

Class Size Effects on Early Labour Market Outcomes

by

Beatrice Schindler Rangvid*

Eskil Heinesen†

Version May 2006

Abstract: *We estimate class-size effects and effects of students per teacher hour on employment and wages of labour market entrants. Our identification approach (using discontinuities in the relationship between grade enrolment and school resources) can eliminate bias due to endogeneity of school resources both between schools and grades and within grades. We find no significant effects of school resources for labour market entrants. The insignificance of the estimated effects applies across a variety of different model specifications. However, a tentative analysis on outcomes later in the career suggests that there might be effects on wages for more experienced workers.*

Keywords: class size, employment rate, wages, regression discontinuity design

JEL-code: I2, J23, J31

The Danish Social Science Research Council is gratefully acknowledged for its support of this project, both directly and through its support of the Graduate School for Integration, Production and Welfare. We thank Amelie Constant, Peter Jensen, Nina Smith, Robert Wright, and participants of the ESPE 2005 conference for helpful comments and suggestions. An earlier version of this paper circulated with the title "Class Size Effects and Early Career Employment". The data used in the analysis are available for other researchers.

*Corresponding author: akf, institute of local government studies, Nyropsgade 37, DK-1602 Copenhagen V, Denmark. Phone: (+45) 3311 0300, fax: (+45) 3315 2875, and e-mail: bs@akf.dk.

†akf, institute of local government studies, Nyropsgade 37, DK-1602 Copenhagen V, Denmark.

1 Introduction

Common sense suggests that reducing class size may be a means of improving the quality of education. Reducing the number of students in a class has the potential of affecting how much is learned in a number of different ways, including (i) less disruptive behaviour in the classroom¹, (ii) enabling the teacher to employ different teaching methods², (iii) students receiving more individualized attention, and (iv) teachers having closer contact with parents. It is a convenient policy instrument as it is a parameter parents care a lot about and as it is one of the simplest variables to manipulate. However, student outcomes are dependent on many different factors and the existing evidence of the effectiveness of reducing class size on student outcomes is mixed (see, e.g., Hanushek 1996, and Krueger 2003).

In the present study we analyse effects of class size and the number of students per weekly teacher hour on early career employment and wages. To eliminate bias due to the endogeneity of school resources both between schools and grades, we use a regression-discontinuity design with discontinuities induced by administrative rules as instruments.

The number of students per weekly teacher hour is not equivalent to class size due to variation in the number of weekly lessons and in the number of lessons with more than one teacher. For instance, compensatory allocation of extra teacher hours to larger classes implies that the number of students per teacher hour varies less than class size. If it is the amount of teacher resources per pupil, and not the number of students in the classroom that matters for outcomes, compensatory teacher hours will lower the estimated class size effects, and thus bias the class size effect downward. Below we motivate why it is important to analyse effects on early career outcomes such as employment and wages and why it may be important to use instruments for class size and other measures of school resources.

The overwhelming part of the class size evaluation literature examines the impact of class size on basic skills in, for instance, reading and math measured by test scores. However, economists have long argued that a major role of elementary and secondary education is to prepare students for participation in the labour market, and that the effectiveness of class-size reductions should also be measured by the impact on labour market outcomes such as earnings and employment. Class size may influence outcomes other than test scores, because smaller classes are compatible with a learning environment and teaching methods that may also enhance other skills such as teamwork, creativity and leadership skills that may be valued in the labour market and thus have impact on labour

¹See Lazear (2001).

²See Borghans (2002); Betts and Shkolnik (1999); Blatchford and Martin (1998); and Rice (1999).

market outcomes later in life. Zahorik et al. (2000) argue that possible positive effects of reducing class size may work through more instruction time, less need of discipline, more teacher enthusiasm, more hands-on activities, more class participation, and more student self-direction and responsibility. In big classes, to establish and maintain order in the classroom, most of the class time is typically devoted to activities that can be easily monitored and which keep most students busy, such as lecture and seatwork. When class size drops, a change of balance between different instructional methods becomes feasible³.

A few recent studies examine class-size effects on outcomes after compulsory school such as educational attainment, wages and unemployment experiences, see Dearden et al. (2002), Dustmann et al. (2003), Browning and Heinesen (2005), and Boozer and Maloney (2001). Dustmann et al. find significant negative effects of increasing class size on the probability of staying on at school at age 16, and this decision in turn affects wages later in life, whereas Dearden et al. using the same data set (but unlike Dustman et al. estimating reduced form regressions) find no effects on educational qualifications or male wages, but negative effects on female wages (at age 33). Browning and Heinesen find negative and (marginally) significant class-size effects on years of education. Boozer and Maloney find statistically and economically significant negative effects of class size on both test score growth and on early adult outcomes such as completed education and unemployment.

The key methodological challenge when estimating causal effects of class size is the potential endogeneity of the class size variable. It has long been recognized that class size may be endogenous to outcomes due to such different sources as parents choosing schools by choosing their residences, and educators at the school level pursuing compensatory policies by placing students who have specific problems or are simply doing poorly in smaller classes. Recent findings by West and Wössmann (2005) indicate substantial compensatory sorting of weaker students into smaller classes in many countries. Thus, endogeneity may severely bias naive estimates of the class-size effect.

The majority of studies estimating class-size effects has either ignored endogeneity issues or tried to take account of them by controlling for a large number of important covariates and/or by using measures of class size at a high level of aggregation (for instance, using average class size at school level controls for endogeneity related to within-school allocation decisions). In more recent work, a variety of other approaches has been used to

³Borghans (2002) argues that a possible explanation for the often observed missing class-size effect on test scores in the education literature might be that, generally, when class size is reduced, the time used for class teaching decreases while the time for individual instruction increases. Since each pupil only receives a fraction of the individual instruction provided by the teacher, the total teaching time per pupil may decrease. And, as a consequence, achievement measured by test scores may also go down. However, if individual instruction stimulates non-cognitive skills more than class teaching (and if these are valued in the labour market), this change in the instructional regime might affect labour-market outcomes positively, even though cognitive skill acquisition may be detrimentally affected.

deal with endogeneity problems. Ideally, we would like to conduct explicit experiments, randomly assigning students and teachers to class rooms of different sizes. In his study of the Project STAR experiment, Krueger (1999) finds significant positive effects of smaller classes on test scores in the early grades. However, in practice, experiments also have disadvantages (see Hoxby, 2000; Todd and Wolpin, 2003; and see also Heckman et al., 1999, for a related discussion of advantages and disadvantages of using experiments in the evaluation of labour market programmes), and they are out of reach of most researchers. Other methods to control for endogeneity include value-added models and quasi-experimental methods. Studies estimating value-added models, where achievement is related to prior achievement, do, on the whole, not find strong evidence in favour of class-size reduction in US data (Hanushek, 1997). However, value-added models impose strong assumptions on the underlying production technology and they are highly susceptible to endogeneity bias, see Todd and Wolpin (2003). Studies using quasi-experimental methods have employed exogenous population variation (Hoxby, 2000), and maximum class-size rules (Angrist and Lavy, 1999). Hoxby rules out even small class-size effects on test scores, whereas Angrist and Lavy find significant effects for 5th graders, but not for 3rd or 4th graders.

The present study basically follows the identification approach used by Angrist and Lavy (1999) who use variation in school enrolments combined with an administrative maximum class-size rule that governs class size as a function of enrolment. This rule does not predict the actual class size perfectly because many factors affect class size, but, as Angrist and Lavy show, predicted class sizes fit the actual data reasonably well, as average class size clearly displays a "saw-toothed pattern" induced by the rule. This "regression-discontinuity design" is different from conventional IV methods because the instruments are derived from non-linearities or discontinuities in the relationship between the regressor of interest and a control variable. The analysis of this paper differs in several important respects from that of Angrist and Lavy (1999): Our outcome measure is not test scores, but early career employment and wages; we estimate effects of both class size and the number of students per weekly teacher hour; we use micro data for student outcomes and parental background (whereas Angrist and Lavy use class averages); we use a very large set of controls for family background at the individual level and socioeconomic conditions at the local authority level (whereas Angrist and Lavy use only one control variable, namely the percentage of students from disadvantaged backgrounds at school level); finally, using Danish data we estimate effects of reducing class size from an average level of 20, compared to an average level of 32 in the Israeli data set used by Angrist and Lavy. The regression discontinuity approach can eliminate bias due to endogeneity of school resources between schools and grades, and between classes within grades.

We use the same panel data set as in Browning and Heinesen (2005). Data are merged

from several administrative registers by use of personal civil registration numbers and cover all eight cohorts of 8th graders in public schools in Denmark for the period of 1985-92. Our results suggest that smaller class sizes do not seem to produce benefits that translate into significantly higher early career employment or wages. However, a tentative analysis on outcomes later in the career suggests that there might be effects of class size and teacher resources on wages for more experienced workers.

The paper proceeds as follows. The next section presents the empirical framework. Section 3 presents the data set and descriptive statistics. Section 4 reports estimation results. The last section concludes.

2 The empirical framework

To identify causal effects of school resources on student outcomes, we use discontinuities, induced by administrative rules, in the relationship between school resources and enrolment. We use the fact that when grade enrolment crosses a critical value, this will trigger the establishment of another class in that grade, causing average class size to fall abruptly. This means that at enrolment counts just before the discontinuity point, average class size in the grade will be large, while at enrolments just after the discontinuity points, it will be small. We use the variation around these 'discontinuity points' to instrument class size and students per teacher hour.

The outcome variables studied in this paper are the mean employment rates⁴ and the log of mean hourly wages in the first three years after ended education, i.e. the fraction of this three-year period the person is employed, and the mean hourly wage earned during this period. We discuss these variables in more detail in section 3. Our basic model for estimating the causal effect of class size on employment and wage is given by

$$y_{iskt} = X_{iskt}\beta_1 + X_{skt}\beta_2 + X_{kt}\beta_3 + X_t\beta_4 + \alpha n_{iskt} + \varepsilon_{iskt} \quad (1)$$

$$n_{iskt} = X_{iskt}\gamma_1 + X_{skt}\gamma_2 + X_{kt}\gamma_3 + X_t\gamma_4 + \gamma_5 D_{-ivskt} + \zeta_{iskt} \quad (2)$$

where: y_{iskt} is the employment rate/the log of hourly wages for individual i of cohort t in school s of municipality k ; the cohort index t indicates the year in which student i started in 8th grade; X_{iskt} are characteristics of student i (education, gender and family background); X_{skt} are characteristics of cohort t in school s in municipality k (dummies for enrolment segment and variables for mean parental background); X_{kt} are characteristics of municipality k for cohort t (e.g., the unemployment rate in years relevant for cohort t); X_t

⁴The employment rate measures the percentage of the years where the individual is employed (and thus is not unemployed, enrolled in education or otherwise not active in the labour market).

is a dummy for cohort t ; n_{iskt} is average class size for cohort t in school s of municipality k ; $D_{iv_{skt}}$ is a set of dummies indicating whether enrolment is just to the left or just to the right of the enrolment discontinuity points; and ε_{iskt} and ζ_{iskt} are error terms. The model for estimating the effect of the number of students per weekly teacher hour, h_{iskt} , on the employment rate is similar: In (1) and (2) n_{iskt} is replaced by h_{iskt} , while the set of instruments remains the same as the discontinuities arise at the same enrolment counts for both school resource measures. We do not include both class size and the number of students per teacher hour simultaneously in the model since it turns out that these variables are too highly correlated in our data set to be entered jointly.

