools also has a sound empirical and theoretical basis. We would expect primary school performance to be the principal object of choice by parents seeking to improve the life chances of their offspring. For a start, there is evidence that attainments in the early years are positively correlated with later academic and economic success (Feinstein and Symons (1999), Feinstein (2000), Gregg and Machin (2000)). If gains made in the primary years reap rewards in terms of achievements at secondary school, then the payoff for the investment is higher if the investment is made early on in a child’s life. What is more, investment in good primary education may be a pre-requisite of admission to selective secondary schools. Given the

70 Haurin and Brasington (1996) however, find a 0.52% increase in price for a 1% increase in the proportion of 9th graders passing all sections of the 1990 proficiency test in Ohio.

71 Their later report (Cheshire, et al. (1999)) prices GCSE A-C pass rates at £343 per 1% improvement in Reading, £50 in Darlington, £57 in Nottingham, but only the coefficient for Reading is statistically significant. Even more recently (Cheshire and Sheppard (2002)), they find evidence of secondary and primary school effects for the Reading area.
high fixed costs of moving house, a rational parent will make a once and for all locational choice when their first child enters the education system.

The Survey of English Housing provides some evidence that neighbourhood is an object of choice by homeowners, particularly those with children. Unfortunately, schooling is not one of the reasons on the list of responses available to respondents. However, in the 1997/8 wave of the survey, 29% of 2674 owner-occupier respondents with children gave “to move to a better neighbourhood” as their first reason for their last move. For those without children, the figure was 18%. Whilst this only weak evidence that education is the main concern in a family’s choice of neighbourhood, schools must surely score highly in any list of neighbourhood attributes ranked by desirability to families with children.

Our focus in this Chapter is therefore upon the associations between local house prices and primary school performance using highly disaggregated price and school data. The rest of the paper proceeds as follows. In Section 5.2 we discuss the usefulness of the hedonic framework in the current context, and the extent to which location matters for admission to primary schools in England. Section 5.3 describes our econometric approach. Section 5.4 then moves on to discuss the data we use and, as the data comes from several sources, the matching procedures we adopt. Section 5.5 presents econometric estimates of our house price models. Section 5.6 concludes.

\[ \chi^2 (1) = 44.0. \]
5.2 English primary schools and the housing market

5.2.1 The hedonic approach

We use a standard hedonic property value framework to assess the implicit price of school performance. This framework has been employed frequently in the environmental, land and urban economics literatures to price local environmental amenities (see Rosen (1974), for the classic exposition, or Sheppard (1999) for a modern survey).

Some determinants of school quality may be exogenous to local community characteristics – physical and institutional differences, possibly teacher skills and head leadership styles. But a substantial component of school performance is determined by the education, earnings and other characteristics of parents in the local community. This potential endogeneity presents a problem for empirical analysis. A causal link from school inputs to pupil achievement and a causal link from local family incomes to pupil achievement are observationally equivalent in terms of data on local incomes or house prices and pupil test success rates. Careful discussion, and model implementation based upon this, makes the identification of the demand for school characteristics an important part of our analysis.

We can start the modelling discussion with a simple stripped down model which assumes that school quality affects children’s early educational outcomes, and that variation in school performance is not generated purely by spurious spatial clustering of children of similar innate abilities or clustering of families with similar resources. Early educational attainments lead, in turn, to higher final educational achievements or reflect directly enhanced skills that lead to greater successes in adult life.74 For simplicity,

74 Many studies show this to be the case. For example, Gregg and Machin (2000) show early test scores to be important determinants of educational success and that they have a link with adult labour market outcomes over and above educational qualifications. Dearden, et al. (2002) show the same over and above school characteristics.
assume that we can characterise the effects of primary school quality in terms of an impact on adult expected earnings or lifetime wealth. As such, primary school performance is a desirable commodity for parents, either because they are purely altruistic, or because they expect some form of payback from their children in their later years.

To formalise this in a simple model, we write the preferences of owner-occupier households as a function of household consumption, average lifetime income of own children or young dependants, and characteristics of the property and its location. By lifetime income we mean the expected present value of lifetime income. This means we specify the following utility function:

\[ U = U(c, \bar{y^c}, q, l) \]  

(5-1)

where \( c \) is a numeraire composite consumption commodity, \( \bar{y^c} \) is lifetime income of own children (the bar denoting an average), \( q \) is a vector of structural housing characteristics, \( l \) is a vector of locational characteristics. One can characterise the process of generation of lifetime income in terms of an educational production function, with state-sector school quality and a vector of other inputs. Of course, parents always have the option to transfer wealth directly to their children rather than allocate their resources to the inputs in the educational production function. This results in a production function of the form:

\[ y^c = f(x, z) + y' + y^s \]  

(5-2)

where child income is related to local school quality \( x \) and other inputs into the human capital production function \( z \), together with direct transfers from parents \( (y') \) and the contribution to the child’s household income by any future spouse \( (y^s) \).

House prices are determined as a function of the same attributes, where the attributes are traded at a set of exogenous prices \( \theta \) fixed by demand and supply equilibrium at a broader geographical level:
\begin{equation}
P_h = P_h(x, q, l; \theta)
\end{equation}

So, the household lifetime budget constraint is:
\begin{equation}
y^h = c + P_h(x, q, l; \theta) + P_x(z, k) + ky'
\end{equation}

where \( k \) is the number of children in the household and \( P_x(z, k) \) is other expenditure on their human capital (e.g. on private education). Assuming the choice space is continuous so that households can purchase their optimum bundle we have the first order condition for \( x \):
\begin{equation}
\frac{\partial U}{\partial c} \cdot \frac{\partial f}{\partial x} = \frac{\partial P_h}{\partial x}
\end{equation}

Note that the option of transferring wealth directly to children implies the following relationship between the implicit price of school quality and the returns to school quality in terms of an expected child's lifetime income.
\begin{equation}
\frac{\partial \ln \bar{y}^c}{\partial x} = \frac{P_h}{k \cdot \bar{y}^c} \cdot \frac{\partial \ln P_h}{\partial x}
\end{equation}

This is the case where school performance is valued purely as an input into the production of children's lifetime wealth. Marginal willingness of parents to pay for better schools will be the marginal effect of the school performance measure on the present value of their children's total future earnings.

An estimate of the implicit price of school productivity is available from a simple regression of some transformation of property prices on a measure of local school performance measures. However, the Ordinary Least Squares will only provide consistent estimates of the partial derivatives in (5-5) if admission to schools is linked directly to school property locations, and if school performance is otherwise unrelated to the housing market. As discussed above, the second assumption is implausible, because of family background effects on educational attainments. We discuss our solutions to this problem in Section 5.3. The first assumption also requires some discussion in the English context.
because schools do not operate well-defined and exclusive catchment areas. We discuss this next.

5.2.2 Does location matter for primary school admission?

Any attempt at valuation of schooling using the hedonic technique requires some method of linking property prices in a given area to the performance of schools available to residents in those properties. In our analysis there is an implicit assumption that geographical proximity to a school is an important criterion for admission. Whilst geographical proximity is one criterion, it is certainly not the only one. In England, Local Education Authorities (LEAs) are the administrative bodies responsible for the provision of state schooling. There are around 174 LEAs in England. Local Education Authorities are largely funded by central government grants, so there is no link between local taxation and school funding. But these LEAs do operate their own systems of prioritising applications to a primary school. Legal precedent (the Rotherham Judgement, 1997) has determined that parental preference must be the LEA’s first consideration. However, good primary schools are usually oversubscribed, so the admissions authority must employ some system for ranking applications in order of priority. Typically, for LEA administered schools, priority is assigned according to the following over-subscription criteria:

i) those with siblings at the school;

ii) those with special educational or medical needs;

iii) those resident in a local “catchment” or “neighbourhood area”;

iv) children of those employed in the school;

v) those ranked first by other geographical criteria (e.g. walking distance to the school).

The exact details and order vary across Local Education Authorities. For religious schools, some statement or evidence of religious affiliation is usually the first criterion to
be met. Even then, parents must attend the local church regularly, or the school must be the nearest of the same denomination for children to be eligible to attend. Although catchment area boundary data might be helpful where catchment areas are well defined, we suggest that close proximity to a primary school is a reasonable proxy for meeting the geographical criteria for admission\textsuperscript{75}.

What is clear is that choosing a location within the LEA and close to a school will maximise the chances of school admission for a family moving house for this purpose, whatever other criteria have been met. It will also minimise the costs of delivering children to school. But a more specific statement of our underlying assumption about the link between residential location and school access might be helpful at this stage. We maintain that location of residence defines the probabilities of admission to a number of schools, depending on the residence-to-school distance. Expected school quality at a specific residential location is thus an admission-probability weighted average of the performance of local schools.

Since we are interested in the price premium generated by those actively seeking school quality, and since catchment areas are non-exclusive, we argue that the relationship between mean neighbourhood property prices and mean neighbourhood school performance will provide just as much information as data based on individual schools and catchment areas. We use the association between property prices and primary school performance averaged at the Postcode sector level (a spatial unit of around 2500 households). We also test this assumption by comparison with fixed effect models that rely on property price and school performance differences between adjacent Postcode sectors separated by Local Education Authority boundaries.

\textsuperscript{75} The London Borough of Hackney publishes information on the maximum distance of residence for successful applicants in the previous year. The median distance in 1999/2000 amongst 27 schools was 580m. Weighting by the difference between applications and intake gives a demand-adjusted median of 450m.
5.3 The Empirical Approach

5.3.1 Empirical model

We follow a similar approach to that used to value neighbourhood human capital in Chapter 4. But this time we have some time-series variation in the data. Again we use a Government Land Registry property price data set in which transactions are aggregated to provide an average of prices in four property-type categories – flat/maisonette, detached, semi-detached, or terraced – at Postcode sector level. The Postcode sector is a geographical area containing an average of 2500 households. We adopt these Postcode-sector-property-type means as our unit of empirical analysis. For the hedonic price function, we specify a semi-log functional form, with an unknown function \( g(c_i, t) \) mapping geographical location \( c_i \) to house prices in each time period \( t \). This imposes the constraint of a constant percentage response in house prices to a one percentage point absolute increase in school performance. Our specification of the log-price of a house of type \( r \) in neighbourhood \( i \) at time \( t \) is then:

\[
\ln P_{irt} = \beta x_{it} + \gamma z_{irt} + g(c_i, t) + u_{irt} \tag{5-7}
\]

\( P_{irt} \) is the mean property price of property type \( r \) in Postcode sector \( i \) at time \( t \). The vector \( z_{irt} \) includes exogenous property and neighbourhood characteristics. In our case, this will mean property type dummies, the Postcode sector proportion in social housing, and the mean number of rooms occupied by home-owning households\(^7\). These can be treated as exogenous under the assumption of a fixed housing supply, which underlies hedonic modelling. Our key school performance variable, \( x_{it} \), is the effectiveness of the school in producing educated children. Specifically, we assume this is the probability that

\(^7\) We define location is defined by the co-ordinate pair \( c_i \)

\(^7\) From the 1991 Census
a child sent to the school reaches the age-11 target level in national assessment tests. These *Key Stage 2* tests are standard across all schools in the state sector. This success probability is unobserved, but we will proxy it by the proportion of children reaching this grade, as published in the DfEE school league tables. This is exactly what parents do when they inspect these tables as the basis for school choice.

### 5.3.2 Estimation strategy

Estimation of a full structural specification of the mapping of location to house prices requires data on local amenities, local housing characteristics, the proximity of neighbourhoods to transport services, local labour demand, environmental quality and other unknown local goods. The general function $g(c_i,t)$ could then be replaced by a specific function of known variables. The difficulty of this approach is knowing exactly what to include in the hedonic price function. Researchers often include a large set of property characteristics in the regression, alongside an ad-hoc selection of socio-economic characteristics to proxy unobserved local characteristics. This is a poor strategy, since neighbourhood socio-economic composition is determined by sorting processes that are driven to a large extent by the housing market. So, local socio-economic composition is endogenous in a property value model. As a result, parameter estimates obtained by OLS regressions may be biased. What we do here instead is allow for general effects of location on property prices using non-parametric estimates of the impact of location on our model variables – in effect abstracting from the unobserved area-specific effects on prices, $g(c_i,t)$.

A simple and approximate solution would be to include geographical area dummy variables to proxy unobserved area effects[^78]. A drawback of this area fixed-effect approach is its reliance on an arbitrary specification of the comparison neighbourhood

[^78]: We present results based on this strategy for comparison purposes in Table 5-3.
In our case we could assume Postcode \textit{district} fixed effects since Postcode sectors are nested within Postcode districts. But there is no theoretical basis for believing that Postcode districts are the appropriate controls, and using noisy measures of area fixed effects will lead to inconsistent estimates of the model parameters.

What we do instead is compute spatially weighted means of the variables in our model at each observation in the data, in which the nearest observations receive the highest weights. These averages capture general, unobserved, area and amenity impacts on the housing market, \textit{centred at the location of the unit of observation}. We then transform the data into deviations from these spatially weighted, means, and use the transformed variables in our regressions. An extension on the model of Chapter 4 is that we have multiple time periods, so we can estimate a separate non-parametric surface for each period\(^{79}\). But we still need to specify how rapidly the weights in our spatial averages decay as we move away in space from a given observation at location. We use a weighting function \( m(c_i, b) \), with a bandwidth \( b \) that specifies how rapidly the weights decay with distance. Expressing the model in deviations from these estimated spatial averages, the regression model becomes:

\[
\ln P_{irt} - m(\ln P_{irt} \mid c, b) = \beta [x_{it} - m(x_{it} \mid c, b)] + \gamma [z_{irt} - m(z_{irt} \mid c, b)] + \epsilon_{irt} \tag{5-8}
\]

Parameters \( \beta, \gamma \) and their variance covariance matrix can then be estimated by applying Ordinary Least Squares to the transformed variables. This smooth spatial effects (SSE) estimator is an application of the semi-parametric partial linear model (see, for example Robinson (1988), Hardle (1990), Stock (1991))\(^{80}\).

\(^{79}\) In practice we have: \( g(c_{i,t}) = \sum d_i \cdot g_i(c_{i_t}) \) where \( d_i \) is a time dummy. This allows for differential growth in house prices across geographical space.

\(^{80}\) The spatially weighted mean is estimated by the Nadaraya-Watson method. See Chapter 4 p.167
5.3.3 Bandwidth choice

The choice of bandwidth is important because we have no way of knowing, other
than by casual empiricism, what geographical area comprises the correct reference group. We therefore experimented with a number of choices of \( b \). A bandwidth of near zero is equivalent to a fixed effects estimator with Postcode sector fixed effects. In this case, the relationship between school performance and house prices would be identified by changes over time alone. This is unsatisfactory since the most sought-after schools tend to near the top of the 0-100% performance range and show low performance growth relative to others. Moreover, this removes nearly all of the useful cross-sectional variation in prices and performance. As we discuss later, short run changes in performance will have little impact on the housing market, so we need retain some cross-sectional variation by including neighbouring Postcode sectors with non-zero weights. However, the land area and household density of Postcode sectors is far from constant in our sample. Postcode sectors in rural locations are much larger than in urban locations, reflecting lower population densities in rural locations. To compensate for this, we vary the bandwidths in inverse proportion to the square root of the local household density as recorded in the 1991 Census. Our main results use a bandwidth corresponding to approximately 3400 households. This bandwidth choice process is discussed in more detail in Appendix A.

5.3.4 Further identification through Instrumental Variables

Our SSE estimation strategy based on deviations from spatially weighted means abstracts from broader area impacts on the housing market, and reduces the potential problem induced by the simultaneity of property prices, local socioeconomic characteristics and school performance. But there are still some issues we need to address. Highly localised, unobserved differences between Postcode sectors may persist and induce simultaneity between property prices and performance.

First we might identify the causal structure by simply assuming a recursive system:
where \( \bar{y}_t \) is a measure of average local housing expenditures, \( \bar{x}_t \) is local school performance, and the variables are in deviations from spatial group means. This model structure implies that this year’s house prices respond to last year’s school performance, but that last year’s performance is unaffected by current house prices. Ordinary least squares applied to the deviations of the variables from local spatial group means gives consistent estimates of \( \beta \) only if the within-group transformation removes all correlation between \( \bar{e}_t \) and \( \bar{\omega}_t \), and serial correlation in \( \bar{e}_t \). Otherwise, estimates of \( \beta \) will be a variance-weighted average of \( \beta \) and \( 1/\gamma \).

Previous studies implicitly assume that controlling for property and neighbourhood characteristics is sufficient to ensure this condition is met, even without our smooth-spatial-effects approach to removing nuisance spatial variation.

