NBER WORKING PAPER SERIES

INDUSTRIAL ACTIONS IN SCHOOLS: STRIKES AND STUDENT ACHIEVEMENT

Michael Baker

Working Paper 16846
http://www.nber.org/papers/w16846

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
March 2011

I thank seminar participants at Dalhousie for helpful comments. The EQAO test scores were accessed through the Public Economics Data Analysis Laboratory at McMaster University. Abigail Payne generously shared her knowledge of these data. I am indebted to my research assistants Winnie Chan, Qing Hong, Davin Raiha, and especially Ekaterina Chmatova and Josh Lewis. The research support of SSHRC (410-2008-0346) is gratefully acknowledged. The views expressed herein are those of the author and do not necessarily reflect the views of the National Bureau of Economic Research.

Industrial Actions in Schools: Strikes and Student Achievement
Michael Baker
NBER Working Paper No. 16846
March 2011
JEL No. I21,J24,J51

ABSTRACT

While many jurisdictions ban teacher strikes on the assumption that they harm students, there is surprisingly little research on this question. The majority of existing studies make cross section comparisons of students who do or do not experience a strike, and report that strikes do not affect student performance. I present new estimates from a sample of strikes in the Canadian province of Ontario over the period 1998-2005. The empirical strategy controls for fixed student characteristics at the school cohort level. The results indicate that teacher strikes in grades 2 or 3 have on average a small, negative and statistically insignificant effect on grade 3 through grade 6 test score growth, although there is some heterogeneity across school boards. The effect of strikes in grades 5 and 6 on grade 3 through grade 6 score growth is negative, much larger and statistically significant. The largest impact is on math scores: 29 percent of the standard deviation of test scores across school/grade cohorts.

Michael Baker
Department of Economics
University of Toronto
150 St. George Street
Toronto, ON M5S 3G7
CANADA
and NBER
baker@chass.utoronto.ca

Introduction

Unions represent primary and secondary school teachers in many developed countries. How teacher unions impact student achievement has been the subject of a number of previous studies. Much of this research focuses on how unionism effects the production of education, through changes in curriculum, changes in the allocation of resources and changes in the incentives for teachers (e.g., Eberts and Stone 1987, Hoxby 1996, Lovenheim 2009).

A much more direct, and perhaps blunt, impact of unions on educational production is through industrial action. Public policy, at least in the U.S., suggests we know what this impact is and it isn't good. Only 13 states provide teachers with the right to strike (Colasanti 2008). Although strikes can have a multitude of effects, including creating major inconvenience for parents, the more often cited impact in the crusade against teachers' right to strike is on student learning. The logic here is simple: students can't learn if they are not in school. In fact teachers' right to strike is often portrayed as a threat to children's right to learn.

What we actually know about the impact of strikes on student achievement is very little. The few studies available offer conflicting results (e.g., Belot and Webbink 2006, Caldwell and Maskalski 1981, Caldwell and Jeffreys 1983, Johnson 2009, Thornicroft 1994, Zirkel 1992, Zwerling 2008), and the identification strategies in many are vulnerable to critique. Given the small number of studies, the more liberal provision of

strike rights to teachers in countries other than the U.S. and the ongoing campaigns to limit teachers' right to strike in some states,[1] new evidence on this issue is timely.

In this study I examine the impact of teacher strikes on student achievement in the Canadian province of Ontario. Teachers in Ontario have had the legal right to strike since 1975 and have exercised it regularly. Between 1975 and 2005 there have been at least 101 teacher strikes that lasted almost 19 instructional days on average. This includes a general strike by both primary and secondary school teachers in 1997 that for 10 days shut down the public education system in the province. At the start of this period some strikes lasted in excess of 50 instructional days.

The focus here is on 11 strikes in the province's primary schools that occurred in the period 1998/99 through 2005/06. For this same period the results of curriculum based standardized tests that Ontario students write in grades 3 and 6 in reading, writing and mathematics are available. The primary empirical strategy is to compare the grade 3 through grade 6 change in test scores for school cohorts that experience a strike to the corresponding changes for cohorts that do not.