In the following presentation, for brevity, we focus on estimation and interpretation of the model with class size as school resource measure, but the interpretation is similar and valid for models using the number of students per teacher hour instead. The average causal effect of changes in class size, indicated by α , is the parameter of main interest. Todd and Wolpin (2003) suggest that this causal effect should be understood as a "total policy effect", because it embodies parental responses to different class sizes. For instance, parents whose children attend smaller classes might invest less of their time in their children's education, or just the opposite, they may see the small class as an incentive to promote learning even more. Another important point is that the interpretation of the causal effect α is the "local average treatment effect" (LATE) interpretation. The LATE interpretation differs from the average treatment effect on the treated (ATET) of a randomized experiment if those who are sensitive to the instrument (i.e. those whose treatment status is sensitive to variation in the instrument) are not a representative subsample of the total population (see Angrist and Lavy, 1999; and Angrist and Imbens, 1995). For the parameter α to be identified by instrumental variables techniques, a necessary assumption is that wage/employment outcomes and instruments are only correlated because of the correlation between instruments and class size (conditioning on the X variables).

We instrument class size with dummies indicating the discontinuity in the relationship between enrolment and class size which is triggered by a maximum class-size rule as in Angrist and Lavy (1999) and Browning and Heinesen (2005). In Denmark, national law sets the maximum class-size permitted at 28 students. However, municipalities may have their own administrative rules determining class sizes (as long as this does not exceed the maximum of 28). In practice, some municipalities have a formal maximum class-size rule of e.g. 23, 24 or 25 applying to all grades, but even municipalities without formal maximum class-size rules will usually be cautious to fill classes up to the maximum of 28 in order to have some scope to accommodate a few additional children in the classes without being forced to add another class.

Thus, what we utilize in our main identification strategy is that, in practice, an addi-

tional class is typically added as soon as enrolment exceeds 24, 48 and higher multiples of 24. Figure A1 in the appendix displays class size as predicted by this maximum class-size rule, and observed average class size by enrolment. As can be seen, the maximum class-size rule of 24 fits the data well. For students per teacher hour, we use an administrative rule for the largest municipality, Copenhagen; see Browning and Heinesen (2005) for details. Figure A2 in the appendix shows the observed average number of students per teacher hour for each value of enrolment and fitted values from a regression of students per teacher hour on the instrument based on the Copenhagen administrative rule. It will be seen that (due to compensatory allocation of extra teacher hours to larger classes), the drop in the number of students per teacher hour at the discontinuity points is less sharp than for class size. Figures A1 and A2 suggest also that predicted school resources fit observed resources quite well, especially for enrolments below 100. Therefore, in the empirical analysis we restrict the sample to the lowest four discontinuity points.

Figure 1 shows the drop in average class size and students per teacher hour between the intervals just below and just above the four smallest discontinuity points for five different discontinuity samples ranging from a rather broad interval of ± 5 enrolment counts around the discontinuity points to the narrowest possible sample of ± 1 students around the discontinuity points. Average observed school resources are higher in a small interval to the left of these points than in a small interval to the right at the four lower discontinuity points, and especially for the two lowest points. This is a necessary condition for the regression-discontinuity design to be able to identify the parameters of interest. As is seen, the drop in average class size is largest for the broadest sample (± 5), and decreases steadily for more narrow samples (except for the discontinuity point at enrolment count 96)⁵. Ideally, one would prefer narrow definitions of the estimation subsamples around the discontinuity points, e.g. the ± 1 discontinuity sample including only students in school/cohort cells with enrolments in 24/25, 48/49, 72/73 and 96/97, but in our data, observed average school resources do not consistently drop just at the discontinuity points. For this reason, we have chosen to focus on discontinuity samples of ± 5 , ± 4 , and ± 3 around the discontinuity points. Also, we see, that the drop in average class size is largest at the lowest two discontinuity points at enrolments of 24 and 48⁶. Therefore, while the main

⁵The school resource variation around the discontinuity point at an enrolment of 96 is different from the three lowest discontinuity points both because the average school resource difference between the treatment and control group is similar for all five discontinuity samples (i.e. for ± 5 , ± 4 , ± 3 , ± 2 , and ± 1 samples), and because the difference is greater than at the third lowest discontinuity point. These patterns can also be seen from Figures A1&A2.

⁶This is not surprising given the maximum class-size rule: at the smallest discontinuity point, when enrolment changes from 24 to 25, expected average class size drops from 24 (one class) to 12.5 ($=25/2$; two classes) - a drop of 11.5 students. At the next discontinuity point, when enrolment increases from 48 to 49, expected average class size drops from 24 ($2 \times 24 = 48$) to 16.3 ($=49/3$; three classes) - a drop of only 7.7 students. Similarly, for the discontinuity points of 72 and 96, the drop is only $24 - (73/4) = 5.8$ and $24 - (97/5) = 4.6$ students.

analysis uses variation at all four lowest discontinuity points (i.e. around enrolments of 24, 48, 72, and 96), additional results using only the lowest 2 discontinuity points (for which the instruments are most powerful) are provided in section 4.3.3.

[Figure 1 about here.]

The instruments are entered as a set of dummy variables indicating treatment (just above the discontinuity point) and control status (just below the discontinuity point) for each discontinuity segment separately (e.g., for the ± 3 discontinuity sample, $z1 = 1$ for $25 \leq e \leq 27$ and otherwise $z1 = 0$; $z2 = 1$ for $49 \leq e \leq 51$ and otherwise $z2 = 0$; etc.). We include different dummy instruments for each segment because the expected class-size reduction at the discontinuity points is larger for lower segments than for higher segments according to the expected class-size function. Moreover, enrolment controls are entered as dummy variables for each enrolment segment (e.g., for the ± 3 discontinuity sample, $d1 = 1$ for $22 \leq e \leq 27$ (otherwise $d1 = 0$), $d2 = 1$ for $46 \leq e \leq 51$ (otherwise $d2 = 0$), etc.). Thus, the only information from the instruments used is that expected class size (or students per teacher hour) is smaller to the right of the discontinuity points than to the left of these points. This specification focuses explicitly on the variation at the discontinuity points and thus highlights the quasi-experimental identification strategy of the regression-discontinuity design: When enrolment is to the right of the discontinuity points we have a high probability of treatment (the expected class size is small), otherwise a low probability of treatment (the expected class size is large)⁷. This is similar to the Wald estimator discussed in Hahn et al. (2001). Yet, in our analysis there are multiple discontinuity points and also additional controls.

As the number of students enrolled in a grade may be related to student outcomes for reasons beyond the effects of varying school resources, the key identifying assumption is that the X variables adequately control for any other effects of enrolment on outcomes. This assumption may not seem undue, as we have a broad set of controls for family background and socioeconomic conditions in the municipalities (see section 3). In addition, we include a set of dummy indicators for enrolment discontinuity segments to help ensure that the estimated class-size effect is identified by the discontinuities in the instrument alone. Moreover, enrolment controls are entered as dummy variables for each enrolment segment (e.g., for the ± 3 discontinuity sample, $d1 = 1$ for $22 \leq e \leq 27$ (otherwise $d1 = 0$), $d2 = 1$ for $46 \leq e \leq 51$ (otherwise $d2 = 0$), etc.).

⁷This is a 'fuzzy' regression-discontinuity design meaning that the rule assigning treatment status is not deterministic, but that it affects the probability of being treated. Yet, treatment is not guaranteed.

Another identifying assumption is the conditional independence assumption: Treatment of student i is independent of α_i conditional on enrolment (and other covariates) near discontinuity points, see Hahn et al. (2001), i.e. parents do not systematically use the maximum class-size rules to place their children in schools with high school resource levels (in i.e. small classes or few students per weekly teacher hour). It is not unreasonable to assume that this holds, as one cannot perfectly predict whether enrolment will be just above or just below a certain cutoff point. Even if there were, it would be difficult for parents to find an alternative school for their children due to rather restricted school choice options within the public sector, and only a rather small private school sector in Denmark in the eighties and early nineties.

While the wage rate distribution is approximately symmetric (but with a slightly fat right tail), unsurprisingly, the distribution of the employment rate has a high frequency of observations of full employment (44%) suggesting the use of Tobit models. Because the employment rates are bounded between 0 and 100 (in percentage terms), an appropriate model is the two-limit version of the Tobit model. Let Y_i^* be the underlying latent variable: $Y_i^* = \beta X_i + \varepsilon_i$. The observed variable Y_i is equal to 0 if $Y_i^* < 0$, Y_i is equal to 100 if $Y_i^* > 100$, and Y_i is equal to Y_i^* if $0 \leq Y_i^* \leq 100$. Thus, the distribution that applies to the sample data is a mixture of discrete and continuous distributions. For the instrumental variables Tobit specification, Newey (1987) proposed an estimator based on Amemiyas Generalized Least Squares (AGLS) and this is the estimator used here including the proposed adjustment of the standard errors therein⁸. We estimate the model by two-stage least squares correcting the standard errors for intra-school correlation in employment outcomes (Moulton, 1990).

3 Data, descriptive statistics and instruments

We use Danish micro data based on administrative registers for the full sample (about 453,000 persons) of the eight cohorts who finished 8th grade in public schools in the summer of 1985-92, and their parents⁹. We have yearly panel data for the period 1985-2002. These micro data on employment and controls are linked with data on enrolment and school resources for each grade at each school, and data on socioeconomic conditions in the municipalities. For each year we have data on about 1,100 public schools. First, we discuss data on hourly wages and the employment rate which are used as outcome variables; next we discuss data on school resources, and finally data on control variables.

⁸To implement this estimator, we use the Stata Version 9 command "ivtobit".

⁹We do not have data on school resources for private schools, and must therefore exclude private school students from the analysis.

Outcomes: Wage and Employment rate We measure the impact of class size on hourly wages and the employment rate for entrants in the labour market. We observe the eight cohorts of students in our sample up to 10 to 17 years after the year they attended 8th grade. For almost all of the students, this corresponds to age 25-32. The reason why we restrict the main analysis to measuring outcomes for labour market *entrants* is because the individuals in our sample are still quite young. This constraint is important, particularly for individuals with a tertiary education who, naturally, are older when they enter the labour market. Also, in an international comparison, individuals in Denmark are on average older than their counterparts in many other countries before completing their education and being ready to enter the labour market: at age 25, 35% of the Danish population is still enrolled in education, compared to 22% in the UK and only 16% in the US (OECD, 2005).

Our data from administrative registers is particularly well suited for the purpose of studying entrants as it is a panel with yearly observations on both ongoing and completed education. We measure the outcomes of labour-market entrants over a three-year period after ended education. Thus, for each person we use information on the three years following the last completed education spell we observe (given that re-enrolling for another education spell does not take place).

We use as the outcome measure the wage and employment rates during the first three years after ended education, i.e. the outcome measures for each person are the log of the 3-year-average hourly wage rate (measured at the wage level of year 2000 using a general wage index) and the 3-year-average fraction of the year the person was employed in the labour market, respectively. The requirement of a three-year spell in the labour market without re-enrolling results in dropping about 81,000 out of 453,000 cases (leaving 372,000) - or almost 18% of the sample, and it has the implication that older cohorts are more likely to be in the estimation sample, a feature which we take account of by including cohort dummies in the estimations. We lose another 9,000/77,000 observations (leaving 363,000/295,000) for the employment/wage regressions due to missing information on the employment/wage rate¹⁰. The estimation sample for the wage analysis is somewhat smaller than the estimation sample for the employment outcome, because - be nature - only the employed have information on hourly wages^{11,12}. The number of observations referred

¹⁰We require at least two years out of the three-year-spell to have valid employment/wage information. Otherwise the individual is not included in the employment/wage regressions. Individuals with more than one year (out of three) of employment data missing are mainly individuals who are permanently inactive in the labour market, either because they have chosen so (housewives), or because they are permanently unemployed (social assistance or unemployment benefits).

