

# Class Size, Education, and Wages\*

Christian Dustmann<sup>†</sup>   Najma Rajah<sup>‡</sup>   Arthur van Soest<sup>§</sup>

January 2002

## Abstract

This paper examines the effects of class size on the decision to stay on in full time schooling at the age of 16 and on wages at later stages in life. Little research exists on the effect of school quality on career decisions, although it has potentially important long term implications. We use micro data for England and Wales that contain an unusually rich set of variables on parental background and previous achievements. We find that class size, measured as the pupil-teacher ratio at the school level, has a sizeable and significant effect on the decision to remain in full time education beyond the minimum age. This finding is very robust and persists when school type variables, exam results, and results from past ability tests are controlled for. We also estimate wage equations including the staying on decision at age 16 as one of the explanatory variables. The effect of staying on is significantly positive for wages at age 33 and age 42, and for females also at age 23. Combining this effect with the effect of class size on the staying on decision, our results reveal that class size significantly affects future wages. Reduced form models in which wages are directly regressed on class size lead to much less accurate and generally insignificant estimates.

**Keywords:** School inputs, Educational Attainment, Training.

**JEL-Classification:** C35, I20, J24

---

\*This paper draws on research funded by the Leverhulme Trust and by the ESRC Research Centre at IFS (grant no M544285001). NCDS data have been provided by the ESRC Data Archive; the authors alone are responsible for the analysis and interpretation in this paper. Some parts of the paper have been written while the first author visited the Center of Labor Economics at the University of Berkeley, California. We thank the Center for its hospitality and financial support, and David Card for invaluable discussions and suggestions. Thanks to Steve Machin and Costas Meghir for comments on earlier drafts of the paper, and Jo Blanden and Alissa Goodman for help with the data.

<sup>†</sup>University College London, and Institute for Fiscal Studies, Gower Street, London WC1E 6BT

<sup>‡</sup>Frontier Economics, 150 Holborn, London, EC1E 2NS

<sup>§</sup>Tilburg University, P.O. Box 90153, 5000 LE Tilburg, Netherlands. Research of the third author was made possible by a fellowship of the Netherlands Royal Academy of Arts and Sciences.

# 1 Introduction

The impact of school quality on academic performance and other outcomes has been subject to much scrutiny, both within the academic literature and in political circles. The last three and a half decades have spawned a vast number of studies on the relationship between smaller class sizes and learning, much of which is based on data from the United States (see for example, Coleman et al. 1966, Betts 1995, Hanushek, Rivkin and Taylor 1996, Goldhaber and Brewer 1998, and Hanushek 1996).

In a recent survey, Hanushek (1996) comes to the conclusion that three decades of research have shown that "school resource variations are not closely related to variations in student outcomes" (p.9) - a view reiterated in his paper in the current issue. Other studies disagree. Positive effects of school quality are found, for instance, by Johnston and Stafford (1973), Card and Krueger (1992), and Heckman, Layne-Farrar and Todd (1996). Card and Krueger (1996a, 1996b) summarise evidence which is largely supportive of the view that school quality is positively related to economic outcomes. This is also the message of Krueger's contribution to this issue.

Much of this debate is centered around class size as a measure of school quality. The reduction of class sizes is an attractive way for governments to increase the resources available to schools. Hoxby (2000) remarks that class size reduction is "probably the most popular and most funded school improvement policy in the United States" (p.1). Advantages of class size reduction as a tool to improve school quality are that they are visible to the voter, and relatively easy to implement (compared to other quality enhancing measures). Yet whether class size reductions have a notable impact on outcomes is heavily disputed in the academic profession.

More recently, this debate has spread into UK politics. In its 1997 Manifesto the Labour Party featured a commitment to reduce class sizes to 30 or under for all 5, 6, and 7 year olds, which stood as one of its five key pledges. The Government has subsequently channelled resources towards reducing class sizes. Local education authorities have also implemented Class Size Plans to manage school admissions so as to meet the statutory limits on infant class sizes, in accordance with the School Standards and Framework Act 1998. In response to these initiatives, the average size of primary classes taught by one teacher in primary schools in England decreased from 27.7 to 27.5 between January 1998 and January 1999.<sup>1</sup> This reduction has continued over time with the latest provisional figures indicating that average size has decreased from 27.1 to 26.7 between January 2000 and January 2001.<sup>2</sup> Over the same period there has also been a slight decrease in class sizes in secondary schools.

The government has also recognised that there may be a link between school quality and career decisions at 16. In its 1998 White Paper *Learning to Succeed*<sup>3</sup> it suggested that "young people can be turned off education and training by poor experiences at any stage in their lives, but critical points usually occur between the ages of 13 - 19." The White Paper subsequently described a number of policy initiatives<sup>4</sup> designed to raise the number of young people continuing in education and training beyond 16.

Concerns relating to the low proportion of British teenagers remaining in education beyond the minimum school-leaving age have prompted a range of empirical studies

---

<sup>1</sup>Source: DfEE and National Assembly Data for Wales Annual School Censuses.

<sup>2</sup>Source: DfEE First Release, Class Sizes in Maintained Schools In England: January 2001 (Provisional).

<sup>3</sup>Learning to Succeed: A Framework for Post 16 Learning, June 1998, Cm 4392.

<sup>4</sup>See DfEE (2000) The Connexions Service and Schools, May 2000, DfEE 0078/2000.

examining the staying on decision (see Rice 1987, Micklewright, Pearson and Smith 1989 and Micklewright 1989). Studies have also shown that many employer-provided training schemes are only accessible to workers with higher secondary school education. Blundell, Dearden and Meghir (1996) show that access to employer provided education is strongly related to previous educational achievement. The choice as to whether or not acquire higher school education is therefore crucial for the individual's entire career path and for the final level of education s/he will achieve. An understanding of the mechanisms by which one could rise the proportion of individuals who continue education at each of these steps is important. Class size may well be one key factor.

There has been limited consideration in the academic literature of the impact that school quality might have on teenagers' decisions to continue in education beyond the age of 16. Card and Krueger (1996a, 1996b) argue that an increase in school quality induces students to attend school longer as a response to economic incentives created by higher returns to schooling, or because school is simply more pleasant. The view that there is a relationship between school quality and the length of education is supported by aggregate data (see Card and Krueger, 1992).

In this paper we examine the impact of class size on the decision to remain in school beyond the age of 16 for pupils in England and Wales. To our knowledge, this is the first study that investigates effects of school quality on the staying on decision based on micro data. Furthermore, we estimate the effect of class size, measured at age 16, on wages at ages 23, 33 and 42, via the effect on the school continuation decision.

Our analysis is based on several waves of the National Child Development Study (NCDS), a panel data survey based on a cohort of children born during one week in 1958. This data provides an unusually large set of background variables that help to

avoid the usual problems of omitted variable bias, prevalent in much of the empirical literature (see the discussion in the Todd and Wolpin contribution to this issue).

We find a significant negative effect of class size on the probability to stay at school at age 16. This finding is robust across specifications that control for various rich sets of background variables, including past performance tests and school types. Moreover, the decision to stay at school beyond the minimum school leaving age has a considerable effect on wages at ages 33 and 42, and for women also at age 23. Combining these two findings implies that bigger class size has a significant negative effect on wages later in life.

This seems to be in apparent contradiction to other studies that find no class size effects on wages, using the same data. The reason is that we explicitly incorporate the staying on decision at age 16 as the mechanism through which class size affects education level and future wages. In reduced form wage equations, as usually estimated, and where the wage is directly regressed on class size, the effect is much less accurately determined and, as we demonstrate, usually insignificant. Based on such reduced form regressions, some authors have concluded that class size has no effect on outcomes later in life. For instance, Dolton and Vignoles (1996), in an analysis of class size effect on earnings, conclude that "we found no significant relationship between additional school resources at the secondary school level and students' subsequent earnings" (p.2). Our conclusion is that there are class size effects on wages that operate through the staying on decision; however, the signal to noise ratio is too small to detect these effects in reduced form regressions, given the available sample size. We provide some simple back on the envelope calculations to illustrate the benefits and possible costs that follow from class size reductions.