Even if these conditions are met, the model implies that parents move house on the basis of a single-year measure of school performance. This is unlikely. The fixed costs associated with housing transactions and family relocations make moves each year in response to league table results highly inefficient. Instead, parents look to longer run indicators of school productivity. They may seek further information from school visits, school inspection reports, teaching staff, and by talking to other parents. Results published in the national tables are noisy measures of long-run school quality, and parents are more likely to seek out schools with proven track records of high performance, or

\[
\hat{\beta} = \beta + \left( \gamma^{-1} - \beta \right) \left( \frac{\sigma_e^2}{\gamma^{-2} \sigma_\omega^2 + \sigma_e^2} \right)
\]

\[81\] The inconsistency is: \( \operatorname{plim} \hat{\beta} = \beta + \left( \gamma^{-1} - \beta \right) \left( \frac{\sigma_e^2}{\gamma^{-2} \sigma_\omega^2 + \sigma_e^2} \right) \]

[82] With the exception of Rosenthal (2000), who also uses an IV approach.
those which exhibit characteristics that are, on average, associated with good long run performance. On this basis, least squares estimation of \( \beta \) in (5-9) will lead to a downward biased estimate of the value of persistent differences in school performance. This problem can be only partly addressed by using time averages of prior school performance, because we only have a short panel \(^3\). So we have two problems: 1) the simultaneity of prices and school performance when there is highly localised variation in unobserved factors affecting prices 2) the downward bias induced by using single-year or short run averages of school performance. But, as usual, consistent estimates can be obtained under both conditions using an Instrumental Variables approach.

This of course assumes we can find suitable instruments. We need characteristics of schools that influence performance but are unaffected by local property prices or neighbourhood socio-economic status. For this, we draw on school characteristics available in the school performance tables.\(^4\) We use historically determined school-type characteristics as instruments. Primary schools in England fall into three main categories: Community, Voluntary Aided or Controlled. Voluntary schools are almost always church schools (mainly Catholic and Church of England). Community schools are distinct in that they are non-religious and that the Local Education Authority employs the staff and administers admissions procedures. Primary school age ranges vary, in that some take pupils before the compulsory school age of 5 years. Others are “Middle” or “Junior”

---

\(^3\) This is simply an application of the classical measurement error model. If permanent school performance is functionally dependent on observable neighbourhood characteristics, such as the proportion of social tenants in the catchment area, then inclusion of this neighbourhood characteristic in the OLS regression will further downward bias the estimate of \( \beta \).

\(^4\) We tried qualified teacher-pupil ratios as an instrument at school level, but the underlying relationship between school performance does not work in the direction that we, and we assume parents, would expect. This suggests that more teachers are assigned to bad schools or
schools and take pupils at age 8 or 9 who have attended separate first-stage primary schools. School performance measured in age-11 tests varies across all these groups, even though there is no selection on the basis of academic aptitude. We assume then that these differences reflect organisational, teaching quality or ‘ethos’ differences that impact on success rates in the national age-11 tests. Indeed, the better performance of church schools is widely recognised by parents. Differences between school age range categories may arise for a number of reasons: there may be benefits from continuity in education between the nursery, infants and junior stages; children in neighbourhoods with nursery units may benefit from earlier introduction to school life; and schools accommodating a wider age-range will be larger, offering potential economies of scale. For all these reasons we believe that age-range and Community status provide good predictors of school performance. This is borne out in the first stage regressions we present (in Appendix C). These are fairly permanent characteristics so are good indicators of expected long-run performance. We also argue that age-range and Community status are unaffected by local spatial variation in current house prices, or by related variation in incomes or other socio-economic status. Indeed, only around 2% of primary schools in the sample opened or changed status during the previous 10 years, so it is unlikely that current house prices have much of an influence on these characteristics. Factors at a broad geographical level may have an impact – Local Educational Authority policy, for example. So, it is important that we incorporate the instruments within our smooth-spatial-effects estimation framework. We refer to this as an Instrumental Variables estimator with smooth spatial effects (IV-SSE).

disadvantaged areas, or that classes are smaller in schools that are less in demand. Either case invalidates its use as an instrument.
5.3.5 Identification from differencing across Local Authority boundaries

If there existed well-defined catchment area boundaries in England, we might do better in assigning property prices to schools. Without this information, our estimates on the price-performance response based on matching mean Postcode sector prices to mean Postcode sector school performance may well be lower bounds, due to the classical errors-in-variables problem induced by the fact that mean school performance in a Postcode sector is a noisy measure of the mean school performance of the schools available to residents of that Postcode sector.

However, we can look at a subset of Postcodes for which we can infer catchment area boundaries. We do this on the assumption that any Local Education Authority (LEA) boundary is also a primary school catchment area boundary. We make this assumption on the grounds that none of the LEAs we contacted drew their catchment or neighbourhood area boundaries to cross LEA boundaries (even though applicants from outside the LEA are not legally excludable). This approach is similar to that taken in Black’s Boston study (Black (1999)) and Leech and Campos’s study of secondary schools in Coventry (Leech and Campos (2001)), though these authors use detailed information on catchment area boundaries and property level data for a single, small geographic area.

We use a London sample of Postcode sectors that each share a Local Education Authority boundary with at least one other. The empirical model is as in equation (1), but the function $g(c_{i,t})$ is replaced by a dummy variable set. These dummy variables indicate pairs of Postcode sectors that are adjacent, but on either side of an LEA boundary, plus LEA dummies and time dummies. This is a difference-in-difference type

---

85 One solution to improving the match between schools and property prices is to average individual school performance within a given radius of the centroid of each Postcode sector. Our initial estimates based on this approach were similar to those obtained by simple Postcode sector matching. However, this procedure introduces an additional bandwidth selection problem, so was abandoned.

- 215 -
estimator. We exploit differences in performance and prices *within* pairs of Postcode sectors that straddle an LEA border, allowing for general differences between LEAs and general differences over time. This assumes similarity in the unobserved attributes of immediately adjacent Postcode sectors. Adjacent Postcode sectors that adjoin LEA boundaries, but are separated by some major physical obstacle are excluded, because the assumption that they form homogenous neighbourhoods is likely to be violated.  

5.4 The Data Set

5.4.1 Data description

The main data set we use comes from four sources, which we splice together at Postcode sector level. We have house price data from the Government Land Registry. Our primary school performance data comes from the public primary school performance tables, available from the Department of Education and Employment. Additional data on the proportion in social housing, average house size, household density and Postcode sector grid references comes from the 1991 Census for England and Wales. Fuller details are available in Appendix B. We end up with an unbalanced panel with up to four property types and property prices in each Postcode sector in each year. Household density, grid-references and the proportion in social housing vary across Postcode sectors but are constant across years in our data.

---

86 This includes, for example, all Postcode sectors separated by the Thames downstream of Richmond. LEA dummies remove differences in local council tax, housing and education policy. Unitary Authorities responsible for other aspects of local government are geographically coincident with LEAs in the London area.
In all, 7444 Postcode sectors and 2060 Postcode districts are represented in our matched house-price and primary-school sample for the years 1996 to 1999. The mean number of households is 2900 per sector, and 12900 per district. There are primary schools in 5681 sectors and 1888 districts. Postcode sectors for which we have house prices, but no school information (because there is no school present here, or because there of no successful match between house-prices and schools) are assigned zero Key Stage 2 results. We include a dummy variable to indicate these in our regressions.

Figure 1 illustrates the geographical relationship between Postcode sectors, districts and primary schools. It shows one Postcode district – E3 in the East End of London. This district, being an inner city area, has a higher density of housing and primary schools than average, but it illustrates the main features used in the analysis. The housing density in Postcode sector E3 4 is 6000/km², so a bandwidth choice of 3400 households in our smooth spatial fixed effect estimator corresponds to a radius of 0.42km. Very little weight is attached to sectors beyond 2.5 bandwidths, so the spatial group for a given Postcode sector, assuming a bandwidth of 3400 households is, roughly speaking, those Postcode sectors whose centre is captured within a 1 km radius from the centre of the observation Postcode sector. Each grid represents 0.5 km on this map. The symbols in Figure 1 represent the school types. Black circles are Community schools covering all primary years. The grey circle is a Community school that takes children from compulsory school age (5 years) only. White circles are schools that take older children only, or have separate organisational units for older children. Black triangles are Voluntary Aided (Church of England and Catholic) primary schools. In this example of an inner city Postcode district we can see a considerable variety of school types and age range within quite localised areas.
In addition to this Postcode sector data, we have obtained a property price data set for the London area from a property valuation firm\(^8\). This gives property level information on sales prices and property characteristics for an area covering around

\(^8\) Ekins, the valuation arm of Woolwich plc
800km$^2$ of the London region. We use this data set so that we can compare our main results with those obtained using more traditional methods.

### 5.4.2 Descriptive statistics

We present our results separately for three broad geographical areas. These areas correspond to grouped Standard Statistical Regions. The grouping scheme was chosen to illustrate any broad regional differences in property markets, whilst retaining a mix of rural, urban and metropolitan areas within each area. Theses groupings are the same as in Chapter 4, but without Wales, for which we have no school data. The upper panel of Table 5-1 reports summary statistics on our Postcode sector property price data set.

House price growth from 1998 to 1999 appears lower than might be expected, considering the recent media attention on soaring house prices in the South East of England. The figures show a growth of just over 11% in Postcode sector mean house prices in the East and South East between 1998 and 1999. Land registry published figures suggest a growth of over 15% in the South East. The anomaly is in part due to our use of annual averages, rather than last quarter prices. Also, our sample includes only those properties with recorded Postcodes. Further investigation indicates that our sample may under represent newer, high-end properties slightly.

The lower panel of Table 5-1 shows some summary statistics for Postcode sector school performance data. The performance measures are fairly similar in each regional group in each year, though The North is always marginally below the other areas. Attainment at Key Stage 2 has improved since the introduction of the performance tables in 1996, though there was little change between 1997 and 1998. School characteristics are also given in Table 5-1 (those recorded in 1999). We see that the East and South East has slightly larger schools, the North has more schools with pre-school and reception years, and the West and South West has more voluntary aided or controlled schools and fewer junior schools. Variation in the age range across areas is attributable to LEA policy – in
some LEAs, primary schools take children from compulsory school age only. In others, primary schools take children from age 4, or even earlier if a nursery is attached to the school.

**Table 5-1: Descriptive Statistics**

<table>
<thead>
<tr>
<th>A. Property prices (£)</th>
<th>East and South</th>
<th>North</th>
<th>West and South</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>East</td>
<td></td>
<td>West</td>
</tr>
<tr>
<td>1996 sector mean</td>
<td>86591</td>
<td>52286</td>
<td>63221</td>
</tr>
<tr>
<td>1997 sector mean</td>
<td>98701</td>
<td>55879</td>
<td>69852</td>
</tr>
<tr>
<td>1998 sector mean</td>
<td>112303</td>
<td>61045</td>
<td>78600</td>
</tr>
<tr>
<td>1999 sector mean</td>
<td>125757</td>
<td>63921</td>
<td>85040</td>
</tr>
<tr>
<td>Mean sales volume</td>
<td>140</td>
<td>95</td>
<td>114</td>
</tr>
<tr>
<td>Detached house mean</td>
<td>165532</td>
<td>94006</td>
<td>110870</td>
</tr>
<tr>
<td>Semi-detached mean</td>
<td>106521</td>
<td>53254</td>
<td>65343</td>
</tr>
<tr>
<td>Terraced mean</td>
<td>96183</td>
<td>40375</td>
<td>54178</td>
</tr>
<tr>
<td>Flat/maisonette mean</td>
<td>69881</td>
<td>40607</td>
<td>43510</td>
</tr>
<tr>
<td>Number of Postcode sectors</td>
<td>2900</td>
<td>2998</td>
<td>1554</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>B: School performance and characteristics</th>
</tr>
</thead>
<tbody>
<tr>
<td>1996 key stage 2, level 4 proportion</td>
</tr>
<tr>
<td>1997 key stage 2, level 4 proportion</td>
</tr>
<tr>
<td>1998 key stage 2, level 4 proportion</td>
</tr>
<tr>
<td>1999 key stage 2, level 4 proportion</td>
</tr>
<tr>
<td>Proportion Community school</td>
</tr>
<tr>
<td>Proportion of schools with pre-school/reception</td>
</tr>
<tr>
<td>Proportion of schools with infants</td>
</tr>
<tr>
<td>School roll</td>
</tr>
<tr>
<td>Number of age 11 pupils present</td>
</tr>
<tr>
<td>Number of schools in Postcode sector</td>
</tr>
<tr>
<td>Number of Postcode sectors</td>
</tr>
</tbody>
</table>

Property prices are matched to lagged Key Stage 2 results in estimation sample
Price means are means of Postcode sector means (unweighted by sales volume)
1999 Key Stage 2 results reported for completeness (not used in estimation sample)
Key Stage 2 assessment tests are sat in Spring and results are released in Autumn

Variation between Postcode districts accounts for 32% of the variance in measured school level performance, and 45% of the variation across Postcode sectors. By contrast we can attribute nearly 80% of the variance in Postcode sector mean house prices to
differences between Postcode districts. The relative variation in school performance across time and geographical space tells us something about the usefulness of exploiting time-series variation in our estimates. Taking the sub-sample of Postcode sectors with primary schools for 1996 and 1999 we find that 75% of the variation in school performance can be explained by Postcode sector fixed effects. Regressing out Postcode sector fixed effects and general time effects, the residual variance is 0.00278, against overall variance of 0.0248. Only 11% of the initial variance is unaccounted for. What is more, if we look at log property prices in the same sub-sample, we find that 95% of the variance is attributable to Postcode sector fixed effects. The residual variance is only 2.5% of the raw variance in log house prices! Clearly, the differences between Postcode sectors in house-price time trends are small relative to other source of variation.

It is also useful to look at how changes over time in school performance are related to initial performance in 1996. Our intuition is that growth will be less in the Postcode sectors with better performing schools in the first period, as it must be at the very top: this is indeed the case. Performance in the bottom 25% of schools in 1996 grew by 23.5 percentage points whilst performance in the top quintile grew by only 5.8 percentage points. The high performing schools which we think might have the biggest effect on prices show the least growth in performance. Because of this, and because there is so little between group variation in our data, variation over time is unhelpful in identifying the response of house prices to school performance.

Before we move on to the regression results, we illustrate the basic association between house prices and school performance in Figure 5-2. This is a kernel regression of the deviations of 1999 log house prices from Postcode district means on the deviation of average 1996-1998 school performance from Postcode district means. The Figure provides a visual test of our log-linear specification; the relationship between house prices and Key Stage 2 performance is upward sloping, monotonic and looks comfortably
linear for all regions. We now move on to discuss the main regression results from the models presented in Section 5.3.

Figure 5-2: Relationship Between Log House Prices in 1999 and Mean 1996-1998 Primary School Performance – Deviations From Postcode District Means

Figure illustrates kernel regression of within-Postcode-district variation in log house prices on within-Postcode-district Key Stage 2 performance. Bandwidths in accordance with Silvermann’s rule of thumb (0.02 for North and South and East, 0.025 for West and South West)

5.5 Results

5.5.1 Baseline results

Baseline results are reported in Table 5-2 for the three area configurations. Each panel of the Table reports three specifications of log-linear regressions of property prices on school performance, controlling for smooth spatial fixed effects. To reiterate, these estimates are regressions using the deviations from the local spatial group means, where these are estimated non-parametrically for each period. The minimum, mean and maximum bandwidths are shown in the Table notes. The distribution of household density on which the bandwidth is based is right skewed, so the median bandwidth is around 1km. To illustrate the way in which the estimator works shows the estimated
function \( g(\epsilon_j,t) \) that defines the smooth spatial fixed effects surface for the London region. Regression estimates are based on deviations of prices from this surface.