¹¹Moreover, hourly wages from self-employment are generally unreliable, and self-employed are therefore not included in the analysis.

¹²The hourly wages are information constructed/estimated from registerdata. A non-trivial share of the wage information is of low reliability. As the results are insensitive to the inclusion of observations with

to above are for the full samples of the employment/wage analyses; the corresponding discontinuity samples are of course smaller (see sections 3 and 4 for details).

[Table 1 about here.]

The distribution of the students by 8th grade cohort in our data is shown in Table 1 which also shows descriptive statistics for the other variables in the analysis. While 16% of the total number of students included in the main estimation sample come from the oldest 8th grade cohort (1985), only 8% come from the youngest cohort (1992) compared to 14% (1985) and 11% (1992) in the full sample. Table 2 shows the distribution of completed education in 2002 for the eight cohorts. Clearly, there are more lower educated individuals in the younger cohorts. Yet, while older cohorts are overrepresented in the estimation sample, both grade enrolment, class size and student per teacher hour are very close to the levels in the full sample (see Table 1). Also, the share of immigrants from the 1st and 2nd generation are virtually the same, as is the male/female ratio. Differences between the two samples become apparent when we look at parental background, as the young adults from more disadvantaged backgrounds are slightly overrepresented in the estimation sample (lower parental education and employment). This is not surprising, as young adults with low levels of education are overrepresented in the estimation sample, and parental socioeconomic status tends to be correlated with child education level. However, as Tables 1 and 2 also show, for the three oldest cohorts (1985-87), the summary statistics do virtually not differ for the full sample for 1985-1992 and the estimation sample for these cohorts. Therefore, in addition to results for the full sample, we also present results for the three oldest cohorts only (section 4.3.1) to make sure that our results are not driven by the use of a selective sample.

[Table 2 about here.]

School resources: Class size and students per teacher hour Average class size in Danish primary and lower secondary schools of 19-20 students is low in an international perspective. In the comparisons of OECD (2002) only Luxemburg and Iceland have average class sizes lower than Denmark, and with respect to students per teacher, Denmark has the second lowest ratio after Italy. Therefore, one may expect relatively low marginal effects of reducing class size in Denmark compared to most other countries.

inferior quality (see section 4.3.2), we have chosen to keep them in the main estimation sample, because dropping them results in a reduction of the number of observations of about 50,000.

In our data set class size, teacher hours and enrolment are recorded at the beginning of the school year (September) and collected by the Ministry of Education. The class-size measure is not actual class size, but school-by-grade average class size, as we have only information on enrolment by grade for each school and the number of classes by grade. The same applies to our other school resource measure, students per teacher hour. This is not a problem for the analysis since the instruments predict school-by-grade average class size and pupils per teacher hour given enrolment.

We have data for each person on which school he/she attended in the 8th grade, but we have no such information for earlier grades. In the statistical analysis, we therefore only use data for school resources at the 8th grade. However, class size at the 8th grade will for most students be highly correlated with class size at earlier grades¹³.

Table 3 shows summary statistics for school level data for enrolment, classes and teacher hours at the 8th grade for all public schools in Denmark in the school years starting in the summer of 1985-1992. The averages for enrolment, class size and students per weekly teacher hour are 52, 19.5 and 0.54, respectively. The distribution of class size is rather concentrated around the mean with 10% of students in classes with less than 16 students and 10% of students in classes with more than 23 students. The distribution of students per teacher hour is even more concentrated around the mean due to the compensatory policy of allocating more teacher hours to larger classes.

[Table 3 about here.]

Controls In the estimations we use a broad set of controls. In all estimations, we control for cohort dummies and a set of dummies for enrolment segments. In estimations controlling for background characteristics we moreover include students' gender, ethnicity and the highest qualification obtained (based on both school and post-school qualifications), and family structure, housing conditions and parents' education, income, wealth, unemployment and labour-market status. Moreover, we control for average family background at the school/cohort cell (years of education, unemployment, and income). We also control for the unemployment rate in the municipalities. All family background variables are measured at the age 15 of the student. The municipal unemployment rate is measured for the year the student attended 8th grade. Descriptive statistics for the explanatory

¹³We have also information on grades 9 and 10, but we have chosen to use 8th grade resources as our measure, because (i) while most students attend the same school from preschool to 8th grade (and thus, for most students school resources in 8th grade are correlated with school resources in earlier grades), it is rather common to change school between 8th and 9th grade (mainly to private (boarding) schools), and because (ii) grade 10 is an optional extra year at lower secondary school and we therefore only have school resource information on a (selective) subsample of 10th graders.

variables in the two main estimation samples are shown in Table A1 in the Appendix (together with estimated parameters for all controls). The background characteristics for the wage estimation sample are slightly more advantaged. However, this is not surprising, because, as explained above, the wage sample is a subsample of the employment sample composed of those individuals who are in (salaried) employment.

4 Estimation results

In this section, we present results from the empirical analysis. Tables 4 and 5 show the main estimation results: Table 4 shows OLS and 2SLS results for the effect of class size and the number of students per teacher hour on the hourly wage rate, and Table 5 shows Tobit and IV-Tobit results for the effect on the employment rate. Each model is estimated with and without background controls for the three discontinuity samples (± 5 , ± 4 , ± 3) at the four lower enrolment segments around the discontinuity points of 24, 48, 72 and 96, and the OLS and Tobit models are also estimated for the full sample. The standard errors are corrected for potential within-school correlations of the error terms which is important since school inputs are measured at the school level (not individual level). For variables to be valid as instruments, a necessary condition is that they are significant predictors of the variables to be instrumented. The first-stage F-statistics displayed in the lower panels of Tables 4 and 5 test the hypothesis that the identifying instruments (i.e. the four dummies indicating treatment status - one for each discontinuity segment) do not enter the first-stage regression. The F-statistics indicate a strong correlation between the set of instruments and observed class size. For students per teacher hour this is also true for the ± 5 discontinuity sample, but there may be some indication of weak instrument problems for the ± 4 and ± 3 discontinuity samples. Thus, according to the literature on weak instruments (Bound et al., 1995; Staiger and Stock, 1997) a rule of thumb is that the F-statistic should be larger than 10, or at least larger than 5¹⁴. The low F-statistics in the students per teacher hour regressions are particularly due to the fact that the instruments for the third and fourth discontinuity points are rarely significant (results not shown), which underlines the need of additional analysis including only observations with enrolment around the first two discontinuity points. Such results are reported in section 4.3.3.

We focus on the effects of class size and teacher hours, but first we discuss briefly the parameter estimates of the controls. Table A1 in the appendix shows full results for the main models for wage and employment (the 2SLS-model for wage and the IV-Tobit for employment for the ± 5 discontinuity sample with class size as school resource measure).

¹⁴The F-statistic is calculated using robust standard errors adjusted for within-school and grade clustering.

In general, the estimates of the control variables are similar whether school resources are measured by class size or students per teacher hour, and whether the ± 5 , 4 or 3 discontinuity sample is used.

The dummy indicators for the discontinuity segments are not significant at the 5% level. The effect of the students' own educational level is highly significant: A higher educational level implies higher employment rates and wages as expected. Even controlling for the students' own education, many parameters of the family background characteristics are significant with the expected signs, especially in the employment regression. However, it is not surprising that some family background controls are insignificant, or have asymmetric effects on the two outcomes, when the very important intervening variables of students' own education are included. Excluding these variables increases the size and significance of parental background effects.

In the rest of this section we first report OLS and 2SLS results for the wage outcome, and then Tobit and IV-Tobit results for the employment outcome. Last, we present some sensitivity results and a tentative analysis of school resource effects on later labour market outcomes.

4.1 Early career (hourly) wage

Table 4 shows results for regressions where the log hourly wage rate (measured in the wage level of the year 2000 using a general wage index) during the individuals' first three years in the labour market is the dependent variable. Estimates that are significant at the 5% are in bold. Table 4 (and the following tables) only reports school resource effects. The upper panel of Table 4 displays results for simple least squares regressions of class size and students per teacher hour on the log of hourly wages. Only the full sample results without additional controls are significantly different from zero: class size is positively related to hourly wages, while students per teacher hour is negatively related to hourly wages. However, when controls are added, the estimates are numerically smaller and insignificant. Moreover, all estimates for the three discontinuity samples are insignificant, whether controls are added or not. However, as explained above, we are suspicious that the coefficients estimated by simple models may be biased due to the endogeneity of school resources.

[Table 4 about here.]

The lower panel of Table 4 shows results from estimations using the regression discontinuity design to account for endogeneity of the class size and students per teacher hour variables. Results are shown for the three discontinuity samples (as the specification only

uses observations around the discontinuity points, there are - by definition - no full sample results for the IV-regressions). Estimates from models both with and without background controls are negative and increase in absolute size the narrower the sample is around the discontinuity points. Parameters of models without background controls are all numerically larger than parameters of models with controls. However, none of the estimates are significantly different from zero.

4.2 Early career employment

Table 5 presents results for the models of the employment rate of labour-market entrants. The structure is similar to that of Table 4. The upper panel of Table 5 displays results from simple Tobit models (i.e. without correction for potential endogeneity of class size and student per teacher hour) for the full sample and for the three discontinuity samples. Estimates from models without controls indicate a positive and significant relation between class size and students per teacher hour on the one hand and employment rates on the other hand. Controlling for background characteristics reduces the coefficient estimates. However, the estimates are still positive for all regressions and remain significant throughout.

The lower panel of Table 5 shows results from the corresponding IV Tobit models. For models excluding background controls, point estimates for class size and students per teacher hour are consistently negative, but they are imprecisely estimated. Estimates from models controlling for background variables are positive for all but one out of six regressions, but insignificant for all estimations.

[Table 5 about here.]

To sum up, results from simple models display a significantly positive effect of larger classes and a higher pupil/teacher hour ratio on early career employment even after conditioning on background controls. However, taking account of potential endogeneity of the class size and teacher hours measures by using regression discontinuity techniques generally reduces the point estimates and none of the estimated coefficients is significantly different from zero. The regression above includes individuals' completed education as controls. However, the education level might also be influenced by school resources (see Browning & Heinesen, 2005). We have therefore rerun the main regressions without including the individuals' own education level as control. Yet, the results are virtually unchanged.

4.3 Sensitivity analyses

In order to assess the reliability of the results found in the main empirical analysis above, we present results from a number of different model specifications for the baseline 2SLS and IV-Tobit models from above. Furthermore, we present results from a tentative analysis of potential school resource effects on outcomes later in the labour market career.

4.3.1 Estimations using the three oldest cohorts only As argued in section 3 above, the estimation sample for all eight cohorts is rather selective, as students with shorter educations are more likely to be included in the estimation sample. Therefore, we also present results using the three oldest cohorts only, where selectivity is much less severe compared to the full sample. Restricting the sample to only three out of eight cohorts implies of course a significant drop in the number of observations: for example, in the ± 5 discontinuity sample, the number of observations is reduced from 115,956 to 52,118 for the wage regression, and from 142,700 to 63,546 for the employment regression, and the standard errors of the estimated school resource coefficients are substantially increased. Results for the main specifications (the 2SLS for wage and IV-Tobit for employment from sections 4.1 and 4.2) are shown in Table 6. The main results are unchanged: as for the 1985-1992 cohort estimation sample, there are no significant class size or teacher hour effects on early career employment or wage.

[Table 6 about here.]