This empirical framework offers a different perspective than the approaches in most previous studies. A more common strategy is a cross sectional comparison of students who experience a strike to those who do not. The criticism of this more common approach is that school boards, schools and students that experience strikes may be different, most importantly in unobserved ways, than those who don't. A standard omitted variables argument raises the possibility that any estimate of the impact of strikes from these comparisons could be biased.

[1] The campaign in Pennsylvania is particularly notable.

The approach adopted here can potentially control for these unobserved factors to the extent they are fixed at the school board, school or student cohort level. I enhance the control for these sorts of factors by also presenting estimates in which school boards that experience strikes are matched with comparable school boards who do not on the basis on the socio-demographic characteristics of the school's local area.

The results indicate that "long" strikes, which last 10 instructional days or more, in grade 6 have significant, negative effects on grade 3 through grade 6 test score growth in reading and especially math. The impact of a strike in grade 6 on math score growth is a reduction of 29 percent of the standard deviation of test scores across school/grade cohorts. The average impact of strikes in grades 2 or 3 on score growth is small, negative and statistically insignificant, although there is some heterogeneity in the impact across school boards. Finally, in either case the estimated impact of strikes using a cross section identification strategy is very small and statistically insignificant.

Teachers and Collective Bargaining in Ontario

Schools in Ontario are organized into 72 school boards, 60 English language and 12 French language.[2] These boards are members of one of two parallel public school systems in Ontario, one secular and one Catholic. Of the 60 English language boards 29 are Catholic and 31 secular. Of the 12 French language boards 8 are Catholic and 4 are secular.

Teachers at these schools belong to one of four professional associations. Two, the Elementary Teachers' Federation of Ontario and the Ontario Secondary School Teachers'

[2] There are also 33 School Authorities that are either small geographically isolated areas or schools associated with specific hospitals.

Federation, represent the elementary and secondary teachers in English language secular schools respectively. The Ontario English Catholic Teachers' Association represents teachers in English language Catholic schools. Finally, L'Association des enseignantes et des enseignants franco-ontariens represents teachers at secular and Catholic French language schools.

Teachers in Ontario gained the right to strike, and school boards the right to lock out teachers, through the *School Boards and Teachers Collective Negotiations Act*, Bill 100, of 1975. The Act specified that collective bargaining should occur at the school board level, between the board and the relevant teachers' association. The *Act* also created the *Education Relations Commission* (ERC) whose duties included overseeing collective bargaining under the *Act*, advising the government when work stoppages might impede students' education and collecting data.[3]

Bill 160, the *Education Quality Improvement Act*, reformed this collective bargaining environment in 1997.[4] This Act repealed Bill 100 and placed collective bargaining under the auspices of province's *Labour Relations Act*, effectively standardizing collective bargaining in the education sector with bargaining elsewhere in the labor market. It also removed principals and vice principals from the bargaining units and reduced the scope of negotiable issues.[5] Finally, the ERC was merged with the *Ontario Labour Relations Board* and its role was reduced to advising government when teacher strikes might jeopardize the school year.

This legislation was part of a package of reforms of the educational sector in

[3] The Act also specified a number of dispute resolution mechanisms particular to collective bargaining in this area.
[4] See Rose (2002) for a detailed overview of this episode.
[5] For example, many workload issues (e.g., class size) were made terms of employment.

Ontario. The government of the time also reduced the amount of money flowing into the educational sector, changed the mechanism used to fund schools, prioritized funding for instruction over administration and introduced the standardized testing of students. These reforms were deeply unpopular with many teachers and led to a 10 day province wide general strike by the teachers in October/November 1997.

An overview of the incidence and duration of teacher strikes in Ontario is provided in figure 1. There is a clear downward trend in the duration of strikes and no real discernable trend in incidence until 2003. Starting in the 2003/04 school year there have been no strikes in the Ontario public school system. This is in part due to the election of a new provincial government in 2002, which had a mandate to bring labour peace to this sector.