**Figure 5-3: Example House Price-Location Surface For London, From Smooth Spatial Effect Model**

Columns (1), (4) and (7) of Table 5-2 include the schools’ lagged, single-year Key Stage 2 score as the measure of performance. Other controls are the observation property type indicator, the Postcode sector in social housing, and the mean number of rooms occupied by home-owning households. Columns (2), (5) and (8) are identical to the above, but regresses 1999 property prices on school performance averaged over 1996-98. Columns (3), (6) and (9) instrument the three-year mean school performance by school type and age range dummies. A common pattern of results appears across all three regional groupings. In all cases there is a positive statistically significant association between house prices and school performance.
<table>
<thead>
<tr>
<th>Region</th>
<th>Proportion reaching target grade, age-11</th>
<th>No primary school</th>
<th>Detached</th>
<th>Terraced</th>
<th>Flat/Maisonette</th>
<th>Mean rooms in owner-occupied housing</th>
<th>Proportion in social housing</th>
<th>Within R²</th>
<th>Overall R²</th>
<th>Overident test p-value</th>
<th>Sample size</th>
<th>Years</th>
</tr>
</thead>
<tbody>
<tr>
<td>South East and East</td>
<td>(0.203, 0.289, 0.711)</td>
<td>(0.183, 0.244, 0.524)</td>
<td>(0.479, 0.481, 0.481)</td>
<td>(-0.185, -0.187, 0.266)</td>
<td>(-0.588, -0.582, -0.582)</td>
<td>(0.196, 0.195, 0.187)</td>
<td>(-0.442, -0.439, -0.257)</td>
<td>0.849</td>
<td>0.934</td>
<td>-</td>
<td>29606</td>
<td>1997-99</td>
</tr>
<tr>
<td>The North</td>
<td>(0.289, 0.372, 0.688)</td>
<td>(0.238, 0.305, 0.502)</td>
<td>(0.524, 0.539, 0.539)</td>
<td>(-0.277, -0.277, -0.277)</td>
<td>(-0.399, -0.387, -0.389)</td>
<td>(0.184, 0.169, 0.170)</td>
<td>(-0.436, -0.488, -0.397)</td>
<td>0.855</td>
<td>0.934</td>
<td>-</td>
<td>1999</td>
<td>1999</td>
</tr>
<tr>
<td>The West and South West</td>
<td>(0.639, 0.289, 0.372)</td>
<td>(0.178, 0.230, 0.304)</td>
<td>(0.261, 0.500, 0.501)</td>
<td>(0.311, -0.188, -0.188)</td>
<td>(0.116, -0.474, -0.480)</td>
<td>(0.555, 0.193, 0.196)</td>
<td>(0.223, -0.319, -0.273)</td>
<td>0.800</td>
<td>0.904</td>
<td>-</td>
<td>1999</td>
<td>1999</td>
</tr>
</tbody>
</table>

Note: 1 Mean is 1996-1998 average school performance

Dependent variable is log of property price

Instruments in columns 3, 7, 11 are Community school status and school admissions age-range dummies, weighted optimally according to clustered error structure

Min, mean, max bandwidth: .23 km, 1.43 km, 9.25 km in South East and East; 0.35km, 1.72km, 26.07 km in The North; 35 km, 1.75 km, 26.46 km in the West and South West

The sample centroids are, by region: (52835,18350), (41279, 41852), (36027, 20062)

Sample comprises unbalanced panel with up to 4 house-price observation types (Detached, Semi-Detached, Terraced, Flat) for each Postcode sector, observed for 1997, 1998 and 1999 (1999 only in columns 2 and 5)

School performance measures are the proportion in the school obtaining Key Stage 2 at Level 4 and above. Key Stage 2 results are average of maths, reading and science scores. Standard errors corrected for clustering on Postcode sectors (in parentheses).

- 224 -
There are some differences between the point estimates in each of the four specifications, but the general pattern of results looks very similar when viewed across regions. In fact, the parameters for each specification are statistically identical across regions. For Columns (1), (4) and (7), the minimum distance parameter is 0.240, and we do not reject equality of the parameters across regions ($p$-value = 0.117). As we expected, the 3-year performance means in Columns (2), (5) and (8) provide better measures of long-run performance than the year-to-year results, so the impact on prices is bigger. The minimum distance estimate is 0.331 and again we do not reject equality across regions ($p$-value = 0.695). Based on the 3-year performance averages, a 10% increase in the mean Key Stage 2 performance in a Postcode sector is associated with a 3.4% premium on property prices – equivalent to £3300 on an average-price property in England in the last quarter of 2000.

We now turn to our Instrumental Variables estimates. Let us first consider the suitability of our instruments. As already discussed in some detail above, we use admissions age-range and school-type indicators as instruments for Key Stage 2 performance. The instruments are good predictors of school performance, as we can see in Table C1 in Appendix C.

The Table shows the estimated coefficients on the variables excluded from the property price equation, and the F-tests for their exclusion from the school performance equation. We see a broadly similar pattern across regions, and the F-statistics and t-statistics are always high. The proportion of children at Community schools achieving Level 4 is between 4.6% and 6.6% lower than those at church schools, and overall the instruments generate a ±10 percentage point range in predicted performance.**

We should note that Neal (1997) finds that the advantage of a catholic secondary education in the US varies by geographical location. He finds that it is only in urban areas that catholic schooling offers clear benefits. If this were true in our sample, we would have to question the
Now return to the main results of our IV procedure in Columns (3), (6) and (9) of $$$. For all three regions, the IV estimates of the implicit price of Key Stage 2 performance are substantially higher than the estimates using three-year means. Note though that the statistics for Sargan test of the overidentifying restrictions imply that our instruments are uncorrelated with the residuals from the regression⁹⁰. That the IV-SSE estimates are higher is, in part, surprising. We might expect the IV estimator to remove residual catchment-area-to-school-performance effects that upward bias the least squares estimates of willingness to pay for school quality. The results suggest that these effects are not the principal source of bias. Instead, the use of year-to-year performance measures, and even 3-year averages, seriously attenuates the estimated impact of long-run school quality on property prices due to the noise components of the raw, transitory performance measures. We have further evidence that this what is happening. As it stands, the models in Columns (1), (4), (7) have period-specific smooth spatial surfaces

usefulness of community/voluntary status as an instrument, without area interaction effects. We investigated whether religious school advantage we detect here varies by area, but found no clear pattern. In London, the religious school advantage rises to around 12.3% (s.e. 1.04%), but it is similar or even higher in some, predominantly rural Postcode areas, for example Peterborough (13.5%, s.e. 4.1%) or Carlisle (11.3% s.e. 5.8%), and lower in other urban areas such as Manchester (7.0% s.e. 3.5%). Unfortunately, Neal's analysis of the impact of catholic schooling sheds little light on the endogeneity of religious status with respect to neighbourhood status, as he disregards sorting by parents on school quality.

⁹⁰ The performance advantage of church schools does not appear to be related to selective admissions procedures by Voluntary Aided schools (who may conduct interviews to determine religious convictions). Voluntary Controlled schools, where the LEA administers admissions, also have better pass rates.

⁹⁰ The exclusion restrictions are rejected if we do not control for social housing. This result is consistent with our observation that Community schools are more likely to be located near Local Authority housing estates. Apparently, the poorer performance of Community schools relative to religious schools is in part due to their catchment areas containing a higher proportion of Local Authority tenants.
allowing area specific time trends. Instead we could use a single cross sectional surface and allow for general time effects with time dummies. Clearly this allows area-specific time series variation a greater role in determining the coefficient estimates. What we find though is that the estimated impact on prices is halved; the conditional variance in school performance increases with the transient noise components without corresponding increases in the covariance with prices.

Based on the IV estimates in Columns (3), (6) and (9) the minimum distance estimate across regions is now 0.663, with equality across regions (p-value = 0.859). This 6.9% premium amounts to about £6600 for a 10 percentage point performance improvement, in England in 2000 prices.

5.5.2 Comparison with private sector fees

It is natural to ask how these estimated implicit prices compare with private sector fees. Private sector fees should provide an upper bound to what households are prepared to pay via the property market. In the private sector, the equivalent of primary schools are “preparatory” and “pre-prep” schools, covering the age-range from nursery to age 13. The total number of accredited nursery, pre-prep and prep schools in England on the Independent Schools Information Service (ISIS) database is 717, with nearly 40% of these in London and the South East. The mean national average reported by ISIS for 515 prep and pre-prep schools is £6324 in 2001. Assuming this is paid for eight years, and discounting at a rate of 5%, the present value of the costs of this investment amount to about £38,000. Unfortunately we have no information on age-11 performance for private schools. But, we can guess that parents paying for private primary education would expect nearly everyone at the school to reach the equivalent Level 4 in Key Stage 2, implying a 25 percentage point advantage over the mean state sector primary school in 1999. In terms of property prices in the last quarter of 2000, this performance advantage would be worth roughly £17400 nationally, or around £35300 in London, where property
prices are highest. This suggests that for families with only one child in London, the
capitalised costs of state-sector primary education (over and above the unavoidable direct
costs of taxation) are close to the costs of a private-sector primary education. For families
with, or intending to have, more than one child of primary school age, and for those in
other areas of the country, moving house is probably a cheaper option. Even in London,
the state-sector is cheaper in annual terms: the mortgage costs associated with a 25
percentage point improvement amount to around £2500 each year.

5.5.3 Robustness

Table 3 summarises our parameter estimates under a range of alternative
specifications. Firstly, in Row (1) we show what happens if we specify Postcode district
geographical effects. These results are very close to those from our smoothed spatial
effects approach. Row (2) shows that we are not identifying a church school effect alone,
since we can eliminate the Community/church school indicator from the instrument set.
The argument that our results reflect possible selection on academic ability by church
schools is refuted by the results in the Row (3); including controls for special educational
needs – used here as a proxy for academic abilities – makes little difference. The IV
estimates actually increase slightly. A particularly strong result is that there is no direct
relationship between special educational needs and property prices in our model; it is the
institutional differences in Key Stage 2 performance that makes our school types and age
ranges effective as instruments.
In Rows (4) and (5) we add further school demographics. Controlling for ethnic mix has no impact on the IV estimates. Controlling for the proportion of children on free meals attenuates the OLS estimates, but then this measure of school intake poverty is partly determined by the very selection process we are trying to observe. Only in the North does the free school meal proportion affect the IV estimates. Even so the coefficient can be restricted to be equal across regions at 0.694 (p-value = 0.770), which is hardly different from our main estimate of 0.663. Rows (6) and (7) of Table 5-3 show...
that the estimates with smooth-spatial-effects are remarkably insensitive to an increase or reduction in the local averaging bandwidth. 91

We have also tested the effectiveness of our smooth spatial effects in removing spatial correlation in the data. We would expect any important neighbourhood influences on property prices, that we have failed to capture using the regressors and spatial effects in our models, to induce spatial correlation between Postcode sector residuals. Calculating the Moran I 92 coefficients for the South and South East regions on the residuals from the model without spatial effects (for one year, without school performance as a regressor, and averaging across property types) we get a value of 0.3299 (z-statistic = 15.576). This implies quite strong and significant spatial correlation, even with our property and social housing controls. Using the residuals from the smooth spatial effects models, this falls to -0.0012 (z-statistic = -0.854); clearly there is virtually no systematic relationship between residuals in nearby Postcode sectors in our full SSE models. Similar results hold for the other regions.

An important consideration is whether we should include secondary school performance in our regressions. Nothing suggests to us that we should. If we include state secondary school GCSE pass rates in our models then we get low and statistically insignificant coefficients, and the primary school effect is virtually unchanged. In part,

91 We have also tested our specification of a constant impact of school performance across property types, by using interactions in the non-IV models. Only for detached properties is there a significant difference: the response of detached property prices to school performance is about 90% of that for other property types. This makes virtually no difference once we calculate the mean response across property types

92 Moran's I is defined by $I = \frac{e'Me}{e'e}$ where $M$ is a spatial weight matrix, and $e$ is the residual vector. We use a matrix of inverse distances between observations. Standard errors are calculated by simulating the null distribution, using repeated random allocation of the data to the Postcode sector grid references.
this could be because of our estimation strategy; secondary school catchment areas are much larger than those of primary schools and we eliminate impacts on property prices at this broader geographical level. We know anecdotally of examples where extremely highly regarded state secondary schools push up local house prices, and Rosenthal (2000), Leech and Campos (2001) and Cheshire and Sheppard (2002) give empirical examples for England. But the existence of such effects does not impact on our results. This issue is pursued further in the next section.

5.5.4 Results for London using different strategies

The results from the cross-LEA boundary model outlined in Section 5.3.5 are presented in the top panel of Table 5-4. Given the detailed map-work and analysis required here, we focus on Greater London. We also compute the SSE estimates for all Postcode sectors within the same geographical boundary. The comparable parameters are in the bottom of Table 5-4. Of course, the samples are not the same. The cross-LEA method relies on a selected sub-sample of Postcodes in the London area adjoining LEA boundaries. In both cases, we provide estimates based on annual performance measures, 3-year averages, and IV models, all with controls for social housing and house size.

Remember that our comparable baseline parameter estimates were 0.240, 0.331 and 0.663 for the annual, 3-year average and IV-SSE models respectively. The comparable cross-LEA models for London are 0.107, 0.353 and 0.747. There is some deviation in the point estimates, but the standard errors are large for the cross London boundary model. The attenuated coefficient when we use transient single-year performance measures is, we suggest, due to the poor signal to noise ratio. But nothing here suggests that our imprecise definition of catchment areas in the SSE framework leads us to underestimate the response of property prices in our baseline results. In fact, we do not reject equality with the baseline SSE estimates: the minimum distance 3-year average estimate is 0.365, p-value 0.998; the minimum distance IV estimate is 0.641, p-
Neither do we reject equality with the SSE estimates computed for London alone – which are close to the baseline estimates across all regions. Both methods are consistent, but the SSE approach is efficient relative to the cross-LEA boundary estimator.

### Table 5-4: Comparison Of Coefficients From Cross-Local Authority Boundary Effects, SSE Estimator and Property-level Micro Data For Greater London Area

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1-Year</td>
<td>3-Yr Mean</td>
<td>IV</td>
</tr>
<tr>
<td><strong>X-LEA model, Postcode sector data</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.107</td>
<td>0.353</td>
<td>0.747</td>
</tr>
<tr>
<td>Sample size</td>
<td>(0.050)</td>
<td>(0.134)</td>
<td>(0.304)</td>
</tr>
<tr>
<td></td>
<td>4012</td>
<td>1351</td>
<td>1351</td>
</tr>
<tr>
<td><strong>SSE model on London area, Postcode sector data</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.195</td>
<td>0.231</td>
<td>0.596</td>
</tr>
<tr>
<td>Sample size</td>
<td>(0.038)</td>
<td>(0.068)</td>
<td>(0.171)</td>
</tr>
<tr>
<td></td>
<td>11051</td>
<td>5274</td>
<td>5274</td>
</tr>
<tr>
<td><strong>P-value of parameter equality:</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hausman test (assuming same sample)</td>
<td>0.007</td>
<td>0.291</td>
<td>0.548</td>
</tr>
<tr>
<td>Assuming independent samples</td>
<td>0.161</td>
<td>0.417</td>
<td>0.665</td>
</tr>
</tbody>
</table>

**Traditional property-level micro data model**

|                                      |         |         |         |
| **Primary schools**                  |         |         |         |
|                                       | 0.454   | 0.664   | 1.041   |
|                                       | (0.055) | (0.014) | (0.232) |
| **Secondary schools**                | -0.010  | -0.033  | 0.134   |
|                                       | (0.038) | (0.045) | (0.086) |

**Property-level micro data with spatial effects**

|                                      |         |         |         |
| **Primary schools**                  |         |         |         |
|                                       | 0.174   | 0.263   | 0.756   |
|                                       | (0.079) | (0.092) | (0.321) |
| **Secondary schools**                | -0.011  | 0.009   | 0.047   |
|                                       | (0.061) | (0.078) | (0.116) |
| **Sample size**                      | 8067    | 8067    | 8067    |

Samples:
- Postcode sectors adjoining Local Authority boundaries, on the Geoplan Greater London Postcode sector map
- All Postcode sectors within the same boundary
- Micro data covering 800km² of the London region in 2001, provided by Ekins surveyors

Controls in micro-data models are property style dummies (12), number of rooms, number of floors, garage, year built, floor area, distance to: CBD, CBD², local town centres, green spaces, underground stations, police stations; plus mean neighbourhood house size, proportion in social housing, housing density, population density, density of purpose built flats, incidents of criminal damage per km².
By way of further comparison, we have estimated models using our property-level data for the London region. We match school performance to properties using spatially-weighted average performance of the nearest 8 primary schools. We can now match secondary schools too, and we choose a spatially weighted average of the nearest 5. The lower panel of Table 5-4 shows these results. Other controls in the regression are described in the table notes. The first two rows in the lower panel provides parameter estimates using models with no spatial controls other than a quadratic in distance to the central business district, distances to various amenities and Local Education Authority dummies. We find weak impacts here from secondary schools, but our primary school coefficients are somewhat higher than those in our main results. In the next two rows we apply our SSE techniques to the micro data to remove unobserved spatial impacts on prices. The resulting coefficients on primary schools, are very close to our results using Postcode sector mean data with few property controls.

We infer from this that we do not need precise mapping of the areas served by schools; nor do we need micro data on properties and their characteristics. The relationship between performance and property prices can be analysed successfully in terms of spatial surfaces measuring expected price and school performance, once we abstract from more general price variation over space using our SSE technique.

5.5.5 Evaluating the returns to Key Stage 2 attainments

In Section 5.2.1 we showed that we could infer the expected private returns to better schooling in terms of increased lifetime earnings, in the case where education increases productivity and is valued purely for its impact on lifetime earnings. Using equation is (5-6) we can tentatively infer the expected returns to future earnings from an
increased probability of attaining of Level 4 in Key Stage 2. However, we need some further assumptions since we have no data on expected mean lifetime income of children at our Postcode sector unit of analysis. But we do have family income at this level, in a commercial marketing data set (from CACI Ltd.). From this we estimate expected family income of the next generation using an intergenerational mobility function:

\[ y^c = \bar{y}^p + \rho \tilde{y}^p \]

where \( y^c \) is the child’s family income, \( \bar{y}^p \) is mean family income of the parent’s generation, and \( \tilde{y}^p \) is the deviation of own parents income from the mean.