4.3.2 Estimations excluding low quality wage information The variable for hourly wage is calculated by Statistics Denmark using data from administrative registers. Estimates for hourly wages are based on earned wages and an estimate of hours worked. For some persons, the number of hours worked is imprecisely estimated, and therefore, the estimate of the hourly wage rate is somewhat less reliable. Therefore, as a supplement for the hourly wage variable, Statistics Denmark provides a variable indicating the quality of each hourly wage observation. Statistics Denmark suggest a cutoff value of 50 for this quality indicator, where individual hourly wage observations with an indicator of less than 50 are considered of acceptable quality. In the main estimation sample in section 4.1, we have kept all observations irrespective of wage data quality, because dropping observation with low-quality wage data would result in a substantial reduction of the estimation sample (from 295,000 to 241,000 observations in the full dataset for wage estimations). However, before deciding to include all observations in the main estimation data, we made sure that

this would not qualitatively affect our main conclusions. We present the results of this sensitivity check in Table 7.

Table 7 presents results for the dataset without low quality wage observations for the main wage estimation model (2SLS) from section 4.1. The results suggest that while the precision of the estimates is slightly increased in spite of the reduced sample, and the predictive power (R-squared) in the estimations with controls is substantially increased (from around 0.25 to 0.35), dropping low quality wage observations from the estimation does not affect the conclusions on the school resource effects that remain insignificant throughout.

[Table 7 about here.]

4.3.3 Estimations using only the two lowest discontinuity segments As laid out in section 2 above, the drop in school resources at the two higher discontinuity segments (at enrolments of 72 and 96) is small and here the instruments seem rather weak, especially in predicting students per teacher hour. Since the instruments have stronger predictive power at the lower discontinuity points, we estimate the model for the two lowest segments of the discontinuity samples. This substantially reduces the number of observations (e.g. for the ± 5 discontinuity sample from 115,956 to 64,536 for the wage regression, and from 142,700 to 80,031 for the employment regression). Table 8 shows the estimation results. The larger F-statistics for the first stage regressions confirm that the instruments in the reduced sample have substantially stronger predictive power than when all four discontinuity points are used as in sections 4.1 and 4.2. Thus, despite the substantially reduced sample size, the standard errors of the estimated school resource coefficients are not increased very much. The point estimates in Table 8 are not systematically different from results in the main models. Qualitatively, the results from the reestimated main models remain unchanged: no effects are significantly different from zero.

[Table 8 about here.]

4.3.4 Later labour market outcomes The main analysis above is restricted to *early* labour market outcomes, because, as explained, many individuals (especially the higher educated) of the younger generations in our sample have not progressed far enough in their labour market careers to be observed at later stages. Therefore, our main analysis is

for entrants' employment and wages. However, the effects of class size (and students per teacher hour) might only become apparent at later stages of the career. To investigate this, we provide a tentative analysis of later labour market outcomes for the oldest cohort only - 8th graders of 1985 - because the oldest cohort can be tracked for more years of their careers. As we are interested in *later* stages of the labour market career, we include only individuals of the 1985 cohort who have entered the labour market at least 12 years prior to the last year of data (2002). This reduces the sample to mainly individuals with lower educations: about 60% of the sample have no education beyond lower secondary school, and the remaining 40% has an upper secondary education. Only 0.3% of the sample has more than upper secondary education. All in all, of the 1985-cohort of 61,400 students, only 17,000 have been in the labour market for at least 12 years. Of these, about 16,000 individuals have information on employment rates and for about 11,000 individuals there is wage information in each year. However, these samples are further reduced, as we can only use observations around the discontinuity points for the regressions: in the ± 5 discontinuity samples about 4,000-6,000 are available, and only 2,500-4,000 in the ± 3 discontinuity samples.

Figure 2 presents results from regressions for the oldest cohort. In the upper panel, for each of the three discontinuity samples, regressions of class size (and controls) on wages earned in each of the first 12 years in the labour market are estimated separately (all in all, 36 separate regressions). Figure 2 suggests that class size might have stronger and more significant impacts on wages at later stages of the career. Many estimates are numerically larger for years 4 to 12 than for the first three years. Moreover, three of the estimates are significant in the ± 5 discontinuity sample, and the estimates of year 7 are significant for all three discontinuity samples. Yet, most estimates are not significant. The corresponding results for students per teacher hour (lower panel of Figure 2) show a similar tendency, but are more mixed and imprecisely estimated throughout. Results for the employment outcome (not reported) do not display the pattern observed for wages.

[Figure 2 about here.]

In an attempt to increase the precision of the results, we have split the labour market career in two: the first three years, and the remaining period (years 4 to 12). We then use average wages for each of the two periods as dependent variables. Results from these estimations are shown in Tables 9 and 10 and confirm the pattern observed above that class-size effects might be more important at later stages of the career. However, it is important to recall the fact that this is a sample almost exclusively based on individuals with at most secondary education. The estimated effects of the second column of Table

9 suggest that a decrease in the class size by 1 (i.e. by about 5% since average class size is about 20) increases mean wages in later years (year 4-12) by about 0.6%. The point estimates for students per teacher hour suggest a larger effect (a fall of 5%, i.e. of $0.05 \times 0.55 = 0.03$, increases wages by 0.9-1.5%), but they are not significant at the 5% level.

[Tables 9 and 10 about here.]

To examine whether the effects are heterogenous by educational level, we have estimated the model above for individuals with lower and upper secondary education separately, see the last four columns of Tables 9 and 10. The point estimates for years 4 to 12 for these two subgroups are not significantly different, but the estimates are significantly different from zero only for individuals with upper secondary schooling (and only in ± 5 and ± 4 discontinuity samples for class size, and in the ± 5 discontinuity sample for students per teacher hour). We have also conducted a similar analysis for males and females, but there does not seem to be heterogeneity by gender in the school resource effects.

Using a regression discontinuity design, large differences in estimated effects depending on whether background controls are included or not may indicate that the identifying assumptions are not valid so that the distribution of pupils between schools with enrolment counts just to the right and just to the left of discontinuity points, respectively, is not random. Therefore, we have investigated, whether these significant results hold in estimations without controls for student background. Such results for the wage estimations are shown in Table A2. Both the point estimates and the standard errors are slightly increased compared to the results from estimations that control for student background (Tables 9 and 10). Nevertheless, the results are not very sensitive to whether or not we include background controls.

To sum up, this - rather tentative - analysis for later years of the labour market career suggests that while there are no significant class size effects on wages for entrants in the labour market, there might be some impact of class-size on wages later in the career. Additional results suggest that the significant class-size effect might be mainly driven by individuals with *upper* secondary schooling. For lower education levels, the point estimates are imprecisely estimated and often smaller. The estimated effect of an increase in school resources by 5% on wages during later labour market years (4-12) estimated in the discontinuity sample with the most precise estimates (the ± 5 discontinuity sample) is an increase of 0.6% for class size and 1.5% for students per teacher hour. These results suggest that teacher resources per student have a much larger impact than mere variation in class size on labour market outcomes. This pattern is also found in the study by

Browning and Heinesen (2005), who report that the number of students per teacher hour has a substantially larger effect than class size on years of education and the probability of completing upper secondary education.

5 Conclusion

We have estimated the effect of class size and students per teacher hour at 8th grade of compulsory school on the employment rate and hourly wage level during individuals' first three years in the labour market after ended education. We have taken the potential endogeneity of the school resources into account by using a regression-discontinuity design which resembles the method used by Angrist and Lavy (1999). In addition to Least Squares models for wages, we have estimated IV-Tobit models to allow for the censored character of the employment outcome.

- We find no significant effects of school resources on labour market outcomes (employment or wage rate) in the first three years in the labour market. The insignificance of the estimated effects applies across a variety of different model specifications. This 'negative' finding is not caused by very large standard errors on the parameters of school inputs. For instance, looking at the main estimation results controlling for background characteristics for the ± 5 discontinuity sample, the standard error of the class size coefficient of the wage regression is 0.001. This implies that we can in our dataset detect a wage effect of reducing class size by 1 student (i.e. by 5%) as small as 0.2%. In the corresponding employment regression the standard error of the class-size coefficient is 0.077 implying that we can detect an employment effect of a 1 pupil (5%) reduction in class size as small as 0.15 percentage points.
- A tentative analysis of class size effects at later stages of the career suggests that there might be effects for more experienced workers. Also, the results suggest that these effects are more significant for individuals with upper secondary education compared to those with less than upper secondary education. However, these results are highly tentative, because (due to data limitations) they are based on workers with an education at the secondary level only (not tertiary) and because they are restricted to the oldest cohort¹⁵. In some years, when more individuals of our dataset will have reached later stages of their labour market career, further analysis of this will be possible.

¹⁵We did not find this trend for the two subsequent cohorts of 1986 and 1987, but this may be due to even smaller sample sizes for these cohorts.

Our results may be compared to the results reported in Dustman et al. (2003) who examine the effects of pupil-teacher ratios on wages using the British National Child Development Survey (NCDS). They find that school resources, measured by the pupil/teacher ratio has no impact on wages at age 23, but has significant effects on wages at age 33. Their estimates suggest that a reduction in class size (pupils per teacher) of 1 increases wages (at 33) by about 0.3%. This effect is only about half the size of the effect indicated by the analysis of the present paper.

References

- [Angrist, J.D.; and G.W. Imbens (1995)] Two-stage least squares estimation of average causal effects in models with variable treatment intensity. *Journal of the American Statistical Association* 90: 431-442.
- [Angrist, J.D. and V. Lavy (1999)] Using Maimonides' Rule to Estimate the Effect of Class Size on Scholastic Achievement. *Quarterly Journal of Economics* 114; no.2: 533-75.
- [Betts, J., and J.L. Shkolnik (1999)] The Behavioral Effect of Variations in Class Size: The Case of Math Teachers. *Educational Evaluation and Policy Analysis*, 21, 2, 193-213, Sum 1999.
- [Blatchford, P.; and C. Martin (1998)] The Effect of Class Size on Classroom Processes: 'It's a bit like a Treadmill – Working Hard and Getting Nowhere Fast'. *British Journal of Educational Studies*, 46, 2, June, 118-137.
- [Boozer, A. B.; and T. Maloney (2001)] The Effects of Class Size on the Long Run Growth in Reading Abilities and Early Adult Outcomes in the Christchurch Health and Development Study. Economic Growth Center, Yale University, Center Discussion Paper No. 827.
- [Borghans, L. (2002)] Class Teaching and Individual Instruction. Manuscript, Research Center for Education and the Labour Market, Maastricht University, The Netherlands.
- [Bound, J., D. A. Jaeger & R. M. Baker (1995)] Problems with instrumental variables estimation when the correlation between the instruments and the endogenous explanatory variables is weak. *Journal of the American Statistical Association*, 90:443-450.
- [Browning, M. and E. Heinesen (2005)] Class Size, Teacher Hours and Educational Attainment. Working paper, akf, institute of local government studies, Copenhagen.
- [Dearden, L., J. Ferri and C. Meghir (2002)] The effect of school quality on educational attainment and wages. *Review of Economics and Statistics* 84(1): 1-20.
- [Dustmann, C.; Rajah, N.; and A. van Soest (2003)] Class Size, Education and Wages. *Economic Journal* 113: F99-F120.
- [Hahn, J., P. Todd and W. van der Klaauw (2001)] Identification and estimation of treatment effects with a regression-discontinuity design. *Econometrica* 69: 201-209.