We also investigate some mechanisms that may explain why class size affects the staying on decision of the teenager. One channel is to improve examination performance in the final exams. We find that class size has a significant and negative effect on the 16 year old's exam performance, conditional on parental background variables and indicators for previous achievements at age 7 and 11. This effect vanishes however when we add school type variables as additional measures for school quality. Other mechanisms could relate to the attachment of the pupil to the school. We find a positive and significant association between class size and truancy for boys, conditional on parental background, past examination performance, and ability tests earlier in life.

We suggest some potential sources of exogenous variation in class size that remains after conditioning on all the other background variables. Although hardly discussed in the literature using non-experimental data for assessing school quality effects, this is crucial for assessing the appropriateness of such studies to investigate causal effects of class size. One possible source could be that larger schools of any type have larger classes. Another source could be the variation across geographical areas in the levels and changes of cohort sizes. We provide some evidence that is compatible with these hypotheses, but leave a detailed analysis for future research.

Considering all the available evidence, we conclude that class size affects educational outcomes of teenagers in the UK by prolonging school attendance, conditional on the type of school the teenager attends. This supports the conjectured relationship between learning experiences and choices in favour of further education, suggested in the 1998 White Paper. It suggests that it is worthwhile to investigate the indirect impact of school quality enhancements, such as class size reductions, over and above its direct impact on educational performance.

The remainder of the paper is structured as follows. The next section gives some background information of the education system in England and Wales, and describes the data. In section 3, we discuss factors that are likely to affect educational outcomes, and discuss our model assumptions. Section 4 provides the results on the staying on decision at age 16. In section 5, the relations between class size, staying on, and wages are analysed. Section 6 discusses sources of class size variation, and concludes.

## 2 Background and Data

We briefly describe the secondary education system of England and Wales, and the changes that took place since the 1970's. We then discuss the sample, and the variables we use in our analysis.

### *Secondary Education in England and Wales*

During the early 1970s a tri-partite selection based system of grammar, secondary modern and technical schools operated within many local education authorities. Performance in the 'eleven plus' examination taken at age 11 or 12 was used to select pupils into grammar schools, secondary modern schools and technical (vocational) schools. This system, however, was criticized because of the selection purely on the basis of performance at the age of 11 or 12. As a result, from the mid-1960s onwards, a number of local education authorities had moved away towards a system with comprehensive schools taking all children in a given local authority, regardless of their ability.

In 1974 there were two sets of public examinations in England and Wales - Ordinary level examinations (O levels) and Certificates of Secondary Education (CSEs). These were usually taken at the age of sixteen and provided the first opportunity for teenagers

to be formally tested in a broad range of subjects on a national basis, before either entering the labour force, continuing in secondary school for another two years, or participating in more vocationally based training.

### *Data and Variables*

Our analysis is based on data from the National Child Development Survey (NCDS). The same data source is used for several other studies in the UK on similar topics, such as Dolton and Vignoles (2000), Harmon and Walker (2000), Feinstein and Symons (1999), Currie and Thomas (1998), Robertson and Symons (1996), and Dearden, Ferri and Meghir (2000). The NCDS followed a cohort of individuals born between 3rd and 9th March 1958 (see Micklewright 1986 for a detailed description of these data). Of particular interest for class size is the data recorded in the third and fourth sweeps of the survey (NCDS3 and NCDS4) and information collected in the Public Examinations Survey (PES), a follow-up survey to NCDS3. NCDS3 was conducted in the spring of 1974, and records extensive information about the respondents, such as educational and physical development, aspirations for the future, spare time activities etc., as well as the information usually gathered in household surveys. Similar information was also collected for NCDS4 in 1981 when cohort members were aged 23. NCDS4 also contains further details on education and employment experience. For wages, we use NCDS4, and the latest surveys NCDS5 and NCDS6, conducted in 1991 and 2000 when individuals were 33 and 42 years old.

The data set we use to study the relationship between school continuation and class size is based on a sub-sample of almost 4,000 cases out of the possible 11,602 who were traced at NCDS3, PES and NCDS4. Differences in the educational system in Scotland restrict our analysis to teenagers living in England and Wales. More important is the



problem of missing or incorrectly recorded information. Information collected at the third sweep was retrieved from four separate sources (from the cohort member, from his or her parents, from the school that the 16 year olds attended and from the teenager's doctor) and many respondents failed to complete one or more of these questionnaires.

For information on school leaving decisions, we draw on NCDS4. This wave contains a month-by-month diary recording various activities from May 1974 to January 1982. We use the information recorded in February 1975 to see whether, at the end of their 16th year, the sample members were in full-time schooling, had a regular job, or were following a training programme.<sup>5</sup>

Table 1 describes the variables used in our analysis and provides means and standard deviations. The table shows that just under a third of the sample continued in education beyond the minimum school leaving age and that the majority of the sample either had entered into the labour force or participated in training by February 1975.<sup>6</sup> Girls who did not remain in full time education after the age of 16 were more likely to join the labour market full time than boys. Two out of three boys who did not continue with full time education after age 16 joined some kind of training scheme.

Average log wages are larger for males than for females, and the difference increases over time. The standard deviation of log wages for both gender increases likewise.

---

<sup>5</sup>We classify all those who have any element of training associated with their job in the "training" category, in addition to those enrolled in full time training schemes. Thus, for example, an individual in part time employment and on an apprentice scheme would be classified as being in training, as would someone who was simultaneously on a government training scheme and in part time education.

<sup>6</sup>Only about 3 percent of the entire NCDS4 sample was not in one of these three states. We excluded these observations.

Table 1: Descriptive Statistics

Description	Females		Males	
	Mean	Standard deviation	Mean	Standard deviation
<u>Main Variables of Interest:</u>				
Continuation decision:				
Dummy Stay on at school	0.320		0.318	
Dummy Enroll in training scheme	0.221		0.393	
Dummy Regular Job	0.459		0.289	
Number of O levels/CSE Grade 1s passed	2.480	2.935	2.269	2.898
Class size (number of pupils per teacher)	17.202	1.898	17.023	2.013
Real log hourly wage, age 23*	1.566	0.321	1.742	0.327
Real log hourly wage, age 33*	1.873	0.487	2.225	0.415
Real log hourly wage, age 42*	1.929	0.499	2.328	0.514
<u>Background Variables:</u>				
Number of older siblings	0.429	0.642	0.435	0.646
Number of younger siblings	1.194	1.258	1.197	1.215
Years of full-time education, father	10.99	1.72	11.02	1.75
Years of full-time education, mother	11.02	1.44	10.99	1.40
Logarithm of household income	3.856	0.435	3.869	0.364
Score on sum of age 7 maths and reading test**	0.679	0.197	0.671	0.204
Score on sum of age 11 maths and reading test**	0.470	0.202	0.480	0.208
<i>Dummy Variables; 1 if condition is satisfied:</i>				
Father working when teenager is 16	0.896	0.304	0.915	0.278
Mother working when teenager is 16	0.684	0.464	0.708	0.454
Teenager attends a secondary modern school	0.242	0.428	0.235	0.424
Teenager attends a technical school	0.005	0.075	0.011	0.107
Teenager attends a comprehensive school (non-selective state run)	0.527	0.499	0.551	0.497
Teenager attends a grammar school (higher ability state run)	0.167	0.373	0.134	0.340
Teenager attends a private school	0.039	0.195	0.045	0.209
Teenager attends a special school (handicapped and special needs children)	0.017	0.130	0.021	0.144
Teenager attends a single sex school	0.286	0.452	0.249	0.432
Teacher considers parents to be interested in teenager's school work	0.759	0.427	0.744	0.436
Parents want teenager to sit A levels	0.282	0.450	0.229	0.420
Parents want teenager to go to university	0.347	0.476	0.371	0.483
Teenager has private room for studying	0.904	0.293	0.883	0.321
Teenager truant at least once over the last year	0.470	0.499	0.484	0.499

(\*): In 2000 prices. (\*\*): Test score variables are an average of the reading and math tests, on a scale from 0 to 1.