To parameterise \( \rho \), we refer to the literature on intergenerational mobility (Dearden, et al. (1997)). Assuming family income is received forever and applying a discount rate of 5% to calculate the present value of lifetime income, we can replace the variables in equation (5-6) with their Postcode sector averages.

Using this method, our estimated median returns to Key Stage 2 Level 4 are around 0.24% to 0.27% for a 1 percentage point improvement in Key Stage 2 performance. The mean result is fairly insensitive to the choice of \( \rho \) (which mainly effects the variance), but will obviously depend on the discount rate applied. The result implies that, on average, parents expect the lifetime income of a child who attains Level 4 in Key Stage 2 to be around 28% higher than one who does not.

### 5.6 Conclusions

In this Chapter we ask how much parents are prepared to pay to get their children into better schools by moving house. We use Postcode sector level data on house prices and primary school performance in England to estimate the magnitude of the association

---

93 This calculation ignores the long-run, intergenerationally-transferable asset value of any increase in property value attributable to school performance.

94 Values of 0.2, 0.4 and 0.8 are used.
between primary school quality and local house prices. We eliminate the effects of catchment area wealth on pupils’ achievements by concentrating on the effects within narrow geographical areas, and by instrumenting measured pupil achievements by characteristics of the school itself. Our best estimates imply a premium on Postcode sector house prices of 6.9% for each 10% improvement in the proportion of children reaching the target level in Key Stage 2 tests at age 11. This translates into capitalised values, at regional mean property prices, between £4500 in the North and £13500 in Greater London (all at 2000 property prices). Interestingly, our estimates of the primary school effect are of the same order as those in (Black (1999)) for suburbs of Boston, Massachusetts. She finds that a 10% increase in primary school mean test scores attracts a 5% property price premium. Using a similar estimation technique and London data, we find a 3.6% property premium. Our lowest estimates based on OLS, within-area estimators put the figure at around 2.4% for a 10% school improvement.

The sensitivity of property prices to local primary school quality implies the existence of a back-door selection of pupils by the incomes of their families. This flies in the face of notions of equality of opportunity, is likely to restrict intergenerational mobility and generates an inequality of educational outcomes that may be unrelated to the abilities of children. If pupil ability is related to parental incomes then selection by income is implicitly selection by academic ability. Indeed, this goes against the principle in the DfEE code of practice on admissions (Section 5.6) that “academic ability should not be used to decide entry into primary education”. The equilibrium arising from local sorting by incomes on primary school quality will be inefficient if the net marginal benefits of state school quality are greater for lower income families. This is almost certainly true given that the alternatives – private sector schooling, private personal tuition – are available at lower marginal cost to wealthier families with sufficient capital or lower borrowing costs. As usual with issues of educational equity, relaxation of borrowing constraints is a fundamental issue here. Linking of property loans to current
incomes means that the marginal costs of borrowing become infinite at lower and lower purchase price thresholds as incomes decrease. This is sensible given the need to match lending to borrowers’ ability to repay the debt, but leads to exclusion of those on low incomes from the benefits of good local schooling.

The primary objective for policy seeking to remove inequities and inefficiencies arising from income-related selection on good state schools is to eradicate differences in primary school quality across geographical space. Current government policy is to increase competition between schools as an incentive for good performance. However, proximity-based restrictions on admissions, together with the house price effects shown in this Chapter, mean that higher income families will inevitably benefit the most. Lower-income home-owners will be priced out of the best school catchment areas. More public information on school performance differences could exacerbate this problem, though there seems to be no evidence that house prices are more sensitive to school quality over the years that the school performance league tables have been available.

The clear message that emerges is that households value improvements in primary school performance. Importantly, this valuation relates to differences attributable to exogenous schooling inputs, not simply to exogenous neighbourhood status. From this we infer that school inputs must matter. Lack of suitable data means we cannot empirically address the question of which inputs matter most. This is the appropriate question for policymakers looking for a policy lever, and more research on this question using detailed data on children and schools is vital. Nonetheless our findings are important as they show parents to strongly value better school performance.

We should however note an alternative explanation for our results. It is possible that certain observable school characteristics act as a focal point for high-income parents seeking high-income or high-ability peer groups for their children. Our use of community/voluntary status as an instrument for school performance may be open to this objection in that non-community status may offer no advantages in terms of expenditures,
teaching techniques or other inputs, but historical belief that these schools are better may lead high income parents to converge on them. Their performance advantage would then be purely attributable to the characteristics of the children, or parents of the children, and the peer-group benefits of mutual association. Whilst this is plausible, it seems highly unlikely: we get similar results when using only age range as an instrument, and controlling for differences in ability composition has no effect. As far as we know, age-range is not widely used by parents as a signal of school quality.

If it is really peer groups, and not school inputs that matter, then our results amount to a valuation of a peer-group effect in primary education. If neither peer groups nor inputs matter, so differences in school performance between school types are purely attributable to the distribution of child and parental characteristics at the school, then sorting on school types and the school-property price premium is highly irrational and inefficient. Under this scenario, high income families would do better to send their children to schools which score low in the performance tables, where the attainments of their own children would be identical to their attainments at a ‘good’ school.

Let us suppose that we have correctly measured willingness to pay for school quality improvements. Extrapolating from our results, we can say that any technology which raises primary school standards by one percentage point has a social valuation per household equivalent to between 0.33% and 0.67% of the local mean property price. But some caution is needed in extending this result to aggregate educational improvements: parents may value good schools because of the advantage they offer their children relative to others in the same generation. In this case, a national, across-the-board one percentage point increase in school standards may be worthless. Assuming better schooling really increases productivity or otherwise brings social benefits we can make a simple valuation of national-level policy changes. For a national population of 21 million households, and mean property price of £96700 at the end of 2000, our figures imply a maximum aggregate social valuation of £13,600 million, or about £1.09 million per school for a one
percentage improvement, in present value terms. If we include only the 18.5 million households who are resident in Postcode sectors containing primary schools (under the assumption that those elsewhere place no value on primary schools) and take the lower estimate we get a lower bound of £5,900 million, or £0.47 million per school. This means that a sustained one percentage point improvement in primary school performance scores is valued somewhere between £39 and £90 per year, for each child of primary school age or younger. By way of comparison, it is worth noting that current annual expenditure on the National Literacy and Numeracy Strategy of £190 million is worth about £50 per pupil per year (Department of Education and Skills (2001) p.28).

---

95 These calculations assumes 12500 Key Stage 2 primary schools, 7.6 million children age 11 and under and a 5% social discount rate.
5.7 Appendix A: Bandwidth choice

The choice of bandwidth for the kernel in our SSE estimator is important, since the appropriate comparison needs to encompass more than one Postcode sector, without averaging over too broad an area. Since the appropriate area (in terms of geographical distance) depends on local household density, we need to take this into account. Postcode district HG4, just north of Harrogate, has an area of roughly 270 km² and Postcode sector household densities that range from 20 to 1300 per km². By contrast, E3 around Bow and Tower Hamlets in east London has an area of roughly 4.5 km² and household densities between 6000 and 6800 per km². No fixed bandwidth can accommodate this variation: a suitable bandwidth choice at HG4 will average over much of the London area if applied to a sector in E3. A bandwidth suitable for E3 if used in HG4 will apply virtually no weight to any observations beyond the Postcode sector. Consequently, we weight the neighbourhood bandwidth using data on household density matched in from the 1991 Census.

Fixing the number of households \( n \) in a circular spatial group of radius \( b \), gives us a bandwidth weighting rule dependent on housing density \( h \):

\[
b = \sqrt{\frac{n}{\pi h}}
\]  

(5-11)

In order to choose a bandwidth regulator \( h \), it is useful to know how our Postcode sectors relate to primary school catchment areas. This is made more difficult by the fact that we could obtain almost no information on this from our enquiries to LEAs, as catchment areas are rarely precisely defined, and vary with demand. Data on addresses of pupils actually attending is considered confidential, and is usually held only by the schools themselves. We were unable to obtain this. From our primary school performance data, the total number on the school role of primary schools recorded in the 1999 performance tables is 3.77 million and the total number of imputed households in our CACI data is 20.1 million. The ratio of households to primary school children is 5.33, implying an average catchment area of around 1400 households, which is about half a Postcode sector. This is consistent with the fact that there are, on average, two primary schools per Postcode sector in the school performance tables. Choosing bandwidths corresponding to groups of roughly one, two and three Postcode sectors and adjusting downwards by 40% to compensate for the use of a Gaussian kernel (which applies non-zero weights to observations outside the bandwidth window), suggests corresponding household groups of roughly 1700, 3400 and 5000 respectively. The main results we present use bandwidths corresponding to 3400 households, but comparisons are made with other bandwidth choices.
5.8 Appendix B: Details of the data sources

Data on individual housing transactions is unavailable in Britain, so we have used the best available alternative: house prices aggregated to Postcode sector level. This data set covers the whole of England and Wales, and is available from 1995 to 2000. It contains mean house prices and total sales volumes at Postcode sector for each Postcode sector, where annual sales numbered 3 or more. Properties sold for under £10,000 and over £1,000,000 are excluded. This amounted to only 0.5% of all property sales in 1999.

In the UK, Postcodes contain up to seven alphanumerical characters, and contain four hierarchical components. The first two alphabetic characters define the Postcode Area, the broadest postal zone. Examples are N, EX and YO representing North London, Exeter and York. Within Postcode Areas, the next level down is the Postcode District. This is defined by a single or two-digit number following the Postcode Area. Examples are N6, EX24, and YO10. A single letter further subdivides some Postcode districts in central London. Below this, we have Postcode Sectors. This is the unit of observation in our house price data set.

The school performance tables for England compiled by the Department for Education and Employment (DfEE) provide the basis for our school performance measures. We have the 1999 primary and secondary school tables, which include background information on the schools in 1999, plus the performance measures for years 1996 to 1999 inclusive. We also have the original data for the years 1996-1998 which includes the school background characteristics for these years. The primary performance measures are proportion of pupils reaching Level 4 (the target level of attainment) in the Key Stage 2 standard assessment tests administered at age 11. We average the measures for Maths, Reading and English tests. We average these school performance measures and characteristics across schools within each Postcode sector to provide a Postcode sector level primary school performance indicator and characteristics. Here, we experimented with simple means and school-size weighted means, but opted for the former on the basis that weighting by school size conflates school size and school performance issues. In practice, the choice of scheme made little difference to our results.

We match Postcode sector house prices to the Postcode sector school performance and characteristics from the school data set, giving us up to four house prices (detached, semi-detached, terraced, flat/maisonette) for each Postcode sector in each year.

Additional variables at Postcode sector level are derived from the 1991 Census, and from the 1998 Postcode to Census Enumeration District directory, which relates 1998 Postcodes to corresponding 1991 Census area codes. These sources give us geographical data including the national grid reference, the proportion of social housing, and the density of households per kilometre-squared. Although Postcode-sector aggregated Census...
data is available, the Postcodes relate to the 1991 Postcode geography, so the Census variables we use are means of the values in the Enumeration Districts which are wholly or partly included in a given Postcode sector. Grid references are taken as the mid point between the maximum and minimum in each direction.

No population bases are available at the Postcode sector level later than 1991, though we have household figures in our CACI data set on household incomes. The mean number of household addresses per Postcode sector in the CACI data in 1999 is 2800. In the UK there are 26 million postal addresses, 2901 districts and 9624 sectors, so a crude average is 9000 per Postcode district, 2700 per sector. These numbers change over time with changes in the Postcode geography. In 1996, the number of households in England was 20.2 million, implying an average of around 9600 households per Postcode district, and around 2560 in each Postcode sector.

Appendix C: Underlying prediction equations for IV estimates

Table C1: Coefficients on instruments in first stage IV equation

<table>
<thead>
<tr>
<th></th>
<th>South East and East</th>
<th>North</th>
<th>West &amp; South West</th>
</tr>
</thead>
<tbody>
<tr>
<td>Community school – LEA appointed governors and admissions</td>
<td>-0.066 (0.006)</td>
<td>-0.046 (0.007)</td>
<td>-0.064 (0.008)</td>
</tr>
<tr>
<td>School has pre-school/ reception years</td>
<td>-0.034 (0.009)</td>
<td>-0.014 (0.009)</td>
<td>-0.032 (0.016)</td>
</tr>
<tr>
<td>School has infants and junior years</td>
<td>0.022 (0.006)</td>
<td>0.069 (0.008)</td>
<td>0.042 (0.009)</td>
</tr>
<tr>
<td>F-test of instruments</td>
<td>F(3, 843) = 72.0 P = 0.0000</td>
<td>F(3,713) = 67.6 P = 0.0000</td>
<td>F(3,502) = 37.5 P = 0.0000</td>
</tr>
</tbody>
</table>

Predicted Key Stage 2 performance, from identifying instruments (all areas):
- s.d. = 0.02
- max = 0.102
- min = -0.097

Models include property type dummies, proportion of local social housing, and are estimated within Postcode-district-year groups

Standard errors (and F-tests) corrected for clustering on Postcode districts.

Results shown for illustration only; estimation of main models does not use 2-stage least squares method.
6 The Costs of Urban Property Crime

6.1 Introduction

Urban crime has a powerful influence on perceptions of area deprivation. Criminal damage to public and private property symbolizes urban decay, and fear of burglary and theft promotes insecurity and anxiety. Crime prevention and control policy is top of the political agenda in developed countries, and these problems are particularly acute in the urban environment. Although no place is crime-free, the fear of crime and the direct costs associated with property crime can have particularly severe consequences in urban areas, in discouraging local regeneration and catalysing a downward spiral in neighbourhood status. This ‘tipping’ process has a prominent role in criminological explanations of community change and crime Bottoms and Wiles (1997). Policy makers in Britain apparently share this view, arguing that “neighbourhoods have been stuck in a spiral of decline. Areas with high crime and unemployment rates acquired poor reputations, so people, shops and employers left. As people moved out, high turnover and empty homes created more opportunities for crime, vandalism and drug dealing” (Social Exclusion Unit (2001) p.7). Certainly, casual observation suggests that persistently high local crime rates deters new residents and motivates those who can to move out to lower-crime rate neighbourhoods. We would expect this demand for low-crime neighbourhoods to be revealed in a property or land price gradient between residences in high and low-crime localities.

The evidence from the US based on hedonic models (Hellman and Naroff (1979), Thaler (1978), Lynch and Rasmussen (2001)) suggests that crime rates do effect property values, although the effects may be small below high-crime thresholds. Lynch and Rasmussen (2001) find that a one percent increase in violent crime rates reduces prices by
0.05 percent, but report positive associations of property crime rates with prices. This they attribute to higher reporting rates in wealthier neighbourhoods, but higher victimisation rates may provide a better explanation. Properties are heavily discounted in high-crime neighbourhoods. For the UK, however, there is no existing evidence on the relationship between urban crime and property values. We address this here by estimating the effect that crime rates have on property prices in the inner London area, using spatial property crime data provided by the Metropolitan Police. Following the traditional hedonic literature, we interpret this as measuring households' marginal willingness to pay to avoid crime, or the implicit costs of crime.

One problem with existing studies is that identification relies on inclusion of an ad-hoc set of control variables at the household and neighbourhood levels. No attempt has been made to deal with the potential endogeneity of crime rates in a property value model. In this Chapter we deal carefully with this issue. We apply a semi-parametric regression approach that is useful for abstracting from unobserved price variation induced by access to local amenities and changes in the unobserved physical characteristics of property over geographical space.

The paper is structured as follows. The next section sets out the empirical framework for our estimates, and goes into some detail on how we think we can identify the impact of crime density on property values. Section 6.3 discusses our data sources. Section 6.4 presents the results, and discusses their interpretation. Section 6.5 concludes.

6.2 Empirical model and methods

Our task is to measure the impact that property-based crimes in the immediate neighbourhood have on the price of residential property. But this highlights a general problem with the use of property value models to infer the implicit price of local characteristics that reflect the behaviour of local residents. Clearly, the behaviour of neighbours will depend on their individual characteristics, and these may well be
systematically related to unobserved determinants of property prices. Consequently we may falsely infer a causal relationship between local characteristics and property prices, when in fact it is the unobserved component of property values that drives neighbourhood composition. Consider this example: local land prices will attract low-income residents, and if low-income residents are prone to commit crimes in their own neighbourhood we will find more crime in low land price neighbourhoods. Unless we can observe land prices, regression estimates of the impact of crime on property prices will be biased towards finding a negative relationship.