- [Hanushek, E.A (1996)] School resources and student performance. In G. Burtless, ed.: Does Money Matter? The Effect of School Resources on Student Achievement and Adult Success. Washington D.C.: Brookings Institution.
- [Hanushek, E.A. (1997)] Assessing the Effects of School Resources on Student Performance: An update. *Educational Evaluation and Policy Analysis*; 19, 2: 141-64.
- [Heckman, J.; Lalonde, R.; and J. Smith (1999)] The Economics and Econometrics of Active Labor Market Programs. *Handbook of Labor Economics*, Volume 3, Ashenfelter, A. and D. Card, eds., Amsterdam: Elsevier Science.
- [Hoxby, C.M. (2000)] The Effects of Class Size on Student Achievement: New Evidence From Population Variation; *The Quarterly Journal of Economics*, 115, 1239-1285
- [Krueger, A. (1999)] Experimental Estimates of Education Production Functions. *Quarterly Journal of Economics*, 64, 497-532.
- [Krueger, A. (2003)] Economic considerations and class size. *Economic Journal* 113: F34-F63.
- [Lazear, E.P. (2001)] Educational production. *Quarterly Journal of Economics* 116(3): 777-805.
- [Moulton, B.R (1990)] An Illustration of a Pitfall in Estimating the Effects of Aggregate Variables in Micro Units. *Review of Economics and Statistics*, 32: 334-338.
- [Newey, W. (1987)] Efficient Estimation of Limited Dependent Variable Models with Endogenous Explanatory Variables. *Journal of Econometrics*, vol. 36, 231-50.
- [OECD (2002)] *Education at a Glance*; OECD, Paris.
- [OECD (2003)] *Education at a Glance*. OECD, Paris.
- [OECD (2005)] *Education at a Glance*; OECD, Paris.
- [Rice, J. K. (1999)] The impact of class size on instructional strategies and the use of time in high school mathematics and science courses. *Educational Evaluation and Policy Analysis*, 21(2), 215-130.
- [Staiger, D. and J. H. Stock (1997)] Instrumental variables regression with weak instruments. *Econometrica*, 65(3): 557-586.
- [Todd, P.E. and K.I. Wolpin (2003)] On the specification and estimation of the production function for cognitive achievement. *Economic Journal* 113: F3-F33.

[West, M.R.; and L. Wössmann (2005)] Which School Systems Sort Weaker Students into Smaller Classes? *International Evidence*, *European Journal of Political Economy*, in press.

[Zahorik, J.; Molnar, A.; Ehrle, K.; and A. Halbach (2000)] Smaller Classes, Better Teaching? *Effective Teaching in Reduced-Size Classes*. In Sabrina W.M. Laine and James G. Ward, eds: *Using What We Know. A Review of the Research on Implementing Class-Size Reduction Initiatives for State and Local Policymakers*.

Table 1: Descriptive statistics

	1985-1992					1985-1987		
	Main estimation sample				Full sample	Estimation sample	Full sample	
	Mean	Std. Dev	Min	Max	Mean	Mean	Mean	
Employment rate	90.119	16.520	0	100	90.119	87.318	87.318	
<i>No. of obs.</i>	362,847				362,847	166,179	166,179	
Hourly wages (year 2000-prices)	131.749	49.902	3.072	3288	131.749	136.082	136.082	
<i>No. of obs.</i>	294,823				294,823	135,322	135,322	
Enrolment	60.162	21.866	6	168	59.934	62.544	62.575	
Class size	19.729	2.722	6	31.5	19.789	19.476	19.487	
Teacher hours per student	0.552	0.061	0.194	0.871	0.554	0.544	0.544	
Own education: Lower secondary education	0.222	0.416	0	1	0.187	0.196	0.185	
Own education: Upper secondary education	0.549	0.498	0	1	0.519	0.500	0.479	
Own education: Short college education	0.061	0.240	0	1	0.062	0.072	0.070	
Own education: Long college education	0.131	0.337	0	1	0.169	0.166	0.174	
Own education: University education	0.037	0.190	0	1	0.064	0.066	0.092	
8th grade cohort of 1985	0.159	0.365	0	1	0.136	0.346	0.337	
8th grade cohort of 1986	0.157	0.363	0	1	0.137	0.341	0.341	
8th grade cohort of 1987	0.144	0.351	0	1	0.130	0.313	0.323	
8th grade cohort of 1988	0.134	0.341	0	1	0.127	0.000	0.000	
8th grade cohort of 1989	0.128	0.334	0	1	0.129	0.000	0.000	
8th grade cohort of 1990	0.108	0.310	0	1	0.118	0.000	0.000	
8th grade cohort of 1991	0.092	0.289	0	1	0.112	0.000	0.000	
8th grade cohort of 1992	0.079	0.270	0	1	0.111	0.000	0.000	
Female	0.485	0.500	0	1	0.495	0.495	0.495	
Danish native	0.938	0.241	0	1	0.941	0.943	0.944	
Immigrant	0.015	0.120	0	1	0.015	0.010	0.010	
Second generation immigrant	0.008	0.089	0	1	0.009	0.004	0.004	
No. Siblings	0.902	0.857	0	11	0.904	0.908	0.910	
No. of younger siblings	0.488	0.500	0	1	0.493	0.498	0.500	
Lives with both parents	0.716	0.451	0	1	0.725	0.738	0.740	
Lives with single mother	0.131	0.337	0	1	0.128	0.120	0.120	
Lives with single mother&stepdad	0.093	0.290	0	1	0.089	0.085	0.084	
Lives with single father	0.026	0.161	0	1	0.026	0.024	0.024	
Lives with single father&stepmum	0.018	0.134	0	1	0.018	0.017	0.017	
Lives without natural parent	0.016	0.125	0	1	0.015	0.016	0.016	
Lives without natural parent, but natural parents live together	0.001	0.038	0	1	0.001	0.002	0.002	

Data on mother missing	0.015	0.123	0	1	0.015	0.015	0.015
Data on father missing	0.051	0.220	0	1	0.049	0.054	0.054
Teen-mum	0.095	0.294	0	1	0.087	0.091	0.088
Teen-dad	0.022	0.146	0	1	0.020	0.021	0.020
Both parents only lower secondary education	0.235	0.424	0	1	0.212	0.249	0.239
At least one parent has upper secondary educ.	0.493	0.500	0	1	0.475	0.479	0.469
At least one parent has short college education	0.048	0.215	0	1	0.051	0.047	0.047
One parent has long college or university educ.	0.140	0.347	0	1	0.160	0.142	0.152
Both parents have long college or university educ.	0.058	0.234	0	1	0.080	0.062	0.073
Parental education missing	0.022	0.146	0	1	0.021	0.017	0.017
Mother wage earner	0.777	0.416	0	1	0.787	0.764	0.766
Father wage earner	0.737	0.440	0	1	0.744	0.734	0.735
Mother student	0.090	0.286	0	1	0.088	0.101	0.100
Father student	0.160	0.367	0	1	0.159	0.169	0.168
Mother self-employed	0.005	0.070	0	1	0.005	0.005	0.005
Father self-employed	0.001	0.035	0	1	0.001	0.001	0.001
Mother social benefits	0.015	0.123	0	1	0.014	0.012	0.011
Father social benefits	0.009	0.093	0	1	0.008	0.005	0.005
Mother not active in labour market	0.097	0.297	0	1	0.091	0.105	0.102
Father not active in labour market	0.041	0.199	0	1	0.039	0.037	0.036
Mother's income	13.446	10.121	0	196	14.044	13.000	13.223
Father's income	21.626	18.195	0	2466	22.414	22.441	22.785
Mother's unemployment rate	7.770	20.911	0	100	7.382	6.886	6.716
Father's unemployment rate	4.919	16.323	0	100	4.707	3.923	3.824
Mother's wealth	0.340	5.153	-939.2	1311	0.397	0.513	0.551
Father's wealth	1.982	35.828	-371.5	20850	2.057	2.829	2.918
Lives in owned dwelling	0.725	0.447	0	1	0.741	0.736	0.741
Lives in rented dwelling	0.235	0.424	0	1	0.220	0.223	0.217
Other type of dwelling	0.016	0.127	0	1	0.017	0.016	0.017
Type of dwelling missing	0.008	0.090	0	1	0.008	0.009	0.009
Rooms per person	1.207	0.458	0	7.5	1.222	1.208	1.214
Rooms per person unknown	0.004	0.067	0	1	0.004	0.005	0.005
Peers: Mean parental years of schooling	11.508	0.706	9	14.875	11.557	11.432	11.453
Peers: Mean parental unemployment rate	6.013	3.233	0	29.876	6.044	5.308	5.270
Peers: Mean parental income	18.066	3.780	0	52.291	18.229	17.929	18.004
Unemployment rate in municipality (year of 8th grade)	9.246	3.005	1.9688	23.287	9.440	7.939	7.929
<i>No. of obs.</i>		372,182			453,076	170,772	182,446

Table 2: Cohorts of 8th graders in the estimation sample by completed education

	Unskilled	High-school /Voc. training	Short college	Long college	University	PhD	All	No. Obs	% of cohort included in estimation sample
1985	28%	52%	4%	9%	7%	0.2%	100%	59056	96%
1986	28%	53%	4%	9%	6%	0.2%	100%	58257	94%
1987	28%	54%	4%	9%	5%	0.1%	100%	53459	91%
1988	28%	57%	3%	8%	4%	0.0%	100%	50024	87%
1989	30%	58%	3%	7%	2%	0.0%	100%	47536	82%
1990	30%	61%	3%	5%	0.7%	0%	100%	40062	75%
1991	31%	63%	3%	2%	0.2%	0%	100%	34322	67%
1992	34%	64%	2%	0.6%	0.0%	0%	100%	29466	59%

Table 3: Summary statistics for all public schools 1985-92 (8,764 observations)

<i>Variable</i>	<i>Mean</i>	<i>SD</i>	<i>Min</i>	<i>Max</i>
Enrolment	51.99	20.4	6	168
Number of classes	2.68	0.98	1	8
Number of teacher hours/week	95.04	34.49	16	275
Class size	19.45	2.98	6	31.5
Students/teacher hour/week	0.54	0.07	0.19	0.87

Table 4: Estimation results for effects of school resources on wage

Sample	Full sample		Discontinuity points 1-4					
			+/-5 discontinuity sample		+/-4 discontinuity sample		+/-3 discontinuity sample	
Background	Without controls	With controls	Without controls	With controls	Without controls	With controls	Without controls	With controls
OLS								
Observations	294,823/1128		115,956/1071		91,893/1031		68,623/950	
Coef. (class size)	0.0006	0.0002	0.0003	0.0001	0.0005	0.0002	0.0002	0.0000
se	0.0003	0.0002	0.0003	0.0003	0.0003	0.0003	0.0004	0.0003
p-value	0.0300	0.3720	0.3260	0.6900	0.1370	0.4590	0.6660	0.9220
R-squared	0.0094	0.2531	0.0090	0.2509	0.0089	0.2498	0.0091	0.2494
Coef. (stud/teacher hour)	-0.0288	-0.0118	0.0011	0.0065	0.0069	0.0068	-0.0094	-0.0032
se	0.0135	0.0111	0.0172	0.0149	0.0192	0.0163	0.0220	0.0182
p-value	0.0330	0.2880	0.9490	0.6640	0.7180	0.6760	0.6710	0.8600
R-squared	0.0094	0.2531	0.0090	0.2509	0.0089	0.2498	0.0091	0.2494
2SLS								
Observations			115,956/1071		91,893/1031		68,623/950	
Coef. (class size)			<i>-0.0012</i>	0.0000	<i>-0.0019</i>	<i>-0.0005</i>	<i>-0.0026</i>	<i>-0.0019</i>
se			0.0011	0.0010	0.0015	0.0013	0.0021	0.0018
p-value			0.2650	0.9990	0.1950	0.6790	0.2110	0.2960
R-squared			0.0088	0.2509	0.0082	0.2497	0.0082	0.2490
F 1st stage			60.6	60.5	32.3	32.3	14.9	14
Coef. (stud/teacher hour)			<i>-0.1249</i>	-0.0126	<i>-0.1797</i>	<i>-0.0453</i>	<i>-0.1989</i>	<i>-0.1197</i>
se			0.1141	0.1019	0.1472	0.1271	0.2081	0.1833
p-value			0.2740	0.9020	0.2220	0.7220	0.3390	0.5140
R-squared			0.0085	0.2509	0.0076	0.2497	0.0077	0.2489
F 1st stage			17.8	16.8	10.57	10.03	5.17	5.17

Note: Robust standard errors corrected for within-school correlation are reported. All regressions include a constant term, dummies for the year of 8th grade school attendance, and dummy variables for enrolment segment. 'With controls' means inclusion of the full set of control variables described in section 3 and as shown in Table A1. 'Observations' refer to the number of students/the number of schools.