The Impact of Strikes

The most straightforward impact of a teacher strike is the withdrawal of educational services during the period of the work stoppage. Unless this time is made up, for example by extending the school year, students lose a corresponding number of days of instruction. However, this is the only impact of a strike on instruction if we can imagine teachers and students making a clean break at the end of the day before a strike begins and then picking up smoothly at the same place at the beginning of the day after the strike ends.

There are a number of reasons to believe this scenario is false. First, many of the strikes examined are in excess of 10 instructional days. This is longer than both the winter and spring holidays in the Ontario school system. So just as students take a while

to get back to work after a school break, we might expect the same to be true after a strike. Second, curriculums are presumably designed to accommodate holidays. Difficult concepts are not left hanging over a break. In contrast, strikes are not anticipated in curriculums, and so extensive review of material may be required to get students back up to speed after a work interruption. Third, strikes may interact with holidays to result in an extended period away from school for the students. Many of the strikes analyzed are either close to the winter break or at the beginning or end of the school year, significantly reducing the period of *sustained* instruction. Finally, in the days leading up to the strike both students and teachers may be distracted further reducing instruction. Similarly, strikes that are ended by legislation may lead to hard feelings or resentment. While difficult to quantify, each of these factors suggest that actual amount of student instruction lost from a teacher strike might exceed the number of days of the actual work stoppage.

Previous Literature

Previous studies of the impact of teacher strikes on student achievement provide a range of estimates. Zwerling (2008) is a recent example of research finding no impact of strikes on student achievement. He studies Pennsylvania schools in the period 2002/03 through 2005/06. During this period the average strike lasted just over 10 days. While the primary empirical framework is cross sectional, specifications using lagged dependent variables and a two year change in a school district's average results at a

specific grade are estimated to control for omitted variables.[6] The main message of this study is no relationship between either strike incidence or duration and student achievement in math and reading. This conclusion is consistent with the findings of Thornicrolt's (1994) study of Ohio and Zirkel's (1992) review of the literature.

Evidence of negative effects is mostly from an earlier period. Caldwell and Maskalski (1981) and Caldwell and Jefferys (1983) are examples that find some negative effects. More recently Belot and Webbink (2006) report that teacher strikes in Belgium had a negative impact on students' subsequent labor market earnings.

Johnson (2009) examines many of the same strikes in Ontario investigated in this study. Using a cross sectional empirical design he reports that strikes have a modest negative effect on student achievement in grade 6. For example, a 10 day strike is estimated to reduce the proportion of students achieving the provincial standard (defined below) on the standardized tests in grade 6 by one percentage point. The negative effect is concentrated in disadvantaged schools as measured by their social and economic characteristics.

Related evidence comes from the literature investigating the impact of teacher absences on student achievement. In this case instructional time is not reduced, but it is disrupted through the use of substitute teachers. Miller et al. (2008) is a recent study examining this issue. They use teacher fixed effects and an instrument based on local weather conditions to account for unobserved teacher quality correlated with absenteeism. They report that each 10 days of teacher absence leads to a decline in math achievement of 3.2 percent of a standard deviation. Clotfelter et al. (2009) and Herrmann

[6] The main dependent variable is the proportion of students scoring "advanced" and "proficient" on the test.

and Rockoff (2009) provide further evidence of negative effects.

From the perspective of this research the lack of consistent evidence of a negative impact of teacher strikes is surprising. Teacher absences cause a disruption in the delivery of instruction but, strictly speaking, no reduction in the amount of material covered.

Data

The data on student test scores are from the Education Quality and Accountability Office (EQAO) accessed through the Public Economics Data Analysis Laboratory (PEDAL) at McMaster University. These are the results of annual testing of grade 3 and 6 students in math, reading and writing, which commenced in the 1996/97 academic year. The EQAO tests are based on the Ontario curriculum, which the public schools in the province deliver. The timing of the tests corresponds to the completion of the curricula of the primary (grades 1-3) and junior (grades 4-6) divisions.

The test results in each subject area categorize students on a four point scale:

1. Level 1 means their skills fall below the provincial standard.
2. Level 2 means they are close to meeting the provincial standard.
3. Level 3 means they are at the provincial standard.
4. Level 4 means they have surpassed the provincial standard.