### 3 Modelling Educational Outcomes

Our modelling framework is in the tradition of the educational production function literature, where inputs are combined to produce some measure of learning, such as school continuation. When isolating the effect of one particular measure of school quality on achievements - class size in our case - endogeneity issues arise, which are due to the omission of other input variables that affect both educational outcomes, and the respective school quality measure.

Our estimation strategy is to condition on other input factors that are correlated with both achievement and class size, using the unusually rich set of background variables provided by the NCDS data. In this way we eliminate or at least substantially reduce the correlation between the error term and the class size variable. We first discuss the variables we include and then discuss our model assumptions.

#### *Factors affecting Exams and Career Choice*

Family background almost certainly plays a major role in determining educational outcomes (see Ermisch and Francesconi 2001 for a recent UK study). It may affect pupils' achievements in various ways. Becker (1981) distinguishes between financial and time resources allocated to the child. Financial resources may be used to choose better schools for the child, and to provide a more suitable environment for studying. Time inputs may consist of the time parents spend with the child explaining homework exercises, for instance. Better educated parents may also be more efficient in aiding their children's education (e.g. when helping the child with homework) and may provide more support for their academic development.

In the empirical analysis, we measure financial resources of the family by family

income.<sup>7</sup> As a measure of time inputs, we use the labour market status of the parents, particularly of the mother. As further family background variables, we include occupational and social status of the parents.

The quality of the time parents spend with their children may be as important for their learning success as the quantity of time. As proxies for the quality of time, we include parental education. In addition to parental inputs, a teenager's performance may be also affected by studying conditions within the household. Here an isolated working environment for undisturbed study may be particularly important. We include an indicator variable for whether the child has a separate room for studying.

In families with more than one child, children are likely to compete for resources. Becker's (1981) work suggests that parental attention is reduced as family size increases. Hanushek (1992) finds that the birth order plays an important role for children's academic performance. To capture these effects, we include the numbers of older and younger siblings as regressors.

Parental interest in the child's academic performance may not be entirely captured by the above set of family background variables. Keeping wealth and education constant, parents may still differ in their preferences regarding the education of their child. Hanushek, Rivkin and Taylor (1996) emphasise that correlation between these preferences and school quality – which may depend upon the parents' choice – may contaminate estimates of the effect of school quality if the parents' preferences are omitted. It is therefore desirable to include variables that capture parents' commitments to their children's educational career. We use a variable that reflects the teacher's assessment

---

<sup>7</sup>The income information in NCDS3 is recorded in a banded form. We construct a continuous measure of income, taking into account all sources of household income, following Micklewright (1986).

of parental interest in the teenager's school performance, and variables that indicate whether the parents want the teenager to complete Advanced levels (A levels) or to follow a university education.<sup>8</sup>

When isolating the effects of school quality variables on academic achievement, it is important to consider both current inputs and differences in previous academic achievements. Pupils with different previous achievements may attend schools of different quality. Also, previous achievements capture unobserved ability or intelligence, which is likely to be correlated with current school quality measures as well as achievements. We follow Hanushek, Rivkin and Taylor (1996) and use standardised test scores to control for these differences. Test scores also reflect differences in ability between children. We use combined test scores from attainment tests in mathematics and reading comprehension that respondents sat at the age of 7 and 11. These should reflect variation in unobserved ability and past inputs across children.

Environmental factors, such as economic characteristics of the environment where the child grows up (keeping family background constant) may also affect school attainment. We include regional indicators in all regressions.<sup>9</sup>

As a measure of class size, we use the pupil teacher ratio in the school attended

---

<sup>8</sup>Like O levels, A levels are a formal set of public examinations that teenagers take after two years of study of a set syllabus (usually at the age of about 18), normally in one to five subjects. A level results are used to determine eligibility for entry into tertiary education institutions such as universities.

<sup>9</sup>We have also experimented with variables which reflect the immediate environment of the child. We included variables which measure the rate of unemployment, as well as the percentage of unskilled manual workers, on a local authority level. The local authority data covers around 500 separate areas, and therefore relates to narrowly defined labour market areas. These variables were never significant in our estimations, and including them did not change the other parameters.

at age 16. This is the ratio of the school roll to the number of full time equivalent teachers. Aggregation to school level avoids the simultaneity problem that arises if weak pupils are assigned to small classes, or if highly able students are assigned to specialised classes. The pupil teacher ratio is related to the type of school the child attends. In the empirical analysis, we estimate specifications that use this ratio as the only measure of quality, and specifications that also include dummies specifying the type of school that the 16 year old attends. Depending upon the local authority, this can be a grammar school, a secondary modern school or technical school (to which teenagers are allocated on the basis of ability) or a non-selective comprehensive school. Teenagers with special educational needs may attend a special school. Those that are educated outside of the state system will attend private schools.

### *Models and Estimation*

As we pointed out earlier, the estimation strategy for isolating the causal effect of class size is to condition on a very broad set of factors that may affect both achievement and class size. This will reduce or eliminate the correlation between the error term and the class size variable. This estimator, based upon a flexibly-specified regression model, can be seen as a matching estimator. Since our models are non-linear, however, the required assumptions are somewhat different from those in the standard matching model. We therefore formulate them explicitly.

Consider the following latent relationship between outcome ( $O^*$ ) and class size ( $S$ ).

$$O_i^* = \alpha_0 + X_i\alpha_1 + \gamma S_i + \epsilon_i. \quad (1)$$

Here  $i$  is an index for individuals, and the vector  $X_i$  contains all other conditioning variables. A standard model (probit, ordered probit, tobit, ..., depending on the nature

of the observed outcome) is obtained if the error terms  $\epsilon_i$  follow a normal distribution with mean zero and constant variance  $\sigma^2$ , independent of the regressors  $X_i$ .

Standard probit or tobit assumptions imply that  $S_i$  is exogenous, while we have argued that common factors may drive  $S_i$  and  $O_i^*$ . It is easy to show that the Maximum Likelihood (ML) estimator for  $\gamma$  – the main parameter of interest – remains consistent under the weaker assumption that, conditional on  $X_i$  and  $S_i$ ,  $\epsilon_i$  is normal with some variance  $\sigma^2$  and a mean that is linear in  $X_i$  and does not depend on  $S_i$ . In other words, ML is consistent for  $\gamma$  if the endogeneity in  $S_i$  is due to observables, and, conditional on all regressors, the variation left in the error term is not correlated with class size.

## 4 Results

### *The decision to continue full time schooling*

We begin by using probit models to analyse the teenager’s decision to remain in full time education after the minimum school leaving age. The numbers in table 1 indicate that only about one third of teenagers in the sample continued in school beyond age 16. In table 2, we present the probit results in terms of the estimated marginal effects, i.e. the estimated partial derivatives of the probability of staying in school with respect to the explanatory variables, computed at the sample mean. The specifications we present are reduced form estimations in the sense that we do not condition on exam success – we discuss this later. In addition to the variables in the table, all specifications include regional dummies and social class and occupational class variables. The latter refer to both parents’ labour market states, including dummies for being a professional, skilled, unskilled or semi-skilled worker, working in the service sector, being self-employed, and

working on a farm.