The 'F 1st stage' refers to the F-statistics for the instrument in the first-stage regression; it uses standard errors clustered by school.

Significance at the 5%/10% levels is indicated by bold/italic types.

Table 5: Estimation results for effects of school resources on employment

Sample	Full sample		Discontinuity points 1-4					
			+/-5 discontinuity sample		+/-4 discontinuity sample		+/-3 discontinuity sample	
Background	Without controls	With controls	Without controls	With controls	Without controls	With controls	Without controls	With controls
Tobit								
Observations	362,847		142,700/1,071		112994/1031		84389/951	
Coef. (class size)	0.1325	0.0353	0.1486	0.0558	0.1722	0.0698	0.1813	0.0790
se	0.0176	0.0167	0.0211	0.0202	0.0229	0.0220	0.0259	0.0249
p-value	0.0000	0.0350	0.0000	0.0060	0.0000	0.0020	0.0000	0.0020
Pseudo R-sq.	0.0072	0.0232	0.0074	0.0232	0.0075	0.0233	0.0076	0.0238
Coef. (stud/teacher hour)	5.7891	3.2105	6.0027	3.2909	7.7972	4.5425	8.6637	5.5420
se	0.8218	0.7845	1.1502	1.1036	1.2685	1.2173	1.4428	1.3856
p-value	0.0000	0.0000	0.0000	0.0030	0.0000	0.0000	0.0000	0.0000
Pseudo R-sq.	0.0072	0.0232	0.0073	0.0232	0.0075	0.0233	0.0076	0.0238
IV-Tobit								
Observations			142,700/1,071		112994/1031		84389/951	
Coef. (class size)			-0.0957	0.0295	-0.0710	0.0953	-0.1372	0.0280
se			0.0985	0.0771	0.1243	0.1000	0.1798	0.1504
p-value			0.3310	0.7020	0.5680	0.3410	0.4460	0.8520
Log pseudolikelihood			-716611	-710121	-571943	-566621	-429474	-425291
F 1st stage			60.6	60.5	32.3	32.3	14.9	14
Coef. (stud/teacher hour)			-9.3260	1.7410	-8.4646	7.3668	-21.3157	-4.1560
se			10.2030	8.1700	12.2513	10.0214	19.0860	15.9844
p-value			0.3610	0.8310	0.4900	0.4620	0.2640	0.7950
Log pseudolikelihood			-148733	-141936	-119951.00000	-114501	-91004.00000	-86704
F 1st stage			17.8	16.8	10.6	10.03	5.2	4.5

Note: Robust standard errors corrected for within-school correlation are reported. All regressions include a constant term, dummies for the year of 8th grade school attendance, and dummy variables for enrolment segment. 'With controls' means inclusion of the full set of control variables described in section 3 and as shown in Table A1. 'Observations' refer to the number of students/the number of schools.

The 'F 1st stage' refers to the F-statistic for the instruments in the first-stage regression; it uses standard errors clustered by school.

Significance at the 5%/10% levels is indicated by bold/italic types.

Table 6: Results for the oldest three cohorts: 1985-87

Sample	Discontinuity points 1-4					
	+/-5 discontinuity sample		+/-4 discontinuity sample		+/-3 discontinuity sample	
Background	Without controls	With controls	Without controls	With controls	Without controls	With controls
Employment: IV-Tobit						
Observations	63546/829		49667/715		37,083/578	
Coef. (class size)	-0.1408	0.0373	-0.1828	0.0895	-0.2202	0.1708
se	0.1389	0.1143	0.1740	0.1428	0.2466	0.2060
p-value	0.3110	0.7440	0.2930	0.5310	0.3720	0.4070
Coef. (stud/teacher hour)	-23.1165	2.3130	-25.3090	12.9090	-39.8450	12.6811
se	19.9210	19.7150	20.7810	19.3720	29.9120	27.8711
p-value	0.2460	0.9070	0.2230	0.5050	0.1830	0.6490
Wage: 2SLS						
Observations	52118/829		40733/715		30,450/577	
Coef. (class size)	0.0003	0.0001	-0.0014	-0.0005	-0.0009	-0.0002
se	0.0018	0.0014	0.0023	0.0018	0.0032	0.0026
p-value			0.5370	0.7960	0.7800	0.9440
Coef. (stud/teacher hour)	0.0164	0.0830	-0.1879	0.1155	-0.2918	0.0431
se	0.2350	0.2112	0.2610	0.2252	0.3425	0.3256
p-value	0.9440	0.6940	0.4720	0.6080	0.3940	0.8950

Note: Robust standard errors corrected for within-school correlation are reported. All regressions include a constant term, dummies for the year of 8th grade school attendance, and dummy variables for enrolment segment. 'With controls' is the full set of control variables described in section 3 and as shown in Table A1. 'Observations' refer to the number of students/the number of schools.

The 'F 1st stage' refers to the F-statistic for the instrument in the first-stage regression; it uses standard errors clustered by school. Significance at the 5%/10% levels is indicated by bold/italic types.

Table 7: Results for effect on wages without low-quality wage observations

Sample	Discontinuity points 1-4					
	+/-5 discontinuity sample		+/-4 discontinuity sample		+/-3 discontinuity sample	
Background	Without controls	With controls	Without controls	With controls	Without controls	With controls
Observations	94,692/1,071		75,016/1031		56,021/950	
Coef. (class size)	0.0004	0.0015	0.0001	0.0012	-0.0015	-0.0009
se	0.0011	0.0009	0.0014	0.0012	0.0020	0.0016
p-value	0.6980	0.9400	0.9290	0.3150	0.4440	0.5950
R-squared	0.0059	0.3538	0.0058	0.3527	0.0053	0.3549
F 1st stage	60.6	60.5	32.3	32.3	14.9	14.0
Coef. (stud/teacher hour)	0.0583	0.1533	0.0393	0.1374	-0.0254	0.0207
se	0.1119	0.0948	0.1423	0.1192	0.2015	0.1696
p-value	0.6020	0.1060	0.7830	0.2490	0.9000	0.9030
R-squared	0.0056	0.3530	0.0057	0.3519	0.0058	0.3549
F 1st stage	17.8	16.8	10.6	10.0	5.2	4.5

Note: Robust standard errors corrected for within-school correlation are reported. All regressions include a constant term, dummies for the year of 8th grade school attendance, and dummy variables for enrolment segment. 'With controls' is the full set of control variables described in section 3 and as shown in Table A1. 'Observations' refer to the number of students/the number of schools.

The 'F 1st stage' refers to the F-statistic for the instrument in the first-stage regression; it uses standard errors clustered by school. Significance at the 5%/10% levels is indicated by bold/italic types.

Table 8: Main results using only observations at the two first discontinuity segments (around 24 and 48)

Sample	Discontinuity points 1-2					
	+/-5 discontinuity sample		+/-4 discontinuity sample		+/-3 discontinuity sample	
Background	Without controls	With controls	Without controls	With controls	Without controls	With controls
Employment: IV-Tobit						
Observations	80,031		62,479		46,256	
Coef. (class size)	-0.0966	0.0232	-0.0697	0.0877	-0.1269	0.0386
se	0.103	0.0806	0.13	0.1054	0.18677	0.16
p-value	0.348	0.774	0.60	0.405	0.497	0.81
Log pseudolikelihood	-413106	-409540	-325835	-322922	-243153	-240834
F 1st stage	109	108.6	57	56.1	25.9	24.4
Coef. (stud/teacher hour)	-10.191	0.792	-7.6172	6.392	-15.522	-1.6857
se	10.27	8.49	12.49	10.421	18.413	16.497
p-value	0.321	0.926	0.54	0.54	0.399	0.919
Log pseudolikelihood	-90848	-87141	-73553	-70471	-55871	-53410
F 1st stage	34.1	32.07	20.2	19.1	9.7	8.5
Wage: 2SLS						
Observations	64,536/865		50,393/813		37,327/732	
Coef. (class size)	-0.00145	0.00020	-0.00225	-0.00050	-0.00240	-0.0013
se	0.00127	0.0011	0.00166	0.0014	0.0024	0.002
p-value	0.254	0.852	0.175	0.716	0.309	0.531
R-squared	0.0005	0.276	.	0.2724	.	0.2725
F 1st stage	109	108.6	57	56.1	25.9	24.4
Coef. (stud/teacher hour)	-0.1391	0.0165	-0.2127	-0.0599	-0.2093	-0.1039
se	0.127	0.1092	0.1637	0.1368	0.234	0.2006
p-value	0.274	0.88	0.194	0.662	0.371	0.605
R-squared	0.0001	0.276	.	0.2722	.	0.2722
F 1st stage	34.1	32.07	20.2	19.1	9.7	8.5

Note: Robust standard errors corrected for within-school correlation are reported. All regressions include a constant term, dummies for the year of 8th grade school attendance, and dummy variables for enrolment segment. 'With controls' is the full set of control variables described in section 3 and as shown in Table A1. 'Observations' refer to the number of students/the number of schools.

The 'F 1st stage' refers to the F-statistic for the instrument in the first-stage regression; it uses standard errors clustered by school. Significance at the 5%/10% levels is indicated by bold/italic types.

Table 9: Effect of class size on wage for later years (1-12)

	All		Lower secondary educ.		Upper secondary educ.	
	Av. year1-3	Av. Year4-12	Av. year1-3	Av. Year4-12	Av. year1-3	Av. Year4-12
+/-5 discontinuity sample						
Observations	5,544/429	5,833/429	3,006/426	3,273/425	2,538/423	2,560/425
Coef. (class size)	-0.0010	-0.0056	-0.0029	-0.0039	0.0010	-0.0075
se	0.0038	0.0026	0.0064	0.0036	0.0047	0.0030
p-value	0.7920	0.0320	0.6460	0.2780	0.828	0.0130
R-squared	0.1147	0.2411	0.0616	0.1879	0.1563	0.2860
+/-4 discontinuity sample						
Observations	4,513/342	4,738/342	2,443/340	2,653/339	2,070/338	2,085/340
Coef. (class size)	-0.0046	-0.0062	-0.0111	-0.0063	0.0010	-0.0068
se	0.0052	0.0034	0.0099	0.0053	0.0056	0.0034
p-value	0.3750	0.0690	0.2620	0.2380	0.8600	0.0470
R-squared	0.1170	0.2417	0.0515	0.1743	0.1530	0.2976
+/-3 discontinuity sample						
Observations	3,405/262	3,571/262	1,849/262	2,001/260	1,556/258	1,570/260
Coef. (class size)	-0.0062	-0.0055	-0.0137	-0.0067	0.0032	-0.0059
se	0.0064	0.0040	0.0112	0.0059	0.0064	0.0040
p-value	0.3340	0.1740	0.2230	0.2610	0.6240	0.1360
R-squared	0.1064	0.2438	0.0372	0.1829	0.1493	0.2922