On the other hand, estimation of the implicit price of crime presents an additional problem. Burglars will target properties where the expected return in terms of the market value of stolen goods is highest. Since high land price neighbourhoods will have high proportions of high-income residents, the returns to burglary in high land price neighbourhoods will be high. We can expect to find high burglary rates in these areas, other things equal. To proceed, we must pay careful attention to the unobserved components of property values that are area specific, and attempt to control for these in our estimation technique.

To understand and tackle the problem, we need to structure what we are doing fairly carefully. We assume the following structure for the joint determination of crimes and property prices:

\[ \ln P_i = \beta x_i + \gamma z_i + m(u | c_i, h_i) + \nu_i \quad (6-1) \]

\[ x_i = \rho m(x | c_i, h_i) + \delta z_i + \lambda m(u | c_i, h_i) + \sigma \nu_i + \epsilon_i \quad (6-2) \]

Equation (6-1) says that the log-price of property \( i \) is dependent on the incidence of property crimes in the neighbourhood \( x_i \), a vector of exogenous property and location characteristics \( z_i \), plus spatially correlated unobserved components \( u_i \) and a random error term \( \nu_i \). Equation (6-2) says that crimes in the neighbourhood of a property depend on crimes in the broader geographical area \( m(x | c_i, h_i) \), on the observed property and
location characteristics $z_i$, on the unobserved property and location characteristics $m(u|c_i,h_i)$, $v_i$, and on a random error term $\epsilon_i$. The function $m(\xi|c_i,h_i)$ represents a locally weighted average of $\xi$, with weights on each observation determined by their distance from the location $c_i$ of observation $i$. We can think of this as the expected value of a random variable $\xi$ in the broader geographical area of observation $i$, and it captures the impact of location and local amenities.

In more detail, the unobservable components in the property price equation (6-1) are as follows. Firstly, $m(u|c_i,h_i)$ represents factors jointly influencing crime and the prices of properties in the broader geographical area – let us call this the district. A prime example is the land price, which determines property prices, the supply of criminals and the expected returns to crime in the area. Parameter $\lambda \neq 0$ in equation (6-2) implies that crimes in the district and average district property prices are jointly determined. Secondly, error term $v_i$ represents factors jointly influencing the price of a specific property or properties in its immediate neighbourhood and criminal activity at that same location. We might think of large windows or secluded gardens that make a residential area attractive to both burglars and home-buyers, or poorly maintained property that attracts vandals and a low market price. For example, it is known that victimisation rates vary with type of household and so in principle with types of property (Tseloni, et al. (2002)). Hence, recorded crime rates will be endogenous to housing prices unless all housing attributes are observed. So, parameter $\sigma \neq 0$ in equation (6-2) implies that crimes at the property or in the immediate neighbourhood and the property price are jointly determined.

In the crime equation, parameter $\rho$ measures the dependence of criminal activity in the neighbourhood at a given property location on criminal activity in the surrounding district. This might arise for instance through opportunistic burglaries or vandalism in a
street by criminals targeting nearby areas. We allow for spatial correlation in crime rates, since this provides one potential source of identification, as we shall see below.

6.2.1 Identification of the impact of crime on property prices

As it stands, OLS estimation of the hedonic price function in (6-1) produces inconsistent estimates, because of the correlation between \( x_i \) and the unobserved price components, implied by \( \sigma, \lambda \neq 0 \). Let us assume for a start that we can proxy the important local determinants of property prices by some parametric function of observable characteristics (distance to the central business district, local amenities and the like), such that \( m(u | c_i, h_i) = 0 \). Parameters estimated in (6-1) by OLS will still be inconsistent, because we have not dealt with the fact that unobserved property characteristics may determine crime rates in the immediate neighbourhood (\( \sigma \neq 0 \)). But we can obtain consistent estimates by a standard Instrumental Variables estimator, using the spatial lags of crime, \( m(x | c_i, h_i) \), as instruments, since \( E[v_i | m(x | c_i, h_i), z_i] = 0 \) by assumption. The intuition here is that if reported crime density at a given property location is higher because of unobservable attributes of the properties, then the expected number of crimes in the wider district is a suitable instrument — but only once we’ve removed spatial correlation in the unobserved determinants of property prices\(^{96} \).

This is fine if we know what variables to include to get rid of \( m(u | c_i, h_i) \) and remove the residual spatial autocorrelation. The problem with this approach is that it is

\(^{96}\) Note that crime density \( x_i \) at each location depends on a weighted average of crime densities at other locations \( m(x | c_j, h_j) \). Hence, \( x_i \) itself contributes to \( m(x | c_i, h_i) \) through other observations’ spatial lags, i.e. \( x_i = \rho m(x | c_j, h_j) + \cdots + \epsilon_j \). This could invalidate our use of the spatial lags as instruments. However, as the number of observations \( K \) included in the weighted average becomes large, the effect of any individual becomes negligible, so the estimator is consistent if \( K \) is proportional to the sample size.
data intensive, and we need some prior assumptions about which amenities are important enough to warrant data collection. Moreover, proxying neighbourhood attributes with the characteristics of owner-occupying residents will lead to inconsistent estimates, because residents' characteristics are correlated with unobserved determinants of area property prices through sorting and selection processes.

If we do not have this information, the following transformation of (6-1) is useful. In the fashion of a standard fixed effects estimator, we work in deviations from the local spatial average of the variables (centred on observation $i$ at coordinate $c_i$):

$$\ln P_i - m(\ln P_i | c_i, h_i) = \beta [x_i - m(x_i | c_i, h_i)] + \gamma [z_i - m(z_i | c_i, h_i)] + v_i$$  \hspace{1cm} (6-3)

$$x_i - m(x_i | c_i, h_i) = \delta [z_i - m(z_i | c_i, h_i)] + \sigma v_i + \varepsilon_i - m(\varepsilon_i | c_i, h_i)$$  \hspace{1cm} (6-4)

This transformation gets us out of having to specify a full model of price determinants, but means we no longer have a spatial lag instrument for property-specific crimes, and any Instrumental Variables procedure requires exclusion restrictions on $z_i$ in (6-3). A plausible candidate instrument is the number of offences reported on non-residential properties in the immediate vicinity. To see this, consider a house in a residential street located near a parade of retail outlets or other commercial premises. The incidence of crimes reported at the commercial premises and the incidence of crimes reported in nearby dwellings will be correlated in that the same criminals may be active in both. But the returns to crime in each type of premises are plausibly uncorrelated\(^\text{97}\). There is little reason to believe that victimisation rates in commercial and residential premises will be related, except through the local supply of crimes. In Section 6.4.4 we consider another instrument, based on the link between alcohol consumption and crime – the distance to the nearest public house or wine bar.

\(^\text{97}\) Once we have removed common factors like the land price.
6.2.2 Estimation

We use all the approaches described above to estimate $\beta$. Firstly, we use a fairly traditional specification with property characteristics, location descriptors and physical attributes of the neighbourhood on the right hand side of an OLS regression. Secondly, we use crimes on non-residential properties as instruments for crimes at or near the property.

Next, we estimate the model of equation (6-3). To do this we need first to estimate $m(\xi | c_i, h_i)$, the sample estimates of the expected values of the independent and dependent variables at each location. These estimates are just locally weighted averages of the neighbouring observations at each data point. Least squares regression using the deviations of the variables from these spatially weighted averages then gives estimates of the linear parameters, as in (6-3). Details of the procedure for computing the locally weighted averages are presented in Appendix A. Following this, we instrument the deviation of residential crimes from their expected values in the surrounding neighbourhood in (6-3) with crimes in other dwellings (in similar local deviation form).

As final checks, we use distance to nearest public house or wine bar as an instrument, then spatial lags of crime, and deviations in spatial lags as instruments.

6.3 Data sources

6.3.1 Crime data

Many police forces in the UK record crime at a geographically localised level. However, it is nearly impossible to obtain this data at the present time in a form that is useful for mapping to other area characteristics and to properties. One exception is the Metropolitan Police Force for London, which has made available to us a unique data set recording property-based crime on an annual basis for the period April 1999 to March 2001. The numbers of property-based crimes are recorded across the London area on
100m grid references. We have five types of crime: *Burglary in a Dwelling*, *Burglary in Other Buildings*, *Criminal Damage to a Dwelling*, *Criminal Damage to Other Buildings*, and *Theft from Shops*. Criminal damage includes graffiti and vandalism, but excludes damage committed in the course of a burglary, which would be recorded under burglary (Home Office (2002)). Unfortunately, it seems that the Metropolitan Police is unable to Postcode other offences accurately. Only 68% of offence locations in all offence categories were post-coded in 1999 (Home Office (2000)). Property based crimes are the easiest to Postcode, though no information was available on what proportion of recorded property crimes actually appear in our data set.

These crime statistics are far from perfect for other reasons. It is well known from comparison of victimisation surveys and recorded crime statistics that the latter understate the true incidence of crime – the so-called *dark figure*. Unsurprisingly, the probability to report a crime varies with the severity of the incidence. More troubling is the fact that the propensity to report a crime varies with the characteristics of the victim, so presumably varies over space too. Apparently, individuals with a ‘police-neutral’ attitude report only 45% of burglaries involving a loss, but without injury or loss of earnings (MacDonald (2001)). More encouragingly, the figure rises to nearly 100% for burglaries involving injury and loss of earnings. No information is available for reporting rates for Criminal Damage. We also know that the police do not record all reported incidents (Home Office (2000)). Some assessment is made about whether a crime really took place, and the incident is recorded or not on that basis.

Ultimately, we will have to live with these data problems. No victimisation or other crime data exists at sufficient density, or at a useful level of geographical dis-aggregation. It is reasonable to assume that the recorded figures in our spatial data set can be treated as an index of the geographical distribution of the most serious incidents of property crime.
6.3.2 Property price data

Our main data source for property transactions is a sample provided by Ekins Surveyors. Ekins is the trading name of Woolwich Surveying Services Ltd, a wholly-owned but independent subsidiary of Woolwich plc operating in the residential and commercial property sectors. In addition to its work with the Woolwich, Ekins receives survey and valuation instructions from over 100 other lending organisations. The full sample contains 10464 properties in the Inner London Area, covering the E, EC1, N, NW, SE, SW, W, and WC Postcode Areas\(^{98}\), surveyed between December 2000 and July 2001. We assign these properties to grid references and match in local area data from various sources using the address Postcodes.

Although the sample has a good range of variables characterising the property, many of these have missing or implausible zero values. Keeping only those observations with non-missing data means a massive reduction in sample size. To avoid this, we retain all properties with non-missing values for a basic property style/type indicator that takes on ten mutually exclusive values. Missing data elements in other characteristics are zero encoded, and a new dummy variable generated to indicate missing elements for each characteristic. In a regression setting, this takes out mean differences between missing and non-missing groups in the data\(^{99}\). Our final sample with matched grid references amounts to around 8100 properties.

Our second property price data source is the Government land registry, which we use as a comparison sample. This data relates to nearly all post-coded property transactions in England and Wales from January 2000 to December 2001. However, it is

\(^{98}\) In the UK, postal addresses are coded hierarchically by Postcode Area, Postcode District, Postcode Sector and full Postcode.

\(^{99}\) We have compared our results with a sample restricted to properties with non-missing observations on number of rooms and property style, and find no important differences.
only available as spatially aggregated data at Postcode-Sector level, by four property
types: Detached, Semi-Detached, Terraced, Flat/Maisonette.

6.3.3 Matching crimes to property locations

Most of the recorded crimes do not match property locations exactly, and it is not
the intention here to measure attacks on specific properties. Rather, we are interested in
obtaining a measure of the expected density of crime in the neighbourhood of a property
– think of a few blocks or streets. For our property level data we calculate the number of
crimes of each residential crime type recorded within a 250m radius of the property, and
the implied density of crimes per kilometre squared. For non-residential crimes, we
double the distance to compensate for the lower density of non-residential properties.
When we match to our Postcode-sector level data, we measure crime density within a
1km radius of the Postcode sector centroid.

6.4 Results and Discussion

6.4.1 Summarising and visualising the data

Table 6-1 summarises the key variables in the property price and crime data. The
top panel summarises the property valuation sample, the bottom the Postcode-sector
based data. Mean crime densities are much lower in the latter, because it covers a much
wider geographical area and because, as we illustrate below, crime rates are lower in the
suburbs. The focus of our work is on recorded crimes in the categories Burglary in a
Dwelling, and Criminal Damage to a Dwelling. Figure 6-1 and Figure 6-2 and illustrate
the geographical distribution of these crimes for the London area, for the period under
study – those crimes recorded from April 1999 to March 2001100.

100 The map is constructed by counting crimes within a 1km radius of points on a 500m grid. The
figures are regression adjusted for estimated household density at each point.
Table 6-1: Summary statistics

<table>
<thead>
<tr>
<th></th>
<th>Mean</th>
<th>s.d.</th>
<th>Min / Max</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Ekins property valuation data</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Property prices, 12/00-07/01 (£000)</td>
<td>235.4</td>
<td>244.8</td>
<td>14 / 4500</td>
<td>8084</td>
</tr>
<tr>
<td>Criminal damage in a dwelling (yr(^{-1}) km(^2))</td>
<td>50.5</td>
<td>30.5</td>
<td>0.63 / 155.8</td>
<td>8084</td>
</tr>
<tr>
<td>Burglary in a dwelling (yr(^{-1}) km(^2))</td>
<td>121.6</td>
<td>79.4</td>
<td>1.2 / 565.3</td>
<td>8084</td>
</tr>
<tr>
<td>Eastings</td>
<td>53091</td>
<td>676</td>
<td>51470 / 54840</td>
<td>8084</td>
</tr>
<tr>
<td>Northing</td>
<td>18064</td>
<td>664.6</td>
<td>16690 / 19590</td>
<td>8084</td>
</tr>
</tbody>
</table>

| **Land Registry Postcode sector data** |      |       |           |     |
| Property prices, 01/99-12/00 (£) | 218.2| 246.4 | 37.5 / 9535 | 5406|
| Criminal damage in a dwelling (yr\(^{-1}\) km\(^2\)) | 37.8 | 31.9  | 0 / 191.3 | 5406|
| Burglary in a dwelling (yr\(^{-1}\) km\(^2\)) | 78.8 | 63.5  | 0 / 370.2 | 5406|
| Eastings                       | 52926| 992   | 50684 / 55200 | 5406|
| Northing                       | 17911| 864   | 15937 / 19820 | 5406|

Crime densities are for totals for two years, April 1999-March 2001

The maps indicate burglary hot-spots north of Islington in North London, and around Brixton in the south. Criminal damage is high in these areas too, but the hot spots look more dispersed. They extend north from Islington up towards Tottenham on the west side of the Lea Valley, east into the East End of London, and on the south side of the River Thames towards Woolwich. Recorded property crime rates are generally low in the Central London area, rise in the inner city areas, and fall away again towards the suburbs.

The black polygon illustrates the envelope of our property valuation data set.

Turning now to the property valuation data, Figure 6-3 shows the distribution of property prices over the London sample area. Comparing the maps for crimes and burglaries, we see that most of the high-density crime areas are in the east, and outside the highest price districts in the west. But this is not the relationship we want to measure. We need to abstract from these broad geographical trends.
Figure 6-1: Incidents of Burglary in a Dwelling per km², April 1999-March 2001
Figure 6-2: Incidents of Criminal Damage to Dwellings per km², April 1999-March 2001
Figure 6-3: Log Property Prices, First 6 Months of 2001

- 255 -
Figure 6-4: Residual Log-Property Price Surface
Figure 6-4 presents an estimated contour plot of the residual property price surface from our models in the London area, smoothed on to a 500m grid. This is an estimate of the function $m(u | c_i, h_i)$ as it appears in (6-1) (for details of how this is constructed see Appendix A). Our semiparametric method in effect removes this spatial variation before estimating the linear parameters in the hedonic model. It is quite clear that no parametric function can be accurately fitted to this price surface. Any fully parametric property price regression that fails to control adequately for this spatial distribution of unobserved price factors will, in principle, provide inconsistent estimates of the model parameters.

6.4.2 Regression results using property level data

Let us begin though with standard OLS log-property price regressions. These results are shown in Table 6-2, Column (1). Explanatory variables are dictated largely by what is available in our property data set. Column (1) includes a quadratic in the distance to Soho, London. This is an approximation to the Central Business District (CBD). The regression includes various measures of population and household density to adjust for the fact that we measure property crimes on a per-unit-area basis. Otherwise, crime density could proxy for housing and population density. For presentational reasons we do not report the coefficients on ten property style dummies. All the estimated parameters on the property characteristics and distance to the CBD seem plausible.

---

101 Including alternative or additional measures – to the City of London, Victoria, or Docklands – made little difference to the key results, and introduced collinearity problems.