We focus on the effects of class size and briefly discuss the other estimates. Note, however, the implications of the endogeneity discussion given above: the other effects are interpreted under the assumption that  $S_i$  is exogenous, while the effect of school quality is still consistently estimated under weaker assumptions. The results in column 1 show that family background variables are important for the staying on decision. Pupils in larger families are less likely to stay on at school, and older siblings are more likely to continue than younger siblings. The father's and mother's years of education have the expected positive impact on the child's probability of continuing in full time education. Family income is not significant.<sup>10</sup> The availability of an isolated working environment increases the probability of staying on. The effect of the class size variable is quite strong and significant.

In column 2, we add the test score variables at age 7 and age 11 (see table 1 for their definition). The combined test score at age 11 has a particularly strong influence on the staying on decision. Including them reduces the effect of family background variables, since past performance is related to these variables in the same way as the staying on decision. The effect of class size is also reduced, but remains strongly significant.

In column 3, we add parental preference variables. These barely affect the coefficient of the class size variable, although they have sizeable effects on the staying on decision.

Finally, in column 4 we add the school type variables. This reduces the size of the coefficient of the class size variable, but it remains negative and significant. An increase in class size by one standard deviation decreases the staying on probability by

---

<sup>10</sup>The variation in family income is strongly correlated with parental class and occupation. If we omit these variables, the income coefficient increases and becomes significant.



Table 2: Staying on decision, Probit Models, Marginal Effects

Specification	1		2		3		4	
	Effect	StdE	Effect	StdE	Effect	StdE	Effect	StdE
Class size	-0.0399	0.0041	-0.0289	0.0050	-0.0273	0.0050	-0.0199	0.0055
Number of older siblings	-0.0758	0.0137	-0.0364	0.0152	-0.0297	0.0149	-0.0305	0.0151
Number of younger siblings	-0.0260	0.0069	-0.0114	0.0082	-0.0013	0.0081	-0.0006	0.0082
Father working	0.0838	0.0338	0.0186	0.0422	0.0088	0.0423	0.0106	0.0427
Mother working	-0.0176	0.0228	-0.0048	0.0256	-0.0050	0.0252	-0.0025	0.0255
Individual is female	0.0082	0.0160	0.0205	0.0179	0.0161	0.0176	0.0115	0.0178
Years of full-time education, father	0.0399	0.0052	0.0236	0.0061	0.0135	0.0059	0.0125	0.0060
Years of full-time education, mother	0.0574	0.0065	0.0372	0.0074	0.0226	0.0071	0.0217	0.0073
Logarithm of household income	0.0398	0.0256	0.0150	0.0282	-0.0226	0.0278	-0.0281	0.0280
Teenager has private room for studying	0.0969	0.0242	0.0768	0.0276	0.0618	0.0272	0.0594	0.0276
Percent score on sum of age 7 maths and reading test			1.1236	0.0664	0.8053	0.0667	0.7320	0.0697
Percent score on sum of age 11 maths and reading test			0.1055	0.0653	0.0619	0.0641	0.0424	0.0649
Parents take interest in teenager's school work					0.0568	0.0210	0.0550	0.0212
Parents want teenager to sit A levels					0.2354	0.0292	0.2292	0.0294
Parents want teenager to go to university					0.4526	0.0251	0.4440	0.0255
Teenager attends a comprehensive school							0.0537	0.0253
Teenager attends a technical school							0.0896	0.1009
Teenager attends a grammar school							0.1246	0.0407
Teenager attends a private school							0.1760	0.0687
Teenager attends a single sex school							-0.0311	0.0870
Teenager attends a special school							0.0187	0.0236
Log-Likelihood	-1955.18		-1323.06		-1154.27		-1144.86	
Number of Observations	3783		3112		3112		3112	

Note: Presented are estimated marginal effects at the sample means and their standard errors. All specifications include dummies for regions and for parents' socio-economic and occupational class.

about 4 percentage points.

The school type effects are quite strong. Pupils who attend grammar or private schools have a 12 or 17 percentage point higher probability of staying on in full time education than pupils in the base category (modern schools). The school type effects may have several explanations. Peer pressure in grammar or private schools may discourage teenagers from leaving school at the first possible opportunity. Furthermore, specialist staff employed to give informed advice about education and career choices may have an effect on school-leaving decisions.

Results for separate models for boys and girls are presented in table 3. For all specifications and for both genders, the effect of class size is strong and significant. It is always stronger for boys than for girls.

Table 3: Staying on decision, Males and Females

Specification	<b>1</b>		<b>2</b>		<b>3</b>		<b>4</b>	
	Effect	Standard error	Effect	Standard error	Effect	Standard error	Effect	Standard error
<b>Females</b>								
Class Size	-0.0363	0.0060	-0.0243	0.0074	-0.0234	0.0073	-0.0167	0.0080
Log-Likelihood	-1002.99		-689.48		-600.00		-594.49	
Number of Observations	1896		1561		1561		1561	
<b>Males</b>								
Class Size	-0.0429	0.0058	-0.0335	0.0070	-0.0310	0.0069	-0.0230	0.0077
Log-Likelihood	-935.99		-615.75		-535.80		-530.48	
Number of Observations	1887		1551		1551		1551	

Note: Same specifications as in table 2. Presented coefficients are estimated marginal effects at the sample mean and their standard errors.

### *School, Training or Work*

There are two alternatives to full time education: training schemes and labour market participation. Table 1 shows that 22 percent of females, and 39 percent of

males do not enter the labour market directly after leaving school at 16, but enroll in some type of training scheme. We have estimated models where we separate these possibilities and explain which of the three ordered outcomes is chosen, stay at school, training scheme, or regular job. We use a generalised ordered probit model, where one of the cut off point is allowed to vary with the regressors (see Pradhan and van Soest 1995 for details). This uses the implied ordering of educational choices, but is more flexible than standard ordered probit, since it allows for heterogeneity in the threshold parameter. Table 4 gives the results for the rich specification corresponding to column 4 in table 2, for the pooled sample, and for males and females separately. We only report class size effects.

Table 4: School, Training, or Work

Decision:	Stay in School		Training		Labour Market		Log-likelihood
	Effect	StdE	Effect	StdE	Effect	StdE	
All	-0.0214	0.0055	0.0114	0.0047	0.0100	0.0051	-7198.23
Females	-0.0158	0.0080	0.0045	0.0078	0.0113	0.0081	-3580.96
Males	-0.0240	0.0080	0.0186	0.0075	0.0054	0.0067	-3557.50

Note: Specifications as in column 4 of table 2. Presented coefficients are estimated marginal effects of increasing class size by one pupil on the probability to *Stay in School*, go into *Training*, or join the *Labour Market*, evaluated at the sample means, and their standard errors (stdE).

According to the results for the pooled sample, an increase in class size reduces the staying on probability, while it increases the probability of enrolling in training schemes and of joining the labour market by roughly equal amounts. Separate estimates for boys and girls reveal class size effects on the probability to stay at school that are similar to the effects in table 3, with the effect for boys exceeding that for girls. The effects of class size on the other two probabilities are positive for both genders, but size and significance levels differ: boys in larger classes are significantly more likely to

go into a training scheme, while for girls, a larger class size significantly increases the probability to accept a regular job.

### *Class Size and Exam Success*

One possible way class size may affect the school continuation decision is via success in national exams. To investigate this, we take as our measure of academic success the number of Ordinary level (O level) passes achieved by 1974.<sup>11</sup> Table 5 presents tobit estimates, where the dependent variable is the number of O levels achieved. We report coefficients on the class size variable only, for a pooled model and for separate models for boys and girls. The same regressors are used as in columns 1-4 in table 2. The first column is a basic specification that includes various family background variables, and the class size variable. In this specification, class size has a significant and sizeable negative effect on exam results: An increase in class size by one standard deviation decreases the number of O levels achieved by about 0.6.

In the second specification, standardised test score variables measuring past performance and mental ability are added. This changes the coefficient on the class size variable substantially. The effect on exam performance drops to one third of its size, but remains statistically significant.