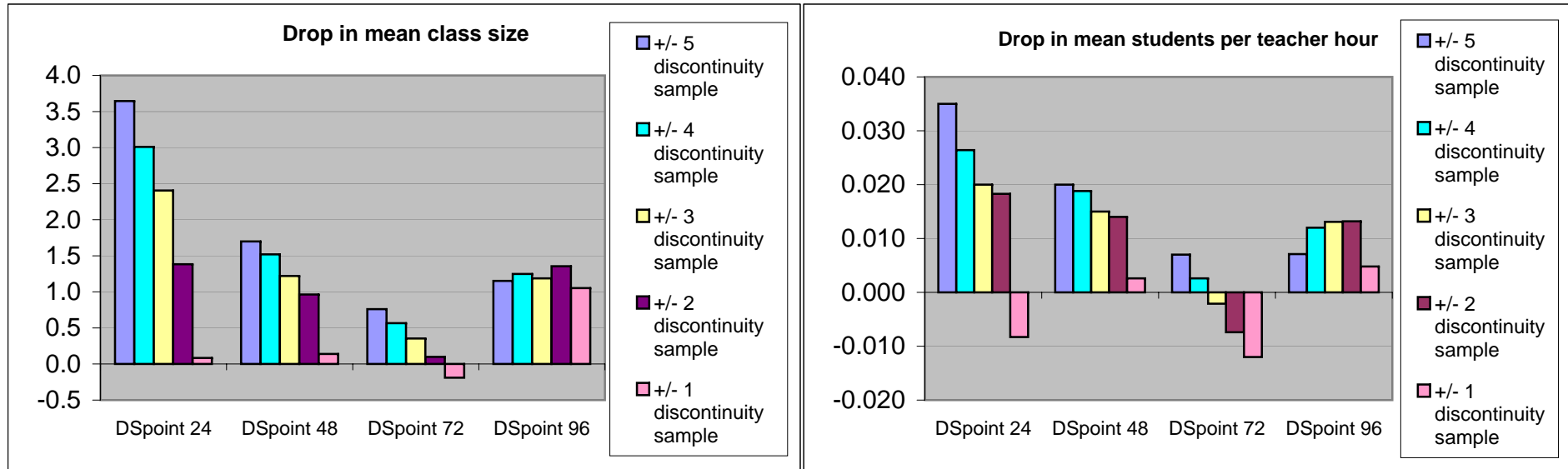
Table 10: Effect of students per teacher hours on wage for later years (1-12)

	All		Lower secondary educ.		Upper secondary educ.	
	Av. year1-3	Av. Year4-12	Av. year1-3	Av. Year4-12	Av. year1-3	Av. Year4-12
+/-5 discontinuity						
Observations	5,544/429	5,833/429	3,006/426	3,273/425	2,538/423	2,560/425
Coef. (stud/teacher hour)	-0.3785	-0.5281	-1.6128	-0.3499	0.2420	-0.7907
se	0.5325	0.3613	1.2115	0.4532	0.5861	0.4457
p-value	0.4780	0.1450	0.1840	0.4400	0.6800	0.077
R-squared	0.1075	0.2220
+/-4 discontinuity						
Observations	4,513/342	4,738/342	2,443/340	2,653/339	2,070/338	2,085/340
Coef. (stud/teacher hour)	-0.6357	-0.3443	-2.5368	-0.2404	0.2828	-0.4132
se	0.6816	0.3655	2.0751	0.5609	0.6388	0.4021
p-value	0.3520	0.3470	0.2220	0.669	0.658	0.3050
R-squared	0.1010	0.2453	.	0.1851	0.1498	0.2986
+/-3 discontinuity						
Observations	3,405/262	3,571/262	1,849/262	2,001/260	1,556/258	1,570/260
Coef. (stud/teacher hour)	-0.8378	-0.2952	-1.6889	-0.3166	0.2969	-0.1738
se	0.6326	0.3130	1.0917	0.4067	0.6214	0.3698
p-value	0.1870	0.3470	0.1230	0.4370	0.6330	0.6390
R-squared	0.0796	0.2471	0.2953	0.1899	0.1492	0.3043

Note: Robust standard errors corrected for within-school correlation are reported. All regressions include a constant term, dummie the year of 8th grade school attendance, and dummy variables for enrolment segment. 'With controls' is the full set of control varia described in section 3 and as shown in Table A1. 'Observations' refer to the number of students/the number of schools.

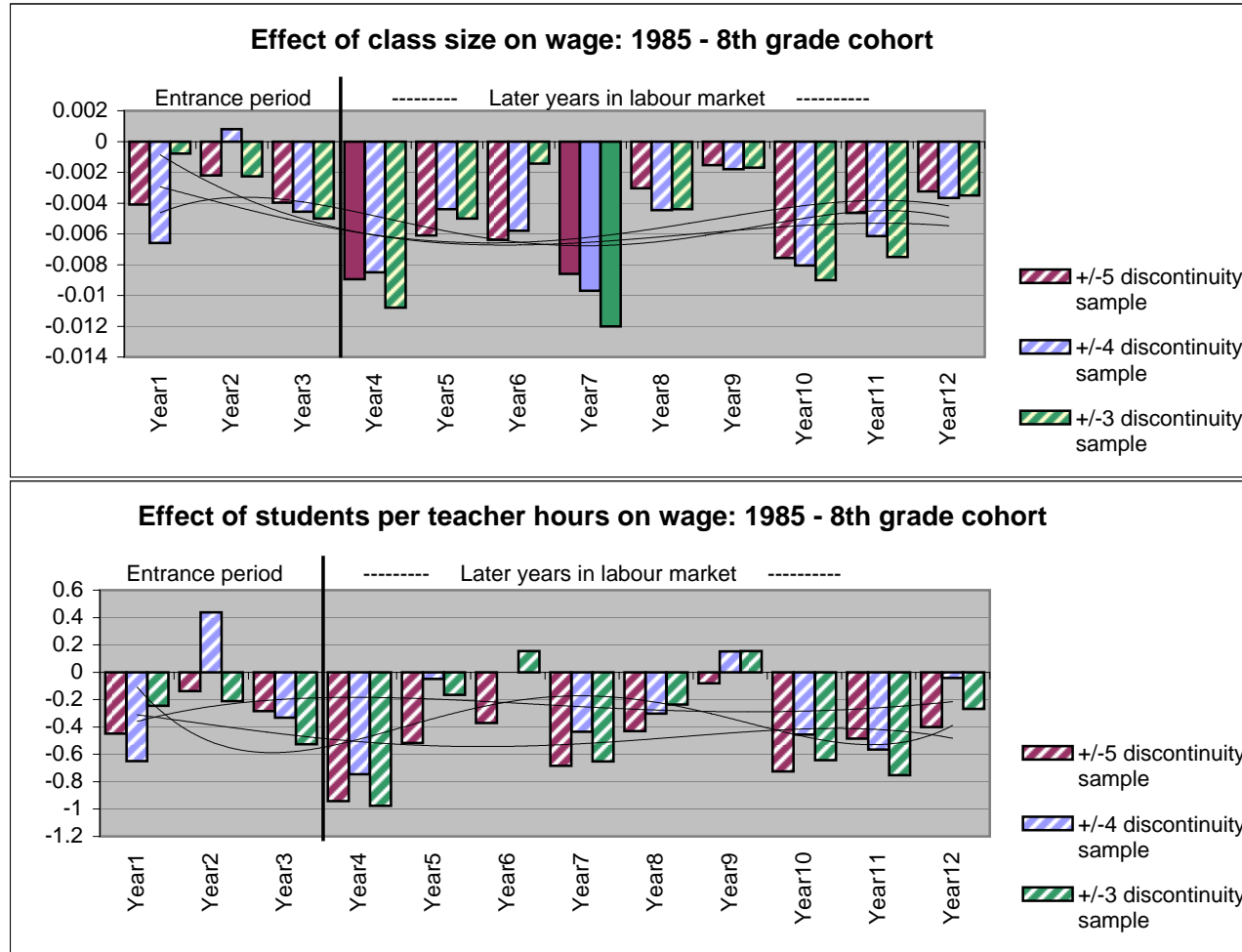
The 'F 1st stage' refers to the F-statistic for the instrument in the first-stage regression; it uses standard errors clustered by school. Significance at the 5%/10% levels is indicated by bold/italic types.

Figure 1: Drop in mean school resources around discontinuity points



Note: Based on individual-level data. The patterns look similar when based on school-level data.

Figure 2: Later career school resource effects on wage for the 1985-cohort



Note: Each column displays the school resource estimate from a separate regression (= results from 36 regressions). Insignificant results (at the 5% level) are indicated by stripes. All regressions include the full set of control variables described in section 3 and as shown in Table A1. Source: Tb. A1.

Table A1: Estimation results for basic wage and employment regressions (± 5 discontinuity sample) and descriptive statistics for these samples

	Wage regression (2SLS)							Employment regression (IV-tobit)						
	Descriptive statistics				Estimation results			Descriptive statistics				Estimation results		
	Mean	SD	Min	Max	Coef.	Robust Std. Err	P> t	Mean	SD	Min	Max	Coef.	Robust Std. Err	P> z
Class size	20.039	3.470	10	29	0.000	0.001	0.876	20.006	3.487	10	29	0.101	0.079	0.199
8th grade cohort of 1985	0.154	0.361	0	1	0.071	0.005	0.000	0.154	0.361	0	1	-20.881	0.355	0.000
8th grade cohort of 1986	0.154	0.361	0	1	0.051	0.005	0.000	0.154	0.361	0	1	-19.491	0.347	0.000
8th grade cohort of 1987	0.138	0.345	0	1	0.044	0.005	0.000	0.137	0.344	0	1	-15.749	0.352	0.000
8th grade cohort of 1988	0.139	0.346	0	1	0.028	0.004	0.000	0.137	0.344	0	1	-12.083	0.332	0.000
8th grade cohort of 1989	0.125	0.331	0	1	0.015	0.004	0.000	0.125	0.330	0	1	-7.770	0.332	0.000
8th grade cohort of 1990	0.113	0.317	0	1	0.007	0.004	0.067	0.114	0.317	0	1	-3.801	0.332	0.000
8th grade cohort of 1991	0.096	0.294	0	1	0.002	0.004	0.715	0.097	0.296	0	1	-1.638	0.341	0.000
8th grade cohort of 1992	0.081	0.273	0	1	<i>Reference</i>			0.082	0.275	0	1	<i>Reference</i>		
Age	21.556	2.817	14	32	0.164	0.006	0.000	21.277	2.884	14	32	-9.050	0.444	0.000
Age-squared					-0.003	0.000	0.000					0.210	0.010	0.000
Age>18 when entering labour market	0.934	0.248	0	1	0.037	0.006	0.000	0.913	0.282	0	1	-9.060	0.368	0.000
d24 (First discontinuity segment)	0.101	0.302	0	1	<i>Reference</i>			0.106	0.307	0	1	<i>Reference</i>		
d48 (Second discontinuity segment)	0.452	0.498	0	1	0.000	0.003	0.999	0.455	0.498	0	1	-0.765	0.259	0.003
d72 (Third discontinuity segment)	0.327	0.469	0	1	-0.003	0.004	0.401	0.322	0.467	0	1	-0.397	0.288	0.169
d96 (Fourth discontinuity segment)	0.120	0.325	0	1	-0.007	0.004	0.104	0.117	0.321	0	1	-0.477	0.335	0.155
Own education: Lower secondary education	0.204	0.403	0	1	<i>Reference</i>			0.270	0.444	0	1	<i>Reference</i>		
Own education: Upper secondary education	0.635	0.481	0	1	0.066	0.003	0.000	0.590	0.492	0	1	13.175	0.207	0.000
Own education: Short college education	0.041	0.198	0	1	0.148	0.005	0.000	0.037	0.188	0	1	7.319	0.416	0.000
Own education: Long college education	0.089	0.285	0	1	0.196	0.004	0.000	0.077	0.267	0	1	10.850	0.360	0.000
Own education: University education	0.031	0.172	0	1	0.348	0.006	0.000	0.027	0.162	0	1	4.804	0.503	0.000
Female	0.465	0.499	0	1	-0.150	0.002	0.000	0.483	0.500	0	1	-4.939	0.136	0.000
Danish native	0.967	0.179	0	1	<i>Reference</i>			0.955	0.208	0	1	<i>Reference</i>		
Immigrant	0.009	0.096	0	1	0.044	0.010	0.000	0.015	0.121	0	1	-7.990	0.599	0.000
Second generation immigrant	0.006	0.079	0	1	0.000	0.011	0.969	0.008	0.090	0	1	-2.718	0.759	0.000
No. Siblings	0.897	0.831	0	11	0.002	0.002	0.164	0.903	0.859	0	11	0.098	0.115	0.391
No. of younger siblings	0.491	0.500	0	1	-0.005	0.002	0.033	0.488	0.500	0	1	0.966	0.189	0.000
Lives with both parents	0.745	0.436	0	1	<i>Reference</i>			0.717	0.450	0	1	<i>Reference</i>		
Lives with single mother	0.117	0.321	0	1	0.006	0.003	0.053	0.131	0.337	0	1	-3.199	0.247	0.000
Lives with single mother&stepdad	0.086	0.281	0	1	0.001	0.003	0.825	0.092	0.289	0	1	-3.107	0.252	0.000
Lives with single father	0.024	0.153	0	1	0.014	0.006	0.022	0.026	0.160	0	1	-3.974	0.457	0.000
Lives with single father&stepmum	0.017	0.129	0	1	0.001	0.007	0.849	0.018	0.134	0	1	-4.187	0.522	0.000
Lives without natural parent	0.010	0.101	0	1	-0.023	0.011	0.031	0.016	0.124	0	1	-9.391	0.658	0.000
Lives without natural parent, but natural parents live together	0.001	0.031	0	1	-0.049	0.028	0.079	0.002	0.040	0	1	0.477	1.682	0.777
Data on mother missing	0.013	0.112	0	1	0.003	0.010	0.796	0.015	0.121	0	1	2.129	0.679	0.002
Data on father missing	0.043	0.204	0	1	0.001	0.005	0.823	0.050	0.218	0	1	3.010	0.400	0.000