These data also provide some information on the students writing the tests and the environment in which the tests were written. The students' characteristics are gender, whether the student is special needs, "gifted", enrolled in French immersion or has English as a second language. Information on whether the student wrote the test under

special circumstance (e.g., received extra time) is also provided.

Unfortunately there are not consistent student numbers that allow students to be followed between the grade 3 and grade 6 tests. As a result the primary unit of analysis is the student cohort at the school level. These cohorts will contain primarily, although not exclusively, the same students in grade 3 and 6, because of student mobility.

The test score data used come from the academic years 1998/99 through 2005/06. The test scores for students in the French school boards are excluded from the analysis. The average change in tests scores between grade 3 and grade 6 are much larger for students in these boards than in the other boards with no obvious explanation. Results including the French board schools in the sample lead to larger estimated impacts of strikes.

The sample only includes school cohorts that have at least 15 test takers in grade 3 and grade 6. Also, cohorts are excluded when the number of test takers rises or falls more than 40 percent between grades 3 and 6.

PEDAL has merged data from the 2001 Canadian census with these test score data. The matching is on the basis of the "Forward Sortation Area" (FSA) in which all addresses share the same first three characters of a postal code. Given a school's address, information on the socioeconomic characteristics of residents in the same FSA are linked from the 2001 census.

Information on industrial actions within the Ontario school system was collected from a number of sources. Primary sources of data were databanks maintained by the ERC and the Workplace Information Directorate of the federal government's Department of Human Resources and Skill Development. The final record of industrial actions was

checked and corrected against a data bank maintained by the Ontario Ministry of Education[7] as well as information gathered from publications of various school boards and the teachers' professional associations and legislation passed to end some of the strikes.

Table 1 lists the strikes in Ontario's elementary schools over the period. Starting in 1997/98, which is one year before the first observation on grade 3 test scores, there are 11 strikes. This does not include the 10 day general strike in the fall of 1997, which is not included in the analysis because it affected all schools in the province. It also does not include rotating strikes—which are listed in the notes to the table—because in these cases it is not possible to determine which schools within a school board were affected.

The number of instructional days lost ranges from 3 to 17, with 6 strikes lasting 2 weeks or more. Both small and large boards from both the Catholic and secular systems experience strikes over the period.

Recall that the ERC's mandate is to decide when a strike jeopardizes the school year. This decision was made for two strikes in the sample: the 17 day strike at the Hamilton Wentworth School Board in 2000/01 and the 13 day strike at the Toronto Catholic School Board in 2002/03. More generally, the ERC does not appear to have a set algorithm to make this decision. Between 1975 and 2002 the completed duration of strikes determined by the ERC to threaten students' education year ranged from 13 to 56 days.

In Ontario, instructional days lost through a strike are typically not made up and this is generally true for the strikes in the sample. The exceptions are the Hamilton-

[7] I am indebted to John Donofrio for providing this file.

Wentworth strike of 2000/01 after which five professional development days were made instructional days, and the York Catholic Board strike of 1998/99 that was followed by seven additional instructional days scheduled during holidays.[8]

Finally, a one month strike of custodial staff closed the schools of the Toronto District School Board for 5 days in April 2001. While this closure is not included in the analysis, the results coding it as a strike are very similar to those reported below.[9]

Empirical Framework

Teacher strikes are modeled as an additional input to an education production function in which students' test results depend on individual level and school level characteristics. In levels the test score equation is

$$(1) \qquad T_{gcs} = X_{gcs}\beta + Z_s\delta + S_{gcs}\phi + S_{g-1cs}\mu + \lambda_{cs} + \varepsilon_{gcs}$$

where T_{gcst} is the g^{th} grade (3 or 6) average test score of cohort *c* at school *s*, *X* and *Z* are student/cohort and school/FSA characteristics respectively, *S* is a measure of a strike experienced by cohort *c* in grade *g*, λ_{cs} is an unobserved cohort/school fixed effect and ε_{gcs} is a transitory error term. In truth *S* varies at the board level so will be the same for all cohorts of a given vintage across schools within a given school board. Equation (1)

[8] The primary specification of the strike variable is dichotomous so the days made up in these cases do not affect the coding (since in neither case were all days lost made up). In some regressions a 0/1 variable for "long" strikes of at least 10 days is used. The days made up in these strikes only change the coding of the York Catholic SB strike used in the analysis of the grade 3 outcomes. The effect of this recoding on the results is minimal.