In column 3, parental preferences regarding the offspring's future academic career are added. Including these variables reduces the size of the class size variable only

---

<sup>11</sup>Number of O levels includes CSE Grade One passes. Since NCDS3 dates from Spring 1974, we observe the cohort members when they are still in compulsory full time secondary education and a few months before they sit their first set of public examinations, O levels and Certificates of Secondary Education (CSE's), in June 1974. The PES conducted in 1978 has detailed information on the examination results of about 95% of respondents to NCDS3, obtained from the schools.

Table 5: Exam Success and Class Size

Specification	<b>1</b>		<b>2</b>		<b>3</b>		<b>4</b>	
	Coeff.	StdE	Coeff.	StdE	Coeff.	StdE	Coeff.	StdE
<b>Pooled Sample</b>								
Class size	-0.3155	0.0378	-0.1072	0.0323	-0.0879	0.0310	-0.0099	0.0330
Log-Likelihood	-6930.84		-4849.07		-4732.98		-4712.00	
N. of Observations	3783		3112		3112		3112	
<b>Males</b>								
Class Size	-0.3397	0.0532	-0.1065	0.0443	-0.0898	0.0424	-0.0210	0.0447
Log-Likelihood	-3575.098		-2453.047		-2388.887		-2377.023	
N. of Observations	1896		1561		1561		1561	
<b>Females</b>								
Class Size	-0.2955	0.0541	-0.1074	0.0472	-0.0840	0.0455	0.0053	0.0490
Log-Likelihood	-3342.78		-2379.68		-2325.86		-2310.27	
N. of Observations	1887		1551		1551		1551	

Note: Specifications in columns 1-4 are identical to those in columns 1-4 in table 2. Reported coefficients (Coeff.) are estimated marginal effects of increasing class size by one pupil on the number of O-levels obtained (and their standard errors (StdE)), based on Tobit estimations.

slightly, and it remains significant.

In column 4 school type dummies as additional school quality indicators are added. The base category refers to secondary modern schools (lower ability state-provided schools). The coefficient of the class size variable is now virtually zero, and its standard error has increased substantially. Estimation results on the school type dummies in this specification for the pooled sample are presented in the first two columns of table 6 (modern schools are the reference category). The order of the effects of school types is inversely related to class size, as shown in the fifth column of table 6. Teenagers attending comprehensive (non-selective state run), technical, or grammar schools (higher ability state run schools), or private (selective non-state run) schools perform significantly better than pupils in secondary modern schools.

The effects of school type dummies are considerably larger than could be expected

Table 6: Exam Performance, School Type, and Class Size

School type	Coeff	StdE	N. Obs.	Class Size	StDev
modern	–	–	947	18.25	1.69
comprehensive	0.4023	0.1556	2021	17.13	1.58
technical	0.6579	0.5513	29	16.67	1.76
grammar	1.2565	0.2142	558	16.11	1.41
private	1.2420	0.3107	178	14.69	2.80
special	-0.0514	0.1468	78	13.21	4.17

Note: Reported coefficients in column 1 are estimated effects of the attending the particular school type on the number of O-levels achieved, relative to modern schools. Standard errors are reported in column 2. Estimates are based on Tobit models, using the pooled sample of males and females. The specification is equal to that in column 4 of table 2, including dummies for parents' socio-economic and occupational class and regional dummies. Columns 4 and 5 give the mean and standard deviation of class size for the different school types.

on the basis of class size differences and the class size effects in column 3 of table 5 only. For example, the average difference in class size between grammar schools and modern schools of -2.1, combined with the parameter estimate of -0.088 would lead to an effect of 0.185, much less than the coefficient of 1.26 for grammar schools in table 6.

Separate estimates for boys and girls are reported in the lower panels of table 5. In all specifications, the effect of class size on exam success is similar for both genders. Coefficients drop to zero when the school type variables are added. The school type effects (not reported) are also similar for boys and girls. We conclude that class size is negatively related to exam performance conditional on a large array of background variables, including previous ability tests. The effect vanishes, however, once we condition on school type as an additional indicator of school quality.

### *Exam Success and Career Decisions*

The models presented in tables 2, 3 and 4 are reduced form specifications in the sense

that they do not include exam performance as a regressor. However, class size can have both a direct, and an indirect effect via exam success on the continuation decision. Since in our most general specification, the effect of class size on examination is basically zero, only the direct effect affects the staying on decision in this specification. Nevertheless, comparisons between models that do and do not include school type variables allow us to assess whether the drop in the class size coefficient in the career choice equation is due to a drop in the direct effect, the indirect effect, or both.

We have estimated examination and career choice equations (distinguishing between staying on in full time education, enrolling in some training scheme, or joining the labour market) simultaneously by maximum likelihood.<sup>12</sup>

In table 7 we display results for the pooled sample. The first panel reports results when school type variables are excluded from both equations. The direct effect of the school quality measure on the continuation decision is five times as large as the indirect effect induced by school quality affecting examination performance. Both effects are significant. The second panel reports results for the most general specification, with school type variables included in both equations. The indirect effect drops to zero, as expected from the results in table 5. The direct effect is hardly affected by the school type variables - it drops only slightly in size, and remains significant. This suggests that the drop in the class size coefficient in the reduced form specifications (when including

---

<sup>12</sup>We allow for correlation between the errors in exam success and career choice equations. To identify this model, we rely on the normality assumption. We also estimated models using indicators of school absenteeism for reasons of illness in the year before the final examinations as instruments, assuming that absenteeism affects career choices only indirectly via examination success. Results were similar to those in 7.

Table 7: Direct and Indirect Marginal Effects of Class Size on Career Decision

Decision:	Stay in School		Training		Labour Market	
Variable	Effect	StdE	Effect	StdE	Effect	StdE
	No school type variables included					
Total Effect	-0.0332	0.0056	0.0184	0.0045	0.0147	0.0049
Direct Effect	-0.0277	0.0053	0.0165	0.0047	0.0112	0.0050
Indirect Effect	-0.0054	0.0020	0.0019	0.0005	0.0034	0.0018
	Log-Likelihood: -7153.02; $\rho = -0.171$ (with t-value -1.54)					
	School type variables included					
Total Effect	-0.0213	0.0060	0.0124	0.0051	0.0089	0.0056
Direct Effect	-0.0215	0.0056	0.0124	0.0051	0.0090	0.0052
Indirect Effect	0.0002	0.0016	-0.0000	0.0000	-0.0001	0.0009
	Log-Likelihood: -7093.52; $\rho = -0.154$ (with t-value -1.37)					

Notes: Identification relies on normality assumptions; estimates in the first and second panel are based on the specifications of columns 3 and 4 in table 2, respectively;  $\rho$ : correlation coefficient between error terms in exam and career decision equations.

school type variables), as reported in table 2, is largely explained by the indirect effect (via examination results) going to zero.

### *Class Size and Truancy*

What are the possible mechanisms by which class size may affect the staying on decision? Card and Krueger (1996a, 1996b) suggest that a positive relationship between school quality and the length of school attendance could be due to the increase in the perceived monetary payoff of an additional year of schooling, or simply to the fact that school is more pleasant if quality is higher. Lazear (2001) emphasises as one channel by which smaller class sizes may enhance the quality and pleasantness of learning a reduction in disruptions.

Our results in the previous section indicate that effects of class size on the staying on decision via a higher success in national exams vanishes when conditioning on school type as additional school quality measures. However, not all of the improvement in



school quality may be reflected in O level results. If smaller class sizes improve the perceived environment in which further education takes place, then this may have effects on the attachment of the pupil to school, and could be reflected by variables which are indirect measures of motivation and commitment. One such variable is truancy. We observe in our data self-reported truancy due to factors other than illness. The numbers in table 1 indicate that nearly 50 percent of pupils report that they have been absent from school at least once over the last year when they should have been there.