Teen-mum	0.090	0.286	0	1	0.007	0.003	0.040	0.097	0.296	0	1	-2.502	0.248	0.000
Teen-dad	0.020	0.141	0	1	0.005	0.006	0.445	0.022	0.148	0	1	-1.208	0.471	0.010
Both parents with lower secondary education	0.223	0.417	0	1	<i>Reference</i>			0.238	0.426	0	1	<i>Reference</i>		
At least one parent has upper secondary educ.	0.513	0.500	0	1	-0.013	0.002	0.000	0.497	0.500	0	1	1.757	0.173	0.000
At least one parent has short college education	0.049	0.217	0	1	-0.024	0.004	0.000	0.048	0.214	0	1	2.357	0.349	0.000
One parent has long college or university educ.	0.139	0.346	0	1	-0.026	0.003	0.000	0.138	0.345	0	1	2.227	0.249	0.000
Both parents have long college or university educ.	0.055	0.228	0	1	-0.036	0.004	0.000	0.056	0.230	0	1	1.475	0.357	0.000
Parental education missing	0.017	0.129	0	1	0.000	0.007	0.984	0.021	0.142	0	1	-0.297	0.505	0.557
Mother wage earner	0.790	0.408	0	1	<i>Reference</i>			0.777	0.417	0	1	<i>Reference</i>		
Father wage earner	0.748	0.434	0	1	<i>Reference</i>			0.737	0.440	0	1	<i>Reference</i>		
Mother self-employed	0.094	0.292	0	1	0.008	0.004	0.051	0.091	0.287	0	1	2.834	0.336	0.000
Father self-employed	0.166	0.372	0	1	0.013	0.003	0.000	0.160	0.367	0	1	5.117	0.291	0.000
Mother student	0.004	0.064	0	1	0.004	0.015	0.813	0.005	0.069	0	1	-1.381	0.976	0.157
Father student	0.001	0.030	0	1	0.012	0.028	0.657	0.001	0.033	0	1	0.818	1.985	0.680
Mother social benefits	0.010	0.101	0	1	0.003	0.009	0.729	0.015	0.122	0	1	-5.534	0.585	0.000
Father social benefits	0.006	0.078	0	1	0.006	0.012	0.621	0.009	0.093	0	1	-2.521	0.736	0.001
Mother not active in labour market	0.089	0.285	0	1	-0.002	0.004	0.612	0.098	0.297	0	1	-1.779	0.307	0.000
Father not active in labour market	0.036	0.187	0	1	0.007	0.005	0.179	0.042	0.202	0	1	-1.084	0.378	0.004
Mother's income	13.790	9.934	0	127.6	0.001	0.000	0.000	13.393	10.055	0	127.6	0.069	0.012	0.000
Father's income	22.054	17.379	0	519	0.000	0.000	0.000	21.422	17.438	0	757.24	0.091	0.007	0.000
Mother's unemployment rate	7.164	20.071	0	100	0.000	0.000	0.085	7.914	21.130	0	100	-0.037	0.004	0.000
Father's unemployment rate	4.440	15.367	0	100	0.000	0.000	0.010	4.994	16.419	0	100	-0.017	0.005	0.000
Mother's wealth	0.343	5.595	-269	1311	0.000	0.000	0.585	0.328	5.247	-268.9	1311	0.127	0.026	0.000
Father's wealth	1.979	10.475	-230	1189	0.000	0.000	0.699	1.814	10.141	-371.5	1189.4	0.072	0.009	0.000
Lives in owned dwelling	0.759	0.428	0	1	<i>Reference</i>			0.727	0.446	0	1	<i>Reference</i>		
Lives in rented dwelling	0.207	0.405	0	1	0.005	0.002	0.063	0.233	0.422	0	1	-2.428	0.193	0.000
Other type of dwelling	0.016	0.124	0	1	0.009	0.007	0.214	0.017	0.128	0	1	-0.830	0.538	0.123
Type of dwelling missing	0.008	0.090	0	1	0.006	0.012	0.598	0.008	0.091	0	1	-0.791	0.925	0.393
Rooms per person	1.219	0.446	0	6.67	-0.006	0.002	0.016	1.207	0.459	0	7.50	1.456	0.188	0.000
Rooms per person unknown	0.005	0.068	0	1	-0.013	0.017	0.449	0.005	0.069	0	1	2.814	1.235	0.023
Peers: Mean parental years of schooling	11.502	0.688	9.50	14.78	-0.017	0.002	0.000	11.492	0.697	9.00	14.78	-0.107	0.169	0.527
Peers: Mean parental unemployment rate	5.961	3.136	0	29.88	0.000	0.000	0.263	6.094	3.252	0	29.88	-0.063	0.030	0.037
Peers: Mean parental income	18.034	3.675	5.98	34.40	0.003	0.000	0.000	17.955	3.696	0.00	34.40	-0.004	0.033	0.891
Unemployment rate in municipality	9.281	2.963	1.97	23.29	-0.001	0.000	0.009	9.365	2.978	1.97	23.29	-0.550	0.032	0.000
Constant					3.050	0.071	0.000					207.946	5.095	0.000
No. Observations				115,956							142,700			
R-squared / LogPseudoLikelihood						0.2135							-710121	

Note: Significance at the 5% level are indicated by bold types.

Table A2: Without controls: Effect of class size on wage for later years (1-12)

	All		Lower secondary educ.		Upper secondary educ.	
	Av. year1-3	Av. Year4-12	Av. year1-3	Av. Year4-12	Av. year1-3	Av. Year4-12
+/-5 discontinuity sample						
Observations	5,544/429	5,833/429	3,006/426	3,273/425	2,538/423	2,560/425
Class size	-0.0009	-0.0069	0.0000	-0.0045	0.0018	-0.0077
	0.0042	0.0031	0.0069	0.0039	0.0049	0.0036
p=	0.8250	0.0260	0.9990	0.2530	0.7210	0.0320
+/-4 discontinuity sample						
Class size	-0.0041	-0.0092	-0.0065	-0.0101	0.0017	-0.0077
	0.0052	0.0041	0.0099	0.0061	0.0056	0.0040
p=	0.4330	0.0270	0.5080	0.0960	0.7580	0.0580
+/-3 discontinuity sample						
Observations	3,405/262	3,571/262	1,849/262	2,001/260	1,556/258	1,570/260
Class size	-0.0038	-0.0070	-0.0098	-0.0083	0.0031	-0.0066
p=	0.5170	0.0930	0.3550	0.1470	0.6380	0.1760

Note: No controls included except for dummies for year of 8th grade attendance and discontinuity segment dummies. Significance at the 5%/10% levels is indicated by bold/italic types.

Figure A1: Mean class size for each value of enrollment and expected class size

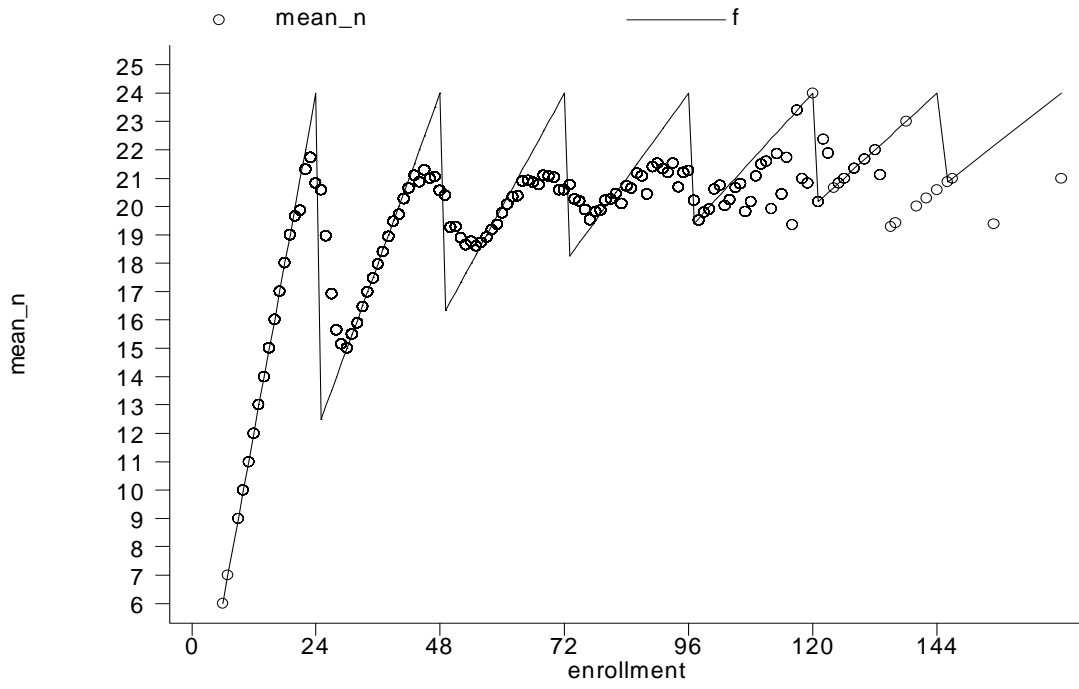


Figure A2: Average number of pupils per teacher hour for each value of enrollment and fitted values from a regression of this variable on the instrument.