[9] As reported below the only strikes of at least 10 days have a significant impact on students' test scores. The closure of the Toronto DSB due to the custodial strike adds another short strike to the sample.

allows for lagged impacts of strikes experienced at grade *g*-1—grades 2 or 5. While in principle longer lags could be important, empirically a one grade lag is sufficient.

Many previous studies have estimated equations such as (1) using cross section data. A criticism of this approach is that $E[S,\lambda] \neq 0$ leading to bias in the OLS estimates of the parameters of (1).

Taking the grade 6/grade 3 difference of equation (1) at the cohort level yields

(2) $$\Delta_g T_{cs} = \Delta_g X_{cs}\beta + \Delta_g S_{cs}\phi + \Delta_{g-1} S_{cs}\mu + \Delta_g \varepsilon_{cs}.$$

$\Delta_g T_{cs}$ is the grade 3 through 6 change in the average test score of cohort *c* at school *s*; Δ_{g-1} denotes the grade 2 through 5 difference. The school/FSA characteristics, *Z*, difference out in this operation because only one observation is available on these variables from the 2001 census. As explained below some control for these characteristics is made through a matching procedure. The school/cohort fixed effect λ_{cs} also differences out eliminating any bias from the correlation of this effect with the strike variables.

The impact of strikes in (1) is not grade specific. This needn't be so, and isn't so in the data. To accommodate this heterogeneity the impact of strikes is estimated separately for two samples of students. One is the sample of students who experienced strikes in grades 5 and 6. The equation estimated is

(3) $$\Delta_g T_{cs} = \Delta_g X_{cs}\beta + S_{6cs}\phi_6 + S_{5cs}\mu_5 + \Delta_g \varepsilon_{cs}.$$

Because these students did not experience strikes in grades 2 or 3 $\Delta_g S_{cs}$ becomes S_{6cs}. All else equal these strikes can only have an impact on grade 6 test scores. The expected impact of these strikes is to depress grade 6 test scores and so also $\Delta_g T_{cs}$, the grade 3 through 6 change in scores.

The other sample is students who experienced strikes in grades two or three. The equation estimated is

$$(4) \qquad \Delta_g T_{cs} = \Delta_g X_{cs}\beta + S_{3cs}\phi_3 + S_{2cs}\mu_2 + \Delta_g \varepsilon_{cs}.$$

In principle, a strike in these grades could affect both the grade 3 and grade 6 test scores. However, (2) imposes the empirically verified restriction of lagged strike effects of one grade, meaning any estimated impact of a strike in grades 2 or 3 is interpreted as an effect on the grade 3 score. Here a negative impact of a strike on grade 3 scores would lead the affected cohorts to exhibit a larger than otherwise grade 3 through 6 score change. For these regressions the dichotomous strike variables are coded 0/-1 so the expected impact of a strike is negative consistent with the grade 5/6 results.

In each regression the school/cohorts experiencing strikes are pooled with school/cohorts that experienced no strikes in these grades.[10] This selection of a "control group" implicitly defines the impact(s) of a strike being identified. School boards that do not experience a strike may nevertheless experience labor tension or more formally a "work to rule" campaign during which teachers services that are not directly specified in the employment contract are withdrawn.[11] Therefore, under certain assumptions we might interpret any impact of a strike as a result of the withdrawal of teaching services.

[10] An alternative would be to pool all the data and define variables to separately capture the impact of the strikes in the different grades. This option is not pursued because estimating the effects in separate samples is a convenient expository device and also facilitates well defined control groups. For example, when estimating the impact of strikes in grade 5 or 6 any cohorts who were affected by strikes in grade 2 or 3 are omitted from the control group. In a pooled sample these cohorts will be part of the control group for the estimation of the grade 5 and grade 6 strike effects. This said, pooling the data leads to estimates that are almost identical in both magnitude and statistical significance.