102 The results shown are based on the maximal sample with basic property style indicators and at least some information on other characteristics. We zero-encode data elements in the property characteristics set with missing values and generate an additional missing data dummy for each variable.
Table 6-2: London property prices and property crimes, 2001

<table>
<thead>
<tr>
<th></th>
<th>No spatial effects</th>
<th>Smooth spatial effects</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>OLS</td>
<td>IV</td>
</tr>
<tr>
<td>Criminal Damage to Dwellings 100s*</td>
<td>-0.768</td>
<td>-0.422</td>
</tr>
<tr>
<td></td>
<td>(-14.10)</td>
<td>(-9.15)</td>
</tr>
<tr>
<td>Burglary of Dwellings 100s*</td>
<td>0.088</td>
<td>0.014</td>
</tr>
<tr>
<td></td>
<td>(4.03)</td>
<td>(0.71)</td>
</tr>
<tr>
<td>Total rooms in property</td>
<td>0.215</td>
<td>0.187</td>
</tr>
<tr>
<td></td>
<td>(25.40)</td>
<td>(27.67)</td>
</tr>
<tr>
<td>Total floor area (100m²)</td>
<td>3.11e-04</td>
<td>3.6e-04</td>
</tr>
<tr>
<td></td>
<td>(2.86)</td>
<td>(3.72)</td>
</tr>
<tr>
<td>Number of floors</td>
<td>0.044</td>
<td>0.048</td>
</tr>
<tr>
<td></td>
<td>(2.86)</td>
<td>(4.04)</td>
</tr>
<tr>
<td>Age of property</td>
<td>2.1e-03</td>
<td>1.5e-03</td>
</tr>
<tr>
<td></td>
<td>(-8.18)</td>
<td>(7.54)</td>
</tr>
<tr>
<td>Garage</td>
<td>0.084</td>
<td>0.120</td>
</tr>
<tr>
<td></td>
<td>(2.83)</td>
<td>(5.21)</td>
</tr>
<tr>
<td>Flat density (1000s/km²)</td>
<td>-0.022</td>
<td>-9.7e-03</td>
</tr>
<tr>
<td></td>
<td>(-4.92)</td>
<td>(-4.37)</td>
</tr>
<tr>
<td>Household density (1000s/km²)</td>
<td>0.060</td>
<td>0.038</td>
</tr>
<tr>
<td></td>
<td>(5.52)</td>
<td>(5.44)</td>
</tr>
<tr>
<td>Population density (1000s/km²)</td>
<td>-0.028</td>
<td>-0.021</td>
</tr>
<tr>
<td></td>
<td>(-6.45)</td>
<td>(-7.17)</td>
</tr>
<tr>
<td>Distance to Soho (km)</td>
<td>-0.161</td>
<td>-0.128</td>
</tr>
<tr>
<td></td>
<td>(-8.96)</td>
<td>(-5.45)</td>
</tr>
<tr>
<td>Distance to Soho squared</td>
<td>3.6e-03</td>
<td>3.1e-03</td>
</tr>
<tr>
<td></td>
<td>(3.52)</td>
<td>(2.67)</td>
</tr>
<tr>
<td>Km to nearest Underground station</td>
<td>-</td>
<td>-0.031</td>
</tr>
<tr>
<td></td>
<td>(-3.09)</td>
<td>(-3.04)</td>
</tr>
<tr>
<td>Km to nearest council office</td>
<td>-</td>
<td>-0.028</td>
</tr>
<tr>
<td>(town centre)</td>
<td>(-2.86)</td>
<td>(-2.91)</td>
</tr>
<tr>
<td>Km to nearest green space</td>
<td>-</td>
<td>-3.4e-03</td>
</tr>
<tr>
<td></td>
<td>(-3.32)</td>
<td>(-0.12)</td>
</tr>
<tr>
<td>Km to nearest police station</td>
<td>-</td>
<td>-6.6e-03</td>
</tr>
<tr>
<td></td>
<td>(-4.44)</td>
<td>(-0.79)</td>
</tr>
<tr>
<td>Mean rooms in neighbourhood</td>
<td>-</td>
<td>0.086</td>
</tr>
<tr>
<td></td>
<td>(9.07)</td>
<td>(8.47)</td>
</tr>
<tr>
<td>Neighbourhood social housing</td>
<td>-</td>
<td>-0.368</td>
</tr>
<tr>
<td></td>
<td>(-11.33)</td>
<td>(-11.17)</td>
</tr>
<tr>
<td>Local Authority dummies</td>
<td>No</td>
<td>Yes</td>
</tr>
</tbody>
</table>

R2 = 0.400 0.718 0.717 0.556 0.586 0.585

Sample size 8084 8064 8064 8084 8064 8064

P-value test of restrictions No No No No 0.797

|                                  |                  |
|                                  |                  |
| Regressions include ten property style dummies, Local Authority area dummies, and missing data dummies. |
| *-statistics adjusted for clustering on Census Enumeration Districts |
| *Crime units are crimes per year per km²: April 1999 to Mar 2001 |

- 258 -
Focussing now on our crime incidence variables, the first coefficients in Column (1) suggest a highly significant 3.9% decrease in property prices for an additional 5 reported incidents of Criminal Damage per square kilometre per year (10% of the mean, or an expected 1 additional reported incidents per year within a radius of 250m). But, taking the results at face value, domestic burglaries appear to push up property values. Following the discussion in 6.2.1, we assume this implausible (in a causal sense) coefficient reflects the dependence of property crime victimisation on unobserved property, household and neighbourhood characteristics. Higher returns to burglaries in higher-price dwellings, and the higher propensity for better-off households to report crime could bias these estimates. Column (2) introduces more neighbourhood and amenity controls. Immediately, the coefficient on Criminal Damage is halved and the impact of Burglaries vanishes to insignificance. Column (3) instruments crimes on dwellings with the reported incidence of Criminal Damage and Burglaries to other buildings. These IV estimates will be consistent even if victimisation rates depend on the characteristics of dwellings or households. In fact, the IV point estimate is slightly higher than the OLS estimate, but not significantly so using a standard Hausman test of exogeneity ($\chi^2 = 0.583$, p-value = 0.445).

In any property value model we must worry about the impact of unobserved local amenities. Column (4) present results for our semi-parametric smooth spatial effects estimator that allows for unobserved spatially correlated effects on property prices. These are just regression estimates obtained after differencing all the variables from their locally weighted averages. Allowing for these spatial effects in Column (4) immediately gives similar results to the more standard property models in Columns 1-3, even though we

---

103 Lynch and Rasmussen (2001) also find a positive, though insignificant association between property crimes and property prices in a property value regression.
include only the most basic property characteristics. Including a few more neighbourhood characteristics – specifically the neighbourhood proportion in social housing – attenuates the estimated impact of Criminal Damage slightly. Instrumenting with incidents on buildings other than dwellings pushes the coefficients back up. This could be because households in higher-price properties have a higher propensity to report acts of criminal damage. The recorded crime density is endogenous to property prices, and means the non-IV estimates are biased toward finding a positive relationship between incidents and prices. In any case, we do not reject the equality of the Criminal Damage coefficients in Columns 5 and 6 ($p$-value=0.452 in Hausman test).

It is worth discussing the effectiveness of our semi-parametric strategy. That it works is clear from the coefficients on the distance from local amenities. Distance from London Underground stations and distance from Council Offices (a measure of intra-urban centres rather than an amenity in its own right) were both significant at around minus 3% per kilometre in Columns 2 and 3 of Table 6-2. Now they are not. This happens because we are effectively exploiting variation only within neighbourhood groups. Distances to anything but immediately proximate amenities will not matter. Working with differences from local averages eliminates distance-to-amenity-related variation and reduces the need for this type of control.

Adding some more community characteristics into the regression, we find few dramatic changes. Table 6-3 presents the main crime coefficients, plus the coefficient on number of rooms, for three additional specifications. Including spatially weighted averages of local school performance and unauthorised school absences in the regression

---

104 Because the mean distance from an amenity to houses in a radius around a given house $j$ is a consistent estimate of the distance from that amenity to $j$. If we work with the deviation of distance from local mean distance, then only amenities that benefit households because of their immediate proximity will matter in the regression: park and riverside locations perhaps. Taking deviations from the local group mean eliminates the impact of other local factors.
makes very little difference. Controls for ethnicity, education levels and unemployment rates have more of an impact, but as we have discussed before, these residential composition variables are likely to be endogenous. Restricting the sample to observations with non-missing rooms data gives us a lower coefficient on the number of rooms, but increases the measured impact of incidents of criminal damage. The impact of burglary rates remains insignificant.

Table 6-3: Robustness checks

<table>
<thead>
<tr>
<th></th>
<th>Criminal Damage</th>
<th>Burglary of Dwellings</th>
<th>Rooms</th>
</tr>
</thead>
<tbody>
<tr>
<td>Baseline specification, full sample</td>
<td>-0.310</td>
<td>8.8e-03</td>
<td>0.191</td>
</tr>
<tr>
<td></td>
<td>(-5.50)</td>
<td>(0.45)</td>
<td>(28.80)</td>
</tr>
<tr>
<td>Plus performance and absence in nearest primary/secondary schools</td>
<td>-0.294</td>
<td>8.8e-03</td>
<td>0.191</td>
</tr>
<tr>
<td></td>
<td>(-5.03)</td>
<td>(0.45)</td>
<td>(28.63)</td>
</tr>
<tr>
<td>Plus higher-educated, black and Indian, unemployment rate, average age (1991)</td>
<td>-0.240</td>
<td>0.013</td>
<td>0.191</td>
</tr>
<tr>
<td></td>
<td>(-4.33)</td>
<td>(0.69)</td>
<td>(28.57)</td>
</tr>
<tr>
<td>Baseline specification, sample restricted to sample with non-missing rooms data</td>
<td>-0.364</td>
<td>0.016</td>
<td>0.170</td>
</tr>
<tr>
<td></td>
<td>(-5.34)</td>
<td>(0.57)</td>
<td>(22.31)</td>
</tr>
</tbody>
</table>

Regressions are otherwise as in Table 6-2, Column (5)

We shall take Column (5) in Table 6-2 as the best specification. A 5 crimes per year per km² increase (+10% at the sample mean) in the expected density of reported Criminal Damage pushes property prices down by 1.6% \( (\exp(0.310 \times 0.05) - 1) \). This is quite a substantial impact considering that mean number of incidents is 50 with a standard deviation of 30 incidents per year per km². Treating criminal damage as an index of visible crime, we can say that a one-tenth standard deviation increase in crime density leads to a 0.94% decrease in property prices. Interpreting the coefficient as an implicit price in a hedonic function gives us a mean implicit price of around £2200 for a one-tenth of a standard deviation reduction in Criminal Damage incidents the Inner London area. We find no impact from domestic burglary rates, despite carefully attention to identification. For interpretation, read Section 6.4.5.
We have assumed so far that the response of log property prices to the density of crimes is linear. Figure 6-6 in Appendix C provides an informal check. It plots the deviation of log property prices from their locally weighted averages, against the deviation of Criminal Damage 2-year densities from their locally weighted averages. The relationship is predominantly linear.

6.4.3 Regression results using Postcode-sector level data

The primary results above are based on a sample of property transactions for the inner London area only. For comparison purposes, we present similar models using geographically aggregated data from the Government Land Registry. This gives near-universal coverage of property transactions. Property prices are aggregated by the Land Registry to Postcode-sector level for four property types (Flats, Detached, Semi-Detached and Terraced houses) for confidentiality reasons. We use Postcode-sectors falling within the envelope of the crime data we have for the Metropolitan Police Force in London. This gives us slightly wider geographical coverage than our property level data. The 1991 Census provides us with information on the average number of rooms for different property types and on other Postcode-sector characteristics, analogous to the neighbourhood measures used in Table 6-2. Table 6-4 presents these results.

Column (1) is a simple OLS regression. What is clear here is that the Postcode sector average data does a pretty good job of measuring the impact of property crimes on property prices, taking our results of Table 6-2, Column (4) as the comparison model. Overall though, the coefficients in Columns (1) and (2) mis-measure the implicit prices of a room, and social housing relative to our baseline. In Column (3) we work with deviations from locally weighted averages of the variables. This puts the implicit prices back in line with the baseline model. Comparison of Column (3) in Table 6-4, and Column (5) in Table 6-2 suggests that micro-geographically aggregated data is quite acceptable for the hedonic analysis of neighbourhood and property characteristics,
provided that we take the trouble to carefully account for unobserved neighbourhood effects.

Table 6-4: Postcode Sector Data, 2000-2001

<table>
<thead>
<tr>
<th></th>
<th>No spatial effects</th>
<th>Smooth spatial effects</th>
</tr>
</thead>
<tbody>
<tr>
<td>Criminal Damage to Dwellings</td>
<td>-0.262</td>
<td>-0.244</td>
</tr>
<tr>
<td>100s/year/km²</td>
<td>(-5.65)</td>
<td>(-3.86)</td>
</tr>
<tr>
<td>Burglary of Dwellings</td>
<td>0.016</td>
<td>-1.5e-04</td>
</tr>
<tr>
<td>100s/year/km²</td>
<td>(0.59)</td>
<td>(-0.40)</td>
</tr>
<tr>
<td>Mean rooms in neighbourhood</td>
<td>0.311</td>
<td>0.198</td>
</tr>
<tr>
<td></td>
<td>(25.59)</td>
<td>(11.79)</td>
</tr>
<tr>
<td>Neighbourhood social housing</td>
<td>-0.214</td>
<td>-0.327</td>
</tr>
<tr>
<td></td>
<td>(-4.38)</td>
<td>(-6.18)</td>
</tr>
<tr>
<td>Flat density (1000s/km²)</td>
<td>-5.4e-03</td>
<td>5.4e-03</td>
</tr>
<tr>
<td></td>
<td>(-1.12)</td>
<td>(1.17)</td>
</tr>
<tr>
<td>Household density (1000s/km²)</td>
<td>0.084</td>
<td>0.013</td>
</tr>
<tr>
<td></td>
<td>(6.58)</td>
<td>(0.95)</td>
</tr>
<tr>
<td>Population density (1000s/ km²)</td>
<td>-0.052</td>
<td>-0.021</td>
</tr>
<tr>
<td></td>
<td>(-8.87)</td>
<td>(-3.46)</td>
</tr>
<tr>
<td>Distance to Soho</td>
<td>-0.133</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>(-19.88)</td>
<td></td>
</tr>
<tr>
<td>Distance to Soho squared</td>
<td>3.0e-03</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>(14.00)</td>
<td></td>
</tr>
<tr>
<td>Local Authority dummies</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>Sample size</td>
<td>5406</td>
<td>5516</td>
</tr>
</tbody>
</table>

Regressions include three property style dummies, one year dummy, plus constant term. Robust t-statistics in parentheses.

6.4.4 Alternative instruments

In Section 6.2.1 we suggested using spatial lags of the crime density as instruments for neighbourhood crime density, on the assumption that averaged crime rates at some radius from a property or neighbourhood should be unaffected by the characteristics of the property or neighbourhood. Using this strategy, we still fail to find any impact from Burglaries in Dwellings on property prices. This reinforces the impression that burglary rates really have no causal impact. The results are in the top panel of Table 6-5. Rather than attenuating the impact of Criminal Damage, this strategy gives us bigger negative

---

105 In practice we use the locally weighted averages computed for each observation as in Section 6.2.2, but excluding any data points within a radius of 1km of the observation.
coefficients: -0.664 (-4.64) using data in levels; -0.680 (-4.02) using the data in deviations from locally weighted averages. As we discussed before in Section 6.4.2, this may be because the higher propensity of occupants of higher-price dwellings to report crime attenuates the non-IV coefficient. But it may also be because average crime density in the wider geographical area suffers from less measurement error and noise than the locally computed crime densities. Instrumenting corrects for errors-in-the-variables-induced attenuation.

Table 6-5: Alternative instruments for criminal damage

<table>
<thead>
<tr>
<th>Instrumentation</th>
<th>Criminal Damage</th>
</tr>
</thead>
<tbody>
<tr>
<td>No spatial effects, spatial lags of crime as instruments</td>
<td>-0.664 (-4.64)</td>
</tr>
<tr>
<td>Spatial effects, second spatial lags of crime as instruments</td>
<td>-0.680 (-4.02)</td>
</tr>
<tr>
<td>No spatial effects, distance to pub as instruments (cubic)(^1)</td>
<td>-0.582 (-3.17)</td>
</tr>
<tr>
<td>Spatial effects, distance to pub as instruments (cubic)</td>
<td>-0.472 (-1.92)</td>
</tr>
</tbody>
</table>

Regressions are otherwise as in Table 6-2
IV regressions using pub distance as instruments include public house density as additional regressor in property-price equation, to allow for amenity effects

Note: 1. The instruments in this model fail the Sargan test for the validity of the overidentification (p-value=0.027). All others pass the test at a p-value of 0.200 or greater.