Table 8: Truancy and Class Size

	<b>All</b>		<b>Males</b>		<b>Females</b>	
	Effect	StdE	Effect	StdE	Effect	StdE
Class Size	0.0135	0.0055	0.0222	0.0078	0.0055	0.0080
Log-Likelihood	-1992.55		-1001.65		-975.70	
Number of Observations	3112		1561		1551	

Note: Estimates are based on the specification in column 4 of table 2.

We estimate probit models explaining truancy from class size and the background variables of the general specification in column 4 of table 2. Results are reported in table 8. The pooled estimate is significant and suggests that an increase in class size by one standard deviation increases the probability of truancy by about 2.6 percentage points. When breaking down the sample in males and females, it becomes obvious that boys' truancy responds to class size reductions more than girls'. While the effect for boys increases to 4.4 percentage points, and is significant, the effect for girls is much smaller and not statistically significant. This is compatible with the larger effect of class size on the staying on decision of males.

## 5 Class Size, Staying at School, and Wages

If class size has a positive effect on the decision of the individual to stay on at school beyond the minimum school leaving age, then we should be able to detect this effect in wages later on in the individual's career.

We estimate the effect that the decision to stay in full time education has on wages when the individual is 23, 33 and 42 years old. To make our sample compatible with the one used for the analysis above, we omit all individuals who obtained their secondary school education in Scotland (see discussion above).<sup>13</sup> Our wage variable is the log of hourly gross wages. We consider salaried workers, either in full-time or in part-time employment.

To infer on the effect of class size on wages, we combine the effect of class size on the staying on decision that we discussed in the previous section with the effect of the staying on decision on wages. Hence, we estimate a structural model, where we specify exactly the mechanism by which class size affects wages at later stages in life. In table 9, we present results that condition on the same set of background variables as in column 4 of table 2, except for the class size variable. The coefficients reported in this panel are thus the effects of the staying on decision on wages, conditional on the whole set of family and school background variables.

At age 23, the difference in log wages for males between those who stayed on in full time education and those who did not is close to zero. The differences at age 33 and age 42 are larger, suggesting a difference in log wages of about 0.15 and 0.13 for those who do, and who do not continue in full time education after age 16. Again, there is

---

<sup>13</sup>Results are very similar when we use the full sample.

Table 9: The effect of the staying on decision on log wages.

	Males				Females			
	Coeff.	StdE	N.Obs.	$R^2$	Coeff.	StdE	N.Obs.	$R^2$
Wage at	Conditional on all background characteristics except class size							
Age 23	-0.0331	0.0285	1172	0.051	0.0807	0.0239	1049	0.193
Age 33	0.1454	0.0309	1030	0.289	0.1763	0.0403	904	0.296
Age 42	0.1331	0.0342	1173	0.281	0.1850	0.0351	1289	0.219

Note: All specifications also include regional dummies at age 16 and at present employment, and all background characteristics as in specification 4 in table 2, except for the class size variable.

no significant effect at age 23. For females, the differences in wages are already sizeable at age 23 (with a log difference of 0.08, conditional upon background characteristics), and further increase later on.

Consistent with the standard human capital earnings model (Mincer 1974), these results suggest that the wage differentials between levels of education is small at early stages of the life cycle (see Card and Lemieux 2001 for evidence for this for Canada, the UK and the US). The continuation decision at age 16 is an important step in determining an individual's ultimate education level, and a higher education level pays off only after some years, due to the implied loss of labour market experience of those with longer schooling. An immediate implication is that school quality effects on wages may be more difficult to detect at early career stages.

The negative effect of class size on the staying on probability and the positive effect of the staying on decision on wages imply a negative effect of class size on wages. The size of this effect can be estimated by multiplying the effects in tables 3 and 9. The results are summarised in the upper panel of table 10. This is based upon the final

column of table 3 and table 9. Standard errors are computed using the delta method.<sup>14</sup> One additional pupil per teacher reduces wages at age 33 by 0.294% for females and 0.334% for males. At age 42, the effects are 0.309% and 0.306% for females and males respectively. The effects are significant at the (two-sided) 5% level for males and at the 10% level for females. At age 23, the effect is insignificant (and of the wrong sign) for males. The effect for females has the right sign and is significant at the 10% level.

Table 10: Effect of Class Size on Wages

Wage at	Males			Females		
	Effect	StdE	t-val.	Effect	StdE	t-val.
<u>Structural Estimates*</u>						
Age 23	0.076	0.070	1.08	-0.135	0.076	-1.78
Age 33	-0.334	0.133	-2.52	-0.294	0.156	-1.88
Age 42	-0.306	0.129	-2.37	-0.308	0.159	-1.94
<u>Reduced Form Estimates**</u>						
Wage at	Effect	StdE	t-val.	Effect	StdE	t-val.
Age 23	-0.321	0.562	-0.57	-0.876	0.535	-1.64
Age 33	-0.613	0.641	-0.96	-1.592	0.793	-2.01
Age 42	-0.150	0.753	-0.20	-0.451	0.751	-0.60

Note: Percentage change in the wage if class size is increased by one pupil per teacher.

(\*): Based upon the results in Table 2, column 4, and Table 9, bottom panel.

(\*\*): Specifications include regional dummies and all background variables of specification 4 in table 2.

These effects are not commonly found in the literature. Papers by Dolton and Vignoles (2000) and Dearden, Ferri and Meghir (2000), for example, use the same

---

<sup>14</sup>The basic assumption required for consistency of the estimate and its asymptotic standard error is that the error term in the wage equation has conditional mean zero given regressors in both equations and given the outcome of the staying on decision. This guarantees that the two effects are asymptotically independent and makes it easy to apply the delta method. Details are available upon request from the authors.

data, but generally find insignificant class size effects on wages. The reason is that we explicitly consider one particular way in which class size relates to wages - via the staying on decision. Incorporating the continuation decision as additional information leads to additional structure and a more efficient estimator, compared to specifications of reduced form models in which wages are directly regressed on class size. To illustrate this point, reduced form estimates are presented in the lower panel of table 10. Here we have included class size and all background variables from column 4 of table 2 that affect the continuation decision. To capture regional wage differentials, we also condition on region of residence at the time the wage information was obtained. Again, we have excluded individuals who received their secondary education in Scotland. Since the effect of continuation on wages should work via the decision to acquire further education, we do not condition on educational outcomes.

The estimated parameter of class size measures the indirect effect via the staying on decision as well as any direct effects of class size on wages. As before, we report the percentage changes in the wage if class size is increased by one pupil per teacher. The effect of class size on log wages is negative for both males and females and at all ages. The coefficients range from -0.15% to -1.5%. It is generally larger for females than for males. For females at the age of 33 the effect is very large and significant at the 5 percent level. The other coefficient estimates are statistically insignificant.<sup>15</sup>

The order of magnitude of the reduced form estimates is similar to that of the "structural" two step estimates, but the standard errors are much larger: the reduced form estimates lack precision, with estimated standard errors ranging from 0.53 to 0.79, while the "structural" estimates have imputed standard errors of about 0.16 for

---

<sup>15</sup>Dearden et al. (2000) find a significant class size effect for females in the same age category.

females and 0.13 for males. Given the limited size of the data set, the sampling error is simply too large to detect a class size effect in reduced form specifications. Since the reduced form estimates do not use the information on the staying on decision, they lead to imprecise estimates of class size coefficients in a wage equation of the type estimated above. The reduction in standard errors using information that class size at age 16 affects wages through the continuation decision at age 16 leads to standard errors that are 3 to 5 times smaller than those in the reduced form regression. To obtain a similar accuracy in reduced form regressions requires a sample size which is about 9 to 25 times as large as the samples we use in this study.