[11] Examples of the types of services withdrawn are supervision of co- and extra-curricular activities and attendance at administrative meetings.

However, other factors may also differ between strike and no strike schools that also contribute to an impact. Note that these issues of interpretation arise in most analyses of the impact of strikes on different outcomes.

This discussion also begs the question of why labor tension boils over into a strike at one board and not at another. The primary concern for the analysis is if this is due to unobserved factors that also affect student test scores. The working assumption of the analysis is that these unobserved factors are fixed at the cohort/school fixed level and therefore difference out in (2).

There are a number of criticisms of this approach. First, it may be unreasonable to assume unobserved cohort characteristics are fixed at such young ages. A more defensible assumption is that any unobserved school/area characteristics are fixed over a 3 year interval. Second, it is not possible to follow individual students over time, and the individuals observed at grade 6 in a cohort will not be exactly the same students observed at grade 3. Therefore, it is necessary to assume that any changes in the unobserved cohort/school factors as a result of student mobility are uncorrelated with the included regressors. While the issue of student mobility is investigated more completely below, note here that because strikes occur at the board level, it not possible to simply switch schools locally to avoid industrial actions.[12] Finally, there may be unobserved time effects coincident with a strike that impact student scores. These sorts of unobserved effects cannot be accommodated in a fixed effect framework. That said, one of the

[12] Another local possibility would be to switch from the secular to the Catholic school system (or vice versa) to avoid, or as a result of, industrial action. However, entrance to Catholic schools is limited to those with Catholic backgrounds and those in Catholic schools have additional reasons to remain in them.

objectives of the paper is to introduce the fixed effect specification to the analysis of teacher strikes and compare the results to the more common cross sectional analysis.

The primary specification of *S* is as a 0/1 dummy variable indicating a strike was experienced during the specified grade. The measures of student achievement (the dependent variable) are a cohort's average score on the reading, writing or math test. The controls for cohort attributes, *X*, are the proportions of the cohort who are female, receiving special support, "gifted", enrolled in French immersion, have English as a second language and who wrote the test.[13] There is also a full set of year effects, a dummy variable for Catholic boards and their interactions.

Note not all the strikes listed in table 1 can be used in each analysis. $\Delta_g T_{cs}$ links grade three and six test scores that are three calendar years apart. When (4) is estimated using the cohorts who experience strikes in grade 2 or grade 3 all the strikes in the table can be coded. However, when estimating (3) for the sample of cohorts who experience strikes in grades 5 or 6, only the strikes in 2000/01 through 2002/03 can be used. To see why, consider a cohort that experienced the 1998/99 strike at Waterloo Region District School Board when they were in grade 6. To construct ΔT_{cs} for this cohort would require their grade 3 scores from 1995/96. However, the data on student tests cores starts in 1998/99.[14]

[13] Information on whether students attended kindergarten is also available for the years 1998-2004. Including a control for the proportion of a cohort who attended kindergarten in regressions using data from this shorter period has no effect on the results for grade 6. In the grade 3 results the estimates are very similar, although the point estimate of the contemporaneous effect of a strike is more likely negative (but still statistically insignificant).

[14] Note if a cohort experienced the one of the 2000/01 strikes in grade 6 data from 1997/98 would be required to construct ΔT_{cs}. Therefore, the strikes in this year can only be used as a one year lagged strike for the cohorts who were in grade 6 in 2001/02.

The regressions are estimated by OLS and standard errors are robust to clustering at the school board level. A set of estimates when boards are matched on the basis of their school/FSA characteristics, *Z*, is also presented. The characteristics available from the 2001 census are average household income, the percentage of households headed by a single mother, the percentage of the population age 20 or older who do not have a high school diploma, the percentage of the population who are immigrants and arrived in Canada in 1981 or later, the percentage of the population who are visible minorities, and the percentage of the population age 16 and up who are unemployed.