Consideration of the possible cultural factors underlying graffiti, vandalism and other forms of criminal damage suggest another plausible instrument. Alcohol consumption is an associated factor in many types of crime, although the lack of official statistics for the UK makes it difficult to quantify the link (Deehan (1999)). A study in one town in England found that 88% of people arrested for acts of criminal damage, over a period of five months, had been drinking in the four hours prior to the incident (Jeffs and Saunders (1983)). Official statistics for local prisons in the United States indicate that 33% inmates convicted for a property crime, and some 56% of inmates convicted for a...
public order offence, had been drinking prior to the offence. Of those inmates, around three-quarters had a Blood Alcohol Content in excess of 0.10g/dl at the time of the offence (Bureau of Justice Statistics (1998)). Although the link between alcohol consumption and crime is not necessarily directly causal, alcohol is often a contributory factor in violent crimes and acts of public disorder. This may be because alcohol encourages aggression, induces psychotic states, or decreases inhibitions. Or it may be that some certain social environments encourage both excessive drinking and disorderly or criminal activity (Deehan (1999), Bottoms and Wiles (1997)). In any case, a link between the location of crimes and the location of licensed premises, and the time of offences and the end of licensing hours is widely recognised (Bottoms and Wiles (1997)).

With these considerations in mind, we would expect the incidence of property crime in our London data to be higher at locations near licensed premises. Indeed this is true. Regressing the criminal damage density at each property location on a 3rd-order polynomial in distance from the nearest public house or wine bar, we find significant negative impacts (F(3,138)=6.43). For the average property, criminal damage density at a property decreases at the rate of 3.5 crimes per km² per year as distance to the nearest pub increases. In the lower panel of Table 6-5 we employ distance to the nearest licensed premises, and its polynomials as instruments for criminal damage in our property price equation. Again, this instrumental-variables strategy increases the estimated negative impact of criminal damage on property values, although the results are not far out of line with the IV estimates in Table 6-2. The use of distance to nearest licensed premises as an instrument assumes that there is no direct amenity value from living close to a pub. This is questionable, though our instruments pass the appropriate tests once we allow for

106 Data on pub locations is from the web edition of the Thomson Local Directory, http://www.infospace.com/uk.thomw/. This result is based on a regression in deviations-from-spatial-means form, with additional controls as in Table 6-2.
general spatial effects. We include local pub density as an additional regressor, under the assumption that accessibility to a variety of drinking establishments if likely to have a higher amenity value than close proximity to the nearest. This does have a positive impact on prices – an additional 10 pubs or wine bars per km$^2$ increasing prices by 2.8%.

6.4.5 Interpretation and discussion

Burglaries do not seem to influence property prices, but Criminal Damage incidents do. This is, at first, quite surprising. True, home-owners can take preventative action against burglars (alarm systems, barriers) but may not be able to prevent damage to property. But we should consider to what extent our estimated impact of Criminal Damage to Dwellings picks up the cost associated with a high incidence of unobserved crimes – violent crime, robbery, vehicle crime for example. Our data is slightly limited by the lack of information on crime in other categories. Some unobserved crime categories are cause for concern, because the estimates of the economic costs of these types of crime are high. Brand and Price (2000) estimate that average cost associated with an act of violence against the person is £19000 with serious wounding carrying total costs of £130000. For robbery the figure is £9700 per incident. Clearly, we can expect the costs associated with increased risk of attack associated with a high persistent high local incidence of robbery or violent crime to be capitalised in property values. On the other hand, incidents of assault and robbery may be more important in individual choices about where and when to walk the streets. The location of property crimes is more directly related to choice of residential location.

Unfortunately there is not much data available that allows us to infer anything about the relationship between rates of crimes in different offence categories at a localised geographical level in urban areas$^{107}$. We can do this at a much broader scale.

$^{107}$ Crime data is collected at Local Authority Level, but not for the Criminal Damage category.
geographical level using recorded crime at the Police Force Area level for England and Wales. Police Force Areas correspond to Counties, with a few exceptions. Whether the cross sectional relationship tells us much about the relationships between types of offences at the neighbourhood level is pretty doubtful. Nevertheless, the relationships between year-to-year changes in crime rates within Police Force Areas will be informative about the links between different types of criminal activity. Table 6-6 reports the coefficients obtained by regressing first differences of various log crime rates (crimes per person) within 43 Police Force Areas on the first differences in log crime rates for Burglary and Criminal Damage. Year dummies to take out general trends.

Table 6-6: Association between year on year changes in Police Force Area crime rates, 1997-1999

<table>
<thead>
<tr>
<th></th>
<th>Violent</th>
<th>Theft</th>
<th>Robbery</th>
<th>Sexual</th>
</tr>
</thead>
<tbody>
<tr>
<td>Criminal Damage</td>
<td>0.237</td>
<td>-0.011</td>
<td>0.417*</td>
<td>0.529*</td>
</tr>
<tr>
<td></td>
<td>(1.46)</td>
<td>(-0.11)</td>
<td>(2.74)</td>
<td>(1.98)</td>
</tr>
<tr>
<td>Burglary</td>
<td>0.232</td>
<td>0.781*</td>
<td>0.500*</td>
<td>-0.159</td>
</tr>
<tr>
<td></td>
<td>(1.07)</td>
<td>(4.74)</td>
<td>(2.22)</td>
<td>(-0.54)</td>
</tr>
<tr>
<td>F-test p-value</td>
<td>0.205</td>
<td>0.000</td>
<td>0.004</td>
<td>0.086</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Vehicle</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Criminal damage</td>
<td>0.376*</td>
<td>0.110</td>
<td>0.215*</td>
<td>0.054</td>
</tr>
<tr>
<td></td>
<td>(4.17)</td>
<td>(0.78)</td>
<td>(3.39)</td>
<td>(0.74)</td>
</tr>
<tr>
<td>Burglary</td>
<td>0.373*</td>
<td>0.245</td>
<td>0.452*</td>
<td>0.598*</td>
</tr>
<tr>
<td></td>
<td>(2.99)</td>
<td>(1.74)</td>
<td>(6.86)</td>
<td>(4.55)</td>
</tr>
<tr>
<td>F-test p-value</td>
<td>0.000</td>
<td>0.155</td>
<td>0.000</td>
<td>0.000</td>
</tr>
</tbody>
</table>

Table shows coefficients and standard errors from regression of first differences of various log crime rates for police force areas, on first differences of log criminal damage and burglary rates.
Regressions include year dummies
Sample size = 172

Crimes in nearly all the offence categories are positively correlated with both Criminal Damage and Burglary, with elasticities of between 0.1 and 0.5. Looking at the joint significance of the coefficients, it seems that only in the case of Robbery and Vehicle Crime are these jointly significantly different from zero. There is a relationship too between Sexual Crimes and Criminal Damage. The number of total crimes in other categories (excluding Criminal Damage, Burglary and Fraud) is quite strongly associated
with Criminal Damage and Burglary. It seems that similar factors influence criminal activity in the Robbery, Sexual Crimes, Vehicle Crimes and Criminal Damage categories, or that the same criminals are active in these categories. But once we exclude Vehicle Crimes from Total Crime (Table 6-6, Column (4), bottom panel), the relationship between Total Crime and Criminal Damage disappears.

We also have some geographically disaggregated crime data for the Derbyshire Police Force Area. This police force records crimes at Census Ward level (around 3000 households). This is not a metropolitan area, but contains a mixture of urban and rural territories. Here we find few significant associations between the annual change in the number of recorded incidents Criminal Damage, Burglary and other crime types. Again, Criminal Damage is significantly associated with Vehicle Crimes (thefts from and of vehicles) and Burglary, but with nothing else.

Figure 6-5: Crime trends in Metropolitan Police Force Area, 1993-2001

Changes in counting rules can make comparison between pre and post 1999 figures misleading. Figures are adjusted for overall effect on offence groups, but the Theft and Handling group cannot be corrected accurately. All vehicle-related crimes (including some criminal damage to vehicles) have been deducted from the Theft and Handling category post January 1998. There were also minor geographical changes to the Metropolitan Police Force boundary in 2000.
The crime trends for the Metropolitan Police Force Area Figure 6-5 also suggest little association between criminal activity in the Criminal Damage category and what we would perceive as serious urban crimes (Violent Crime, Robbery). Whilst recorded crimes in the Burglary, Criminal Damage and Theft categories have been on a general trend down in the last decade, Violent Crime and Robbery have been on the increase.

What then are we to make of our results? On the face of it the impact of Criminal Damage on property prices seems high relative to estimates of the direct, physical and emotional costs associated with Criminal Damage itself. Average costs per incident to the household experiencing it are in the order of £510 (Brand and Price (2000)). In comparison, our estimates say that a household is willing to pay something like this to avoid 14 incidents of criminal damage in a square kilometre in their neighbourhood. But a square kilometre in Inner London holds, on average, some 2800 households. Based on the average value of an incident of criminal damage to the household, these 14 incidents should have an expected cost per household in the order of \((14 \div 2800) \times £510 = £2.55\). By the same calculation, if we translate the impact of an increase in the density of crime into an increase in the probability of victimisation, our results suggest that the cost of victimisation is over £100000 for an incident of criminal damage. It is quite clear that if incidents of criminal damage affect property prices, than it is for reasons other than the expected costs of the incidents themselves!

\[\text{average cost of 14 crimes in one year in one km}^2 = \exp\{(0.14 \times 0.310 - 1) \times £235000 \times 0.05 = £522, assuming the coefficient on crimes in 100s per km}^2\text{ per year is 0.31, mean property price = £235000, discount rate = 0.05}\]

\[\text{We can translate the impact of crimes per km}^2\text{ into crimes per household by multiplying by the population density, and evaluating a marginal change in crimes per km}^2. \text{ The average cost of one crime in one year in one km}^2 = £37 (same assumptions as above). So average cost of a crime per household in an average area of 1km}^2 \text{ containing 2800 households is } 2800 \times £37 = £104000.\]
A more likely explanation is that incidents of vandalism and criminal damage impact on property prices because they induce fear of crime. Graffiti, for example, comes out as one of the few neighbourhood factors which is consistently significantly correlated with several measures of fear of crime (Killias and Clerici (2000)). And yet Criminal Damage rates do not seem highly correlated with other types of crime, except Burglary – which we have controlled for in our property price regressions – and Vehicle Crime – which again imposes relatively low direct and psychological costs. But Criminal Damage is clearly perceived as a problem by individuals. In the 2000 British Crime Survey, 32% of respondents agreed that vandalism was a ‘very/fairly big problem’ (Home Office (2001a)), although only 10% of these considered it had a negative impact on their quality of life. Nevertheless, in the same study, between 33% and 50% of respondents in owner-occupier neighbourhoods consider that disorder in general has a negative impact on quality of life and one in five respondents in affluent owner-occupier neighbourhoods perceive high levels of disorder.

Perhaps the most plausible interpretation of our results is that incoming residents perceive Criminal Damage in the neighbourhood as signalling higher crime in the area, or deteriorating neighbourhood in general. In essence, what we are finding relates to neighbourhood effects of the type described by Wilson and Kelling’s Broken Window Syndrome (Wilson and Kelling (1982)). According to this hypothesis – popular in the environmental criminology literature and with advocates of neighbourhood cleanup campaigns10 – unrepaired damage to property in the neighbourhood encourages further vandalism, perceptions of community disorganisation, upward spiralling crime rates and downward spiralling neighbourhood status. If vandalism and graffiti are seen as predictors of neighbourhood decline and precursors of escalating crime rates, then it is

10 Almost all citations on the web are on community web-sites in the US, encouraging neighbours to clean up their lots.
not surprising that we see them impacting in property prices. Nevertheless, our evidence is that these disorder-related crimes are weakly to moderately associated with more serious crimes, suggesting – like Sampson and Raudenbush (1999) – that the disorder-crime link is not necessarily causal. Physical disorder like graffiti and vandalism may be symptomatic of deeper disruptions in social cohesion and community expectations – what Sampson and Raudenbush call ‘collective-efficacy’.

We should also recognise that vandalism, graffiti and other forms of criminal damage are some of the most visible urban crimes. Uncleaned graffiti and unrepaired damage in the environment is hard to conceal from prospective house purchasers. Whilst sellers may have private information about local incidents of other crimes – by personal victimisation, word of mouth or Neighbourhood Watch newsletters – this information is most likely unavailable prospective home-buyers. In London, information on neighbourhood crime rates is not readily available to the general public. This asymmetry in information means that the hedonic price function does not correctly reveal preferences over most types of crime. Hard-to-observe crimes will have a weak impact on property prices.

6.5 Conclusion

We have estimated the impact of recorded crimes in the Criminal Damage to Dwellings and Burglary in Dwellings categories on property prices in the London area, paying careful attention to identification issues. Crimes in the first category – including vandalism, graffiti and arson – have a significant negative impact on prices. Burglaries have no measurable impact on prices, even after allowing for the potential dependence of burglary rates on unobserved property characteristics. A one-tenth standard deviation increase in the recorded density of incidents of criminal damage has a capitalised cost of just under 1% of property values, or £2200 on the average Inner London property in our sample 2001. In annual terms, this is around £110 per year per household. Aggregating up
to some £340 million per year for all 3.1 million households in the London region. This is a huge impact. By comparison the Safer Communities Initiative offers CDRPs in the London region\textsuperscript{111} a total of 3.7 million for 2002/2003 Home Office (2001b), or around £1.40 per household.

It is, on the face of it, surprising that prices respond more to acts of criminal damage than to burglaries given the apparent physical and emotional costs. The explanation we offer here is that vandalism and graffiti are important factors motivating fear of crime in the community, even though the evidence here suggests that these types of crimes are not strongly correlated with incidents of a more serious nature. More generally, graffiti and vandalism may be taken as signals or symptoms of community instability, disorder, lack of social cohesion and neighbourhood deterioration in general.

\textsuperscript{111} The 1998 Crime and Disorder Act established partnerships between the police, local authorities, probation service, health authorities, the voluntary sector, and local residents and businesses.
6.6 Appendix A: Computing locally weighted averages

This Appendix describes how we compute the locally weighted averages of each variable $\xi$ and so estimate $m(\xi | c_i, h_i)$. We define:

\[
\hat{m}(\xi | c_i, h_i) = \left( \sum_{j=1}^{n} \xi_j \phi[d_{ij} h^{-1}] \right) \cdot \left( \sum_{j=1}^{n} \phi[d_{ij} h^{-1}] \right)^{-1}
\]

(6-5)

\[
d_{ij}^2 = (c_i - c_j)(c_i - c_j) \text{if } d_{ij} < k
\]

\[= \infty \text{ otherwise}
\]

(6-6)

\[
h_i = s.d.(d_{ij}) \text{if } d_{ij} < k
\]

(6-7)

where $\phi()$ is the standard normal density function. This means we are using a Gaussian kernel or distance decay function to weight neighbouring observations. Parameter $k$ sets the maximum distance to the neighbouring observations that will be used to compute these local weighted averages. Our estimator of $m(\xi | c_i, h_i)$ is thus a kernel-weighted nearest neighbour smoother. This is a variation on the Smooth Spatial Effects Estimator of Gibbons and Machin (2001).

Note that the choice of $k$ determines the degree of smoothing. This defines how wide the neighbourhood is over which we compute the locally weighted averages. A higher value of $k$ implies generates a longer spatial lag. The choice of $k$ is somewhat arbitrary, but was found to make little difference in practice over a moderate range. Our baseline choice of $k$ is such that the spatially weighted mean explains around one third of the variation in property prices, as measured by the $R^2$ in a regression of $\ln P_i$ on $m(\ln P_i | c_i, h_i)$.
6.7 Appendix B: Constructing the land price surface

Figure 6-4 illustrates an estimated residual land price surface for the Inner London area. This is an estimate of $m(u | c_g, h)$, the expected value of the residuals from the property price equation at map grid points $c_g$, with a fixed smoothing parameter $h$. To obtain this map, we first estimate the model in equation (6-3) to obtain estimates of the linear parameters $\beta, \gamma$. Note now that

$$m(u | c_g, h) = E[\ln P_i - \beta'x_i - \gamma'z_i | c_g]$$

(6-8)

So we then compute the residuals $\ln P_i - \beta'x_i - \gamma'z_i$. Next we calculate the locally weighted averages of these residuals within 2.5km of each map grid point, using a Gaussian distance decay function (or kernel).