## 6 Discussion and Conclusion

### *What is the benefit of class size reductions?*

To give some idea of the potential effects of a class size reduction on earnings over an individual's life cycle, we perform some back of the envelope calculations in the spirit of Krueger's contribution to this issue. Using data from the British Labour Force Survey (LFS) on wages between 1992 and 2000, we construct hypothetical wage profiles for the 1958 cohort of males and females between the age of 20 and mandatory retirement age (65 for males and 60 for females), controlling for additive cohort effects.<sup>16</sup> These profiles are used to compute the effect of a class size reduction by one pupil on the discounted value of life cycle earnings of an average male and female individual of the 1958 cohort. Based on the estimated class size coefficients in table 10, we assume a

---

<sup>16</sup>The identifying assumption is that time effects are zero. Estimated cohort effects appear to be very small.

constant class size effect of -0.3% per additional pupil on earnings from age 30 until retirement age.

Using the estimated wage profiles and assuming that a regular working year has 1800 hours, we compute the discounted value of annual earnings as  $(1 + \rho)^{-[t-16]} * exp(lw_t) * 1800$ , where  $\rho$  is the real discount rate (all wages are in 1991 prices). Adding up over the years that are assumed to be affected by class size and multiplying by 0.003, gives the life cycle benefit of a class size reduction by one pupil in 1991 Pounds.

According to Krueger (2001), a reasonable real discount rate would be 4 percent. If annual productivity growth is around 2 percent, this would effectively reduce the discount rate by approximately 2 percent. Calculated at that rate, the current value of future earnings benefits of a reduction in class size by one pupil, working solely through the effect of class size on the decision to stay on at school after the age of 16, is about 683 Pounds for males, and 493 Pounds for females. These calculations are conservative in that they consider only wages after age 30; furthermore, they do not take account of secondary effects of education on earnings - for instance that the higher educated typically work more hours over their life cycle.

If we assume that class sizes remain roughly constant during the last two years of compulsory education, a reduction in class size by 5 pupils would increase life cycle earnings by 3415 Pounds for males, and 2445 Pounds for females (in 1991 prices), due to the effect of class size on the school continuation decision. The average benefits per pupil would thus be 2930 Pounds. Average class size in our sample is about 17 pupils, so that a reduction by 5 pupils would cost 5/12 of the average teacher salary for the duration of exposure to this class size.

### *How long is the period of exposure?*

To compare the cost of class size reduction with the possible benefits, an assumption is needed on the length of the period over which a reduction in class size is necessary to obtain the measured benefit. Non-experimental data report usually class size measures at discrete points in time only. For our benefit calculations above, we have assumed that class size remains roughly constant for the last two years of compulsory education.

However, there may be more persistence in class sizes. In this case, the cost of obtaining the class size benefits discussed above will be higher. To illustrate the empirical content of this point, we use information on the class size the pupil was exposed to 9 years earlier, at age 7. We regress class size at age 16 on class size at age 7, including all background variables of our most general specification (see column 4 in table 2). The coefficient estimate of class size at age 7 is 0.031, with a standard error of 0.0035. This estimate is small, but it does indicate some persistence in class size over the school curriculum.

### *What causes variation in class size?*

A key assumption throughout our analysis is that conditional on the rich set of regressors included in the regressions, the variation left in class size is exogenous in an educational performance equation. This raises the question: What causes the variation in the class size variable? Most papers using non-experimental data avoid discussing this crucial point.

We leave a detailed discussion and analysis for future work, but we want to address some possible reasons for class size variation. One hypothesis is that cohort sizes of school children change in different ways across regions and it takes time for local



authorities to adjust school resources. Depending on whether current cohorts are larger or smaller than past cohorts, class sizes may be larger or smaller during the adjustment period. Regressions of average class size on local authority level on the proportion of primary and secondary school children per 1000 inhabitants (obtained from the 1971 census) indeed suggest that class size and cohort size are positively related. We have not exploited the dynamics of this process, but results indicate that further investigation in this direction may be promising.

A second potential source for variation in class size is the size of the school. If schools are expanding or contracting, the number of teachers may adjust only partially. Moreover, the long term elasticity of the number of teachers with respect to the number of pupils may be smaller than one. We examine this by regressing the log of the number of teachers in a particular school on the log of total pupil enrollment, including local authority dummies. We find that the elasticity is smaller than 1 for each school type.

### *Conclusion*

The main result of our analysis is that class size effects on wages are present, but very small. They are unlikely to be detected in simple reduced form regressions when using data sets of moderate size. We illustrate this in our analysis where we concentrate on one particular, and maybe secondary rather than primary effect of class size - its impact on the school continuation decision. By using the information about the channel through which class size affects outcomes allows us to obtain more precise estimates of class size coefficients in wage equations. Although these effects are small, some simple back on the envelope calculations indicate that they accumulate to quite substantial numbers over the life cycle of individuals exposed to reductions.

The main conclusion to be drawn from our analysis is that we should be very cautious in accepting outcomes of non-experimental studies that state that there are no class size effects on wages or other outcomes. This may be misleading; as we show, low precision of class size estimates in data sets of moderate size, using reduced form specifications, does not allow detection of these effects. This however does not imply that effects are zero.

## References

- Becker, G.S (1981). *A Treatise on the Family*. Cambridge: Harvard University Press.
- Behrman, J.R. and P. Taubman (1986). ‘Birth Order, Schooling and Earnings.’ *Journal of Labor Economics*, 4, 121-45.
- Betts, J. (1995). ‘ Does School Quality matter? Evidence from the National Longitudinal Survey of Youth.’ *Review of Economics and Statistics*, 231-250.
- Blundell, R., Dearden, L. and C. Meghir (1996). ‘ The Determinants and Effects of Work Related Training in Britain.’ Institute for Fiscal Studies, London.
- Card, D. and A.B. Krueger (1992). ‘Does School Quality Matter?.’ *Journal of Political Economy*, 100, 1-40.
- Card, D. and A.B. Krueger (1996a). ‘ School Resources and Student Outcomes: An Overview of the Literature and New Evidence from North and South Carolina.’ *Journal of Economic Perspectives*, 10, 31-51.
- Card, D. and A.B. Krueger (1996b). ‘ Labor Market Effects of School Quality: Theory and Evidence.’ *NBER*, Working Paper No 5450.
- Card, D. and T. Lemieux (2001). ‘Can Falling Supply Explain the Rising Return to College Education for Younger Men? A Cohort-Based Analysis’, *Quarterly Journal of Economics*, 2001, 705-746.
- Coleman, J.S., et al (1966). *Equality of Educational Opportunity* Washington DC, US GPO.
- Currie, J. and T. Duncan. *Early Test Scores, Socio-Economic Status and Future Outcomes*, mimeo, UCLA.

- Dearden, L., J. Ferri and C. Meghir (2000). ‘The Effect of School Quality on Educational Attainment and Wages.’ *mimeo*, Institute of Fiscal Studies, London.
- Dolton, P. and A. Vignoles (2000). ‘The Impact of School Quality on Labour Market Success in the United Kingdom.’ *mimeo*, University of Newcastle. *Bulletin of Economic Research* (forthcoming).
- Ermisch, J. and M. Francesconi (2001). ‘Family Matters: Impacts of Family Background on Educational Attainments.’ *Economica*, 68, 137-156.
- Feinstein, L. and J. Symons (1999). ‘Attainment in Secondary School.’ *Oxford Economic Papers*, 51, 300-321.
- Goldhaber, D. D. and D. J. Brewer (1998). ‘Why don’t Schools and Teachers Seem to Matter?’ *Journal of Human Resources*, 32, 505-523.
- Hanushek, E.A. (1992). ‘The Trade Off Between Child Quantity and Child Quality.’ *Journal of Political Economy*, 100, 84-117.
- Hanushek, E.A. (1996). ‘Measuring Investment in Education.’ *Journal of Economic Perspectives*, 10, 9-30.
- Hanushek, E.A. (2001). ‘The Failure of Input-based Schooling Policies.’ *mimeo*, Stanford University.
- Hanushek, E.A., S. G. Rivkin, L.L. Taylor (1996). ‘Aggregation and the Estimated Effects of School Resources.’ *Review of Economics and Statistics*, 78, 611 – 627.
- Harmon, C. and I. Walker (2000). ‘Selective Schooling, School Quality, and Labour Market Returns.’ *Economica*, 67, 19-36.