Results

In table 2 is a comparison of the mean values of cohorts' average standardized test results between school boards that experienced a strike over the period and those that did not. Both the average score on each test and the proportion of students achieving the provincial standard (test scores of 3 or 4) are reported. For example, the average score on the grade 3 math test is very similar in cohorts that experienced a strike in grade 2 or 3 and cohorts that did not experience a strike over the period—2.69 and 2.68 respectively. The proportion of students achieving the provincial standard on this test at 0.63 is identical in the two groups.

More generally, there are few differences on average in the test outcomes in grade 3 between strike and no strike boards. The most notable is the higher achievement in cohorts that experienced a grade 2 or 3 strike in the writing test. The comparison for cohorts that experienced a strike in grade 5 or 6 turns up more differences although most are quite small. Here achievement tends to be marginally higher for the no strike cohorts.

In table 3 are corresponding comparisons for student level control variables used in the analysis. In most cases the averages for the strike and no strike cohorts are very similar. An exception is the proportion of cohorts that are in Catholic schools, which is higher in strike boards, markedly so for the cohorts that experienced a strike in grade 5 or 6. The source of this latter discrepancy can be seen in table 1. Four of the seven strikes for the analysis of grade 5 and 6 strikes occurred in Catholic boards.

Note that the share of the cohorts writing the various tests is similar in the two groups. There is no evidence here that students in cohorts affected by a strike were excused from writing the tests.

An initial set of estimates of equation (3) for students who experienced strikes in grades 2 or 3 are reported in the first panel of table 4. Recall that the control group for these regressions is students who did not experience strikes and the estimated impact is on the grade 3 through grade 6 growth in average test scores. In the first two columns of the first panel are the results using the full sample of boards that experienced these strikes. For each outcome there is no evidence of an impact of a strike in grade 3. All the estimates are statistically insignificant, and all are very small. Somewhat surprisingly, however, there is some evidence of a lagged grade 2 strike effect. The impact is consistently negative and statistically significant at the 10 percent level for the math score growth, indicating a average reduction in the grade 3/grade 6 change in scores of 0.06 points which is 2.2% of the mean score for the no strike cohorts reported in table 2.

One way of trying to discover whether the estimated impact of strikes is causal is to look at the impact of longer strikes. Intuitively any adverse impact of a strike would be increasing to some extent in its length. In the next two columns are the results for

students who experienced "long strikes" defined those lasting at least 10 teaching days or two weeks of school.[15] The estimates for strikes in grade 3 remain very small and statistically insignificant. The estimates for the grade 2 strike effect are all larger (in absolute value) then their counterparts in the first two columns and all statistically significant. The largest impact is for the change in the math score, a reduction of 0.10 points. This perspective suggests that this lagged strike effect is casual.

In the next panel are the estimates using a matching estimator based on the socio-demographic characteristics of the boards. For these estimates a propensity score for experiencing a strike is estimated for each board as a function of the socio-demographic characteristics in the school's FSA and the catholic dummy. The scores are then used to reweight the data for cohorts at the strike and no strike boards respectively.[16] As an alternative, the census socio-demographic characteristics were entered directly into the regression as additional control variables. In this case the estimated parameters on these variables capture the grade 3/grade 6 difference in their impact on the test scores. The results using this alternative method are almost identical to those using the matching method.[17]

The results and the pattern of statistical significance of the estimates from the matching estimator largely match what is seen in the first panel, although in each case the results are "more negative" in moving from the sample of all strikes to the sample of long strikes. The inference remains that there are small and statistically insignificant effects

[15] For these estimates the boards experiencing shorter strikes are omitted from the sample.
[16] See Imbens (2004) on this approach.
[17] An exception is the contemporaneous impact of long strikes on the grade 6 reading test (table 5) is slightly larger: -0.070 (0.023).

for strikes in grade 3—the year the grade 3 EQAO test is written--and that the lagged strike effects are larger, negative and in many cases statistically significant.