6.8 Appendix C: Linearity of the crime-price function

Figure 6-6: Association between local deviations in Criminal Damage density and local deviations in log property prices

Figure shows kernel regression of log prices (vertical axis) on incidents of Criminal Damage to Dwellings in 100s per km$^2$ (horizontal axis). Variables are in the form of deviations from local means. Kernel is Epanechnikov, bandwidth 0.15
7 Summary and Conclusion

7.1 What have we learnt?

A vast theoretical literature documents the relationship between housing bids, residential sorting and community-based public goods. In the US there has been a substantial effort to verify and measure these effects empirically (see Ross and Yinger (1999)). A first contribution of the pieces of research presented in this Thesis has been to extend this measurement of the value of community and local school quality to the British case. There is also substantial evidence from the US on the link between community quality and individual educational and labour market outcomes, and again, the research in this Thesis extends this work to Britain. In both cases, we have employed some novel strategies to the task of empirically identifying the parameters of interest. Each Chapter has presented a variety of econometric evidence on the value of neighbourhood communities, focussing on measuring the value of intangible neighbourhood commodities that are embedded in the populations of local communities – specifically local human capital, neighbourhood school performance and crime-related community disorder. We moved from direct measurement of the impacts of neighbourhood human capital on educational acquisition in Chapters 2 and 3, to more indirect approaches in Chapters 4-6, based on revealed preference in property markets. No claim is made here that this is a complete picture of the role that neighbourhoods play, and no claim is made in terms of new theoretical insights. But we have touched on some of the key issues related to the role of community in human capital formation and on choice of residential location, and revealed previously unseen aspects of the British housing and educational systems. An important contribution of the studies presented in the Thesis (Chapters 4-6) is the application of semi-parametric methods to the empirical analysis of housing
markets. We used these methods to allow for general effects of location on the housing market, by empirically modelling price surfaces that are highly non-linear over geographical space. Overall, we have compelling evidence that neighbourhood community matters to individuals and households. We have found evidence for inter-neighbour spillovers in terms of human capital accumulation at individual and primary school level. These relationships are not merely an artefact of the residential sorting by income and educational class. Moreover, households react strongly to community composition, primary school performance and crime incidence in their demands for residential location.

7.2 A review of the results

We started in Chapter 2 by looking directly at the impacts that neighbourhoods had on a cohort who grew up in Britain in the 1960s and 1970s. Children brought up in the same neighbourhood do end up with similar educational attainments. This association is, however, quite weak. At most, the correlation between an individual’s years in education and the average education of others who lived in the same Ward in the 1970s is around 0.16. And this similarity in educational attainments is, in part, due to the children in the same neighbourhood having similar parents. Allowing for similarities in parental education alone halves this inter-neighbour correlation. An interpretation of these correlation coefficients is that a child could expect to increase his or her time in education by between 3.2 to 11.8 weeks if brought up amongst children destined to stay in education for 1.4 years longer than average.

Another way of looking at this is to rank origin neighbourhoods in terms of the proportion of adults with A-levels and higher qualifications. We showed that children from the top ten-percent of neighbourhoods, ranked in these terms, were between five and seven percentage points more likely to get A-levels themselves than children with similar family backgrounds living in neighbourhoods ranked in the bottom 10%. Children from
educationally advantaged communities are also less likely to end up with no qualifications. One implication of this relationship is that a child brought up in a neighbourhood ranked at the bottom of the educational hierarchy would need parents educated to something like degree level to give him or her the same educational opportunities as another child from an average background.

Importantly, these effects do not operate purely through the quality of local schooling or through association with peer-group pupils from better backgrounds attending the same school. Residential neighbourhood has an impact over and above anything related to local secondary school performance, implying role-model and expectations-related influences.

Many studies have investigated the link between parental social and economic status and that of their children. That a link exists is generally undisputed, although the strength of the link and its causes are open to question, and subject to the usual 'nature-nurture' arguments. Part of this link can be attributed to differences in the status of neighbourhoods inhabited by families at different points in a ranking of social and economic advantages. Educated families are more likely to live in educated neighbourhoods. If educated neighbourhoods matter, then having educated parents provides benefits over and above those advantages arising through parents' direct influence. Having said that, we found that neighbourhoods are not major contributors to social immobility. At most one tenth of the association between our own education and that of our parents is attributable to the type of neighbourhood in which they chose, or could afford, to live.

Our results relate to children raised in the 1960s and 1970s. Are they still relevant today? We do not have the data to test this fully, but we investigate the issue by comparison of broader area effects on this and a later cohort (teenagers in the 1980s). There is no evidence here that the link between spatial location and educational attainment declined between the 1970s and 1980s. One reason to expect a change in the
importance of childhood neighbourhood is if educational and income deprivation has become progressively more concentrated in particular neighbourhoods. But again, there is no evidence for increasing spatial concentration of educational disadvantage in Census wards from 1971 to 1991.

Measurement of neighbourhood effects is about measurement of social interactions in spatial groups. In Chapter 3, we turned to primary schools to look further at neighbourhood effects and to measure the impact that area has on children’s attainments. Primary schools are interesting because they serve highly localised communities, so the link between school neighbourhood and school performance is tight. We looked for geographical clustering of primary school performance as evidence of social interactions within neighbourhoods.

Our analysis showed that good schools do tend to be located near other good schools, and bad schools near other bad schools, even allowing for differences in intake and neighbourhood composition. And better primary schools tend to be clustered in urban areas – especially in the North West and London. Schools in a swathe across rural England from the Severn to the Wash seem to do quite badly on standard performance measures. This contrasts with the common impression of under-performing inner city schools – an impression created by the background disadvantage of pupils in these areas. What this suggests is that the policy idea of ‘Beacon’ schools (Department for Education and Employment (2001)) can work – schools and their pupils can benefit from the presence of successful schools in the neighbourhood.

We also asked to what extent neighbourhood influenced primary school performance measures. As we might expect, there are strong relationships between pupil attainments and the characteristics of the catchment area in which a school is located. Consider pupils attending schools located in the top third of neighbourhoods (measured by a composite index of neighbourhood status). These children are at least twenty percentage points more likely to achieve the target standard in their age-11 tests than their
counterparts attending schools in the bottom third of neighbourhoods. This relationship is not due to better-off parents choosing to live near better schools.

Unpacking the separate effects of neighbourhood characteristics on school performance is tricky. Different dimensions of neighbourhood quality – education, incomes, housing – are closely interrelated. Our evidence is that, although school performance tracks average incomes in the catchment area, other wealth related characteristics – education, housing quality and population age – matter over and above current incomes. Key school intake characteristics – free school meal entitlement, ethnicity and special educational needs entitlement – have a strong influence on performance.

It is natural to ask whether more school resources can overcome neighbourhood disadvantage. We tackled this problem by examining whether additional key resources – more teachers per pupil and higher Local Education Authority expenditures – affect school performance scores. We found that more teachers and more money do help, but the impact is dominated by the influence of neighbourhoods and intake. Inference from the existing pattern of resources and performance suggests that an additional 18,000 primary school teachers would only bump up the proportion of children reaching target standards by 1.3 percentage points (holding teacher quality constant).

All of this does not mean that schools are powerless in the face of neighbourhood disadvantage. It is well known that some schools in disadvantaged areas do well, but that the foundations of these successes are hard to pin down. Here too, we showed that idiosyncratic things that we do not observe about a school its pupils and its environment account for two-fifths of the variation in school performance across the country. Some of these factors lead to persistent differences over time between schools in similar circumstances. Finding these factors and acting on them is clearly an important goal of research and policy. What our results do suggest is that policy directed at raising and
equalising school standards is unlikely to succeed without attention to the substantial geographical basis of these performance differentials.

We then turned to measuring the value of neighbourhoods as perceived by incoming residents in the housing market. We looked first at neighbourhood human capital stocks. The main results from Chapter 4 show that property prices increase by one percent in the South and East of England, and by two percent in Wales, the West and North of England, for each one percentage point shift in the proportion of higher-educated residents. Because mean education levels differ across regions, this amounts to a 0.24 percent increase in prices for a one percent relative change in the education of an average community in any region. This is equivalent to about £156 on 1995 national mean prices. House prices move by 0.52 percent for each one percent change in local mean incomes. Using these figures, and taking into account the empirical relationship between individual earnings and education, we deduce that education is valued as a community commodity for reasons other than its impact on incomes. We find further, that education in adult residents matters over and above other community characteristics like unemployment rates, sick rates, lone-parenthood, age, crime rates and local primary school quality.

Importantly, households pay more for community educational improvements in areas where there are more owner-occupier children. Also, the proportion of home-owners with children is higher in areas where there are fewer social tenants. One interpretation is that families value community educational status because of its influence on children’s development and well-being.

We considered the impact of neighbourhood human capital stocks on prices as a measure of the local social benefits – or community benefits – of education. The monetary value is at least as large as the estimates of the average private returns to education – the increment to earnings arising from educational improvement – which dominate the literature. Given the size of these effects, the community and other wider
benefits of education deserve further analysis. Focussing only on the private returns risks seriously understating the value of education to society, and any policy decisions based on these returns alone may result in sub-optimal provision of educational services.

In Chapter 5, we looked at the value of a specific aspect of neighbourhood communities – primary schools – whilst allowing for more general effects from social tenants and unobserved neighbourhood factors, as in Chapter 4. The amounts households seem to pay are substantial: in London, residents pay around £13500 for an increase of ten percentage points – say from 65% to 75% – in the proportion of pupils succeeding at age 11 tests.

Our main results show that households pay up to 6.9 percent on property prices for each 10 percentage point improvement in the performance score of local primary schools. In terms of a monetary valuation, these figures tell us that the average household is willing to pay around £90 per pupil, per year, for an improvement in school standards which sustains a one percentage point improvement in the performance scores.

We note that a twenty-five percentage point performance advantage in London has a lifetime value of about £35000 in year-2000 prices – roughly equivalent to the cost of a private-sector primary education for one child for eight years. In annual terms, the state-sector still works out cheaper: around £2500 in additional mortgage payments compared to over £6000 in prep-school fees.

This sensitivity of property prices to local school quality implies that there is back-door selection into better performing schools on the basis of family income. This is clearly inequitable, and a barrier to equality of opportunity. The outcome may also be inefficient and un-meritocratic, in that some pupils with the capacities to take advantage of better primary education are excluded on the basis of income or borrowing constraints. If these issues are of concern, then policy makers may need to consider whether residential proximity should take such an important role in school’s over-subscription criteria.
Finally, Chapter 6 analysed the effects of neighbourhood crime rates on property values, using similar empirical tools to Chapters 4 and 5. We find that higher crime rates in an area push prices down. But the property-based crimes that have the greatest impacts are those in the Criminal Damage to a Dwelling category – which includes graffiti, vandalism and arson. Burglary has no discernible effect once we properly control for general location effects and correct for the higher returns to burglary in higher house value neighbourhoods. The magnitudes of these effects imply that households are willing to pay much more to avoid areas with high rates of criminal damage than is rational, given the low probability of victimisation and the low direct and emotional costs of this type of incident. A one-tenth standard deviation increase in the recorded density of incidents of criminal damage has a capitalised cost of just under 1% of property values, or £2200 on the average Inner London property in our sample 2001. This would put the costs of victimisation at over £100000 for an incident of criminal damage.

Why do residents care so much about high levels of relatively trivial crimes? One possibility is that our results are just picking up effects more serious crimes that we do not observe. Or visible criminal damage may induce fear of crime if it is read as an indicator of high local crime levels. But our evidence is that serious crime tends to be less correlated with criminal damage than with burglary – for which we find no property price effect. More likely, we argue, is that neighbourhood criminal damage measures community disorder or signals some deeper community discord that households care to avoid if they can.

An interesting question arises from the analysis of Chapters 3 and Chapters 5. Community characteristics influence school performance and school performance influences community characteristics. Does this imply increasing polarisation of school performance? Consider property prices as a simple index of community wealth. We can infer the relationship between community wealth, as indexed by property values, and primary school performance by running a school production function regression of the
type in Chapter 3, equation (3-3). Average house size and second lags of school performance provide suitable instruments for lagged property values and lagged school performance. If we run this regression for 1998 primary school performance, we find that a one percent exogenous increase in property values is associated with a 0.22 percentage point increase in Key Stage 2 target success rates\textsuperscript{112}. We have shown in Chapter 5 that a one percentage point exogenous increase in school performance, measured by Key Stage 2 success rates, pushes up log-property prices by 0.0067. This in turn could generate a 0.15 percentage point improvement in subsequent school performance and a 0.0010 increase in log-prices. These effects are diminishing over time, so there is no tendency for this process to generate ever-increasing polarization of school performance.

Consider a simple linear model of school performance ($s$) and log-house prices ($p$), of the form, where $\tilde{s}$ and $\tilde{p}$ are independent stochastic error terms:

$$s = \beta p_{-1} + \tilde{s} \quad (7-1)$$

$$p_{+1} = \alpha s + \tilde{p} \quad (7-2)$$

Using the parameters estimated above, the equilibrium variances of log-prices ($p$) and school performance ($s$) are:

$$Var(s) = 0.049\sigma_\tilde{s}^2 + 1.022\sigma_\tilde{f}^2 \quad (7-3)$$

$$Var(p) = 1.022\sigma_\tilde{p}^2 + 0.459\sigma_\tilde{f}^2 \quad (7-4)$$

\textsuperscript{112} This means carrying out an IV regression of the proportion reaching Key Stage 2, level 4, in 1998 on lagged school performance, lagged local log-property prices, school-proportions with special educational needs, non-white ethnic groups, school type and age-range and Local Education Authority dummy variables. Second lag of school performance and mean house size provide instruments for house prices and lagged school performance. The coefficient on log prices is 0.037 (0.011), and the performance autocorrelation is 0.832 (0.021), so the overall effect from log-prices (taking account the lagged dependent variable) is $0.037/(1-0.832) = 0.220$. 

- 283 -
where $\sigma_p^2$ and $\sigma_s^2$ are the variances of the unobserved determinants of property prices and school performance respectively (see the Appendix). From the data (for 1998), we find that $Var(p)/Var(s) = 7.852$, so it follows from (7-3) and (7-4) that $\sigma_s^2 / \sigma_p^2 = 0.084$. Substituting this back, we find that exogenous variation in school performance contributes something like 3.6% to the variance of log-property prices. Variation in exogenous community characteristics associated with log-property prices drive around 36.3% of the overall variation in primary school performance.

7.3 Scope for future research

Whilst there is enormous scope for more work in this field, opportunities for empirical research are currently limited by lack of suitable recent data for the British case. Any future work will obviously benefit from the availability of the 2001 Census data in providing an up-to-date snapshot of neighbourhood conditions, and as a basis for investigating the impacts and causes of neighbourhood change since 1991. Unfortunately confidentiality issues largely preclude the matching of individual micro-data to the Census, since most providers are unwilling or unable to link individuals to residential addresses. So, school-based data will continue to provide, perhaps, the best opportunity for further investigation of the relationship between human capital accumulation and neighbourhoods, since the location of schools is easily ascertained. Many of the interesting issues surrounding neighbourhood effects can be traced out at school level, particularly the interplay between neighbourhood change, school composition and performance. In this context, the recent availability of pupil-level data for England in the Department of Education and Skills' Pupil Level Annual Census offers interesting

---

113 We do not claim a causal link from prices to performance – prices proxy for community characteristics like education and incomes that are school performance-enhancing.
prospects for differentiating between school and pupil-level social interactions in educational attainments. Moreover, school initiatives with an overtly local, spatial dimension – such those within the Department of Education and Skills’ *Excellence in Cities* programme – offer opportunities for the exploration of the impacts of neighbourhood-related educational policy interventions. Certainly, there is scope for more work on the micro-spatial aspects of the housing market, and its relationship with local public amenities, school provision and community characteristics. These are important issues, both in terms of understanding spatial patterns in the housing market and for understanding the role of neighbourhood attributes and local public goods in consumer demand. Empirical approaches developed from the parametric restrictions implied by equilibrium sorting models (Epple and Sieg (1999)) may provide a way forward where non-parametric identification is impossible. A fuller theoretical and empirical understanding of these issues will offer deeper insights into the role that household demand for neighbourhood quality plays in the preservation of inequalities of opportunity and outcome between families and dynasties in Britain.
7.4 Appendix

Our simple linear model for the simultaneous determination of school performance and property prices is:

\[ s = \beta p_{-1} + \tilde{s} \]

\[ p_{+1} = \alpha s + \tilde{p} \]

Which, assuming stationarity, implies equilibrium variances of:

\[ \text{Var}(p) = \sigma_p^2 \left( \frac{1}{1 - (\alpha \beta)^2} \right) + \sigma_s^2 \left( \frac{\alpha^2}{1 - (\alpha \beta)^2} \right) \]

\[ \text{Var}(s) = \sigma_s^2 \left( \frac{\beta^2}{1 - (\alpha \beta)^2} \right) + \sigma_p^2 \left( \frac{1}{1 - (\alpha \beta)^2} \right) \]

Substituting the parameter estimates gives (7-3) and (7-4).

Rearranging, we see that

\[ \frac{\sigma_s^2}{\sigma_p^2} = \frac{1.022 - 0.049 \cdot \text{Var}(p)}{\text{Var}(s)} \]

Taking the ratio of the raw variances as 7.852 from 1998 data, gives us \( \frac{\sigma_s^2}{\sigma_p^2} = 0.084 \).
References


http://www.housing.dtlr.gov.uk/research/hss/hs2000/index.htm#1


- 292 -


- 300 -