- Heckman, J., A. Layne-Farrar, and P. Todd (1996). ‘Human Capital Pricing Equations with an Application to Estimating the Effect of School Quality on Earnings.’ *The Review of Economics and Statistics*, 78, 562-610.
- Hoxby, C.M. (2000). ‘The Effect of Class Size on Student Achievement: New Evidence from Population Variation.’ *Quarterly Journal of Economics*, 115, 1239-1286.
- Johnston, G. and Stafford F. (1973). ‘Social Returns to Quantity and Quality of Schooling.’ *Journal of Human Resources*, 8, 139–155.
- Krueger, A. (2001). ‘Economic Considerations and Class Size.’ mimeo, Princeton University.
- Lazear, E. P. (2001). ‘Educational Production.’ *Quarterly Journal of Economics*, 116(3), 777-805
- Mincer (1974), *Schooling, Experience and Earnings*, New York: Columbia University Press.
- Micklewright, J. (1986). ‘A Note on Household Income Data in ncds3.’ NCDS user support working paper, 18, City University, London.
- Micklewright, J. (1989). ‘Choice at 16.’ *Economica*, 56, 25-39.
- Micklewright, J., Pearson, M., and Smith, R. (1989). ‘Has Britain Got an Early School Leaving Problem?’ *Fiscal Studies*, 10, 1-16.
- Pradhan, M. and A. van Soest (1995). ‘Formal and Informal Sector Employment in Urban Areas in Bolivia.’ *Labour Economics*, 2, 275-298.
- Rice, P.G. (1987). ‘The Demand for Post-Compulsory Education in the UK and the Effects of Educational Maintenance Allowances.’ *Economica*, 54, 465-476.

- Robertson, D. and J. Symons (1996). ‘Do Peer Groups Matter? Peer Groups versus Schooling Effects on Academic Attainment.’ *mimeo*, London School of Economics.
- Todd, P. E. and K. I. Wolpin (2001). ‘On the Specification and Estimation of the Production Function for Cognitive Achievements.’ *mimeo*, University of Pennsylvania.

# Appendix

**Available upon request from the authors; not to be included in the paper**

This appendix contains some details on how the standard errors on the effects of class size via the staying on decision on wages are computed. The model assumptions are as follow:

**Probit model for staying on:**

$$y^* = x_1\theta_1 + \epsilon_1 \tag{2}$$

$$\epsilon_1|x_1 \sim N(0, 1) \tag{3}$$

$$y = 1 \text{ if } y^* > 0 \text{ (stay on); } y = 0 \text{ if } y^* < 0 \text{ (leave school)} \tag{4}$$

**Regression model for log wage  $w$ :**

$$w = x_2\beta + \gamma y + \epsilon_2 \tag{5}$$

$$E[\epsilon_2|x_1, x_2, \epsilon_1] = 0 \tag{6}$$

Probit on the staying on equation gives estimates of  $\theta_1$ ; OLS on the wage equation gives estimates of  $\theta_2$  and  $\gamma$ . It is immediately clear that these estimates together form the pseudo ML estimates for  $\theta = (\theta_1, \beta, \gamma) = (\theta_1, \theta_2)$ , with  $\theta_2 = (\gamma, \beta)$ :

$$\hat{\theta} = \operatorname{argmax} L(\theta) \tag{7}$$

$$L(\theta) = L_1(\theta_1) + L_2(\theta_2) \tag{8}$$

where  $L_1$  is the probit log likelihood and  $L_2$  is the pseudo log likelihood part of the regression equation with contribution of observation  $i$  given by  $-0.5(w_i - x_{2i}\beta - \gamma y_i)^2$ .

Say  $\theta_{11}$  is the parameter on class size in the probit equation. We need the standard error on the estimate of  $\theta_{11}\theta_{21}$ . This is a function of the parameter vector  $\theta$ , and the standard error can thus be approximated using the delta method from the (estimator of the) asymptotic covariance matrix of the complete pseudo ML estimator  $\hat{\theta}$ . According to the general theory of extremum estimators, this asymptotic covariance matrix is given by  $(E[\frac{\partial^2 L(\theta)}{\partial\theta\partial\theta'}])^{-1}E[-\frac{\partial L(\theta)}{\partial\theta}\frac{\partial L(\theta)'}{\partial\theta'}](E[\frac{\partial^2 L(\theta)}{\partial\theta\partial\theta'}])^{-1}$ , computed at the true value  $\theta^0$  of  $\theta$ .

Obviously, due to the additive structure of the pseudo log likelihood, the Hessian is block diagonal. The main issue here is to show that the outer product is also block diagonal. If this is the case, then the estimators of  $\theta_1$  and  $\theta_2$  are asymptotically independent (and applying the delta method is straightforward). The off diagonal part of the outer product matrix is given by:

$$E_{\theta^0}\left[-\frac{\partial L(\theta)}{\partial\theta_1}\frac{\partial L(\theta)'}{\partial\theta_2}\right] = \tag{9}$$

$$= E_{\theta^0}\left[-\frac{\partial L_1(\theta_1)}{\partial\theta_1}\frac{\partial L_2(\theta_2)'}{\partial\theta_2}\right] \tag{10}$$

The first factor is the probit score, the second is the score of the regression pseudo log likelihood, given by  $(w - x_2\beta - \gamma y)(y, x_2)'$ . Using the law of iterated expectations, we can rewrite the whole thing as the expected value over  $x_1, x_2$  and  $\epsilon_1$  of the conditional expectation given  $x_1, x_2$  and  $\epsilon_1$  of a function of the form:

$$E_{\theta^0}[g(x_1, x_2, y)(w - x_2\beta - \gamma y)] \tag{11}$$

$$E_{\theta^0}[g(x_1, x_2, y)\epsilon_2] \tag{12}$$

where  $g$  is some matrix valued function of  $x_1, x_2, y$ , capturing the probit score and the part  $(y, x_2)'$  of the regression model score. The assumption (6) in the regression model that the error term has conditional mean zero (conditional also on the staying on decision, i.e.,



the errors in the two equations are unrelated; this is the essential part of the assumption!) now guarantees that the conditional expectation given above is zero, and this proves that the asymptotic covariance matrix of  $\theta$  is block diagonal, where the diagonal blocks are the separate covariance matrices of the probit estimates for  $\theta_1$  and the OLS (pseudo ML) estimates for  $\theta_2$ . (The White covariance matrix would fit best into the framework as described here; the standard OLS covariance matrix makes stronger assumptions which we do not need in the pseudo ML framework.)

It is now straightforward to apply the delta method to the function  $f(\theta) = \theta_{11}\theta_{21}$ . This function has partial derivative  $\theta_{21}$  with respect to  $\theta_{11}$  and has partial derivative  $\theta_{11}$  with respect to  $\theta_{21}$ . If  $se(\hat{\theta}_{11})$  is the estimated standard error on the probit estimate of  $\theta_{11}$  and  $se(\hat{\theta}_{21})$  is the estimated standard error on the OLS estimate of  $\theta_{21} = \gamma$ , the delta method together with the block diagonality result imply that the standard error of  $f(\hat{\theta})$  can be approximated by:

$$\hat{se}(\hat{\theta}_{11}\hat{\theta}_{21}) = \sqrt{\hat{\theta}_{11}^2 [se(\hat{\theta}_{21})]^2 + \hat{\theta}_{21}^2 [se(\hat{\theta}_{11})]^2} \quad (13)$$