The finding that grade 2 strikes affect grade 3 through 6 test score growth but grade 3 strikes do not is surprising. Further investigation identified one board as critical to this result—the Lambton-Kent District School Board. This board experienced a strike of 14 days in April and May of 2001. In the next panel of table 4 are the estimates when this board is dropped from the sample. The impact is quite dramatic. First the estimates for a strike in grade 3 remain statistically insignificant. Second, the estimates of the lagged, grade 2, strike effects are now much smaller and none are statistically significant. It is clear that this board makes a substantial contribution to the lagged estimates.

The time series of the Board's grade 3 results provides a nice demonstration of the point. In figure 2 is the proportion of the grade 3 students achieving the provincial standard on the math test—a score of 3 or 4—taken from the Board's website.[18] The province wide "pass rate" in each year is also reported. Clear in the figure is a dip in the proportion of the Board's students achieving the standard in 2001/02, the year after the strike took place, relative to the province wide results. In all other years the Board and the provincial pass rates effectively coincide.

Discussions with the Board did not reveal the reason for this dip, nor whether it was related to the strike. The administration at the Board had turned over in the intervening period, so that no one could directly recall the circumstances surrounding the strike or the deviation in grades.

[18] Accessed on January 13, 2010 at http://www.lkdsb.net/Parents/eqao_results1.htm.

In the final panel of the table are the results estimated at the board level. Estimating the regressions at the board level captures any student migration across schools within a board. This might be an important consideration if students fled schools particularly hard hit by a strike. It also has the effect of weighting each strike within the sample equally. In the previous results strikes at larger boards implicitly get greater weight because they have more student cohorts. This can be important if there are heterogeneous impacts of a strike across boards. One way of interpreting any difference in results across these approaches is that the cohort level regressions estimate the impact of a strike on the average cohort that experienced one, while the board level regressions estimate the impact for the average board.

The actual differences between the results in panels 2 and 4 are small but potentially meaningful. First the grade 2 strike effects—due to the Lambton-Kent District Board—remain large and negative, although now mostly statistically insignificant because the standard errors are larger. Second, the impact of a grade 3 strike is now consistently negative, modest and statistically insignificant. Third, in general the point estimates indicate larger negative impacts in the sample using long strikes.

The difference between the cohort and board based results do appear to reflect some heterogeneity in the impact of strikes across boards. Further analysis (not reported) reveals that the impact of the 1998/99 strikes on test score growth is consistently negative, sizable and statistically significant for long strikes.[19] The impact of the 2000/01 strikes is generally very small, statistically significant and negative for reading and

[19] The contemporaneous, negative impact of long strikes ranges from 4 to 6 percentage points on the three tests.

writing. Finally the impact of the 2002/03 strikes is generally positive, statistically insignificant and quite small for long strikes. The anomalies here are the short strike at Windsor Essex, which is estimated to have had a large positive impact on score growth, and the longer strike at Toronto Catholic, which is estimated to have had a small positive impact. Each of these boards gets relatively large weight in the cohort based regressions.

Initial estimates of equation (4) for strikes experienced in grade 5 or 6 are reported in the first panel of table 5. Recall the distinction here is that the strike is experienced in the year of, or the year before, the grade 6 EQAO tests, while the results in table 4 are for strikes experienced in the year of, or year before, the grade 3 EQAO tests. There is a consistently negative impact of a grade 6, although only the result for writing is statistically significant. The grade 5 strike effect is much smaller, statistically insignificant and mostly negative. Once the sample is limited to long strikes, the impacts of grade 6 strikes are larger and now all statistically significant. The estimated impact on math score growth is the largest. It is a reduction of 0.10 points for longer strikes, which is 3.8 percent of the mean grade 6 math score in the no strike boards (table 2), or 29 percent of the standard deviation of grade 6 math test scores across these cohorts. The grade 5 strike effects are also now larger although still substantively small.

In the next panel of table 5 are the results when the boards are matched on the basis of the socio-demographic characteristics of their districts. The parameter estimates are largely the same in these results although the standard errors are generally larger. As a consequence some of the estimates are no longer statistically significant; most notably

the results for the writing test.[20] The results for the math test score growth continue to be larger and statistically significant. New here for this test is evidence of an impact of a grade 5 strike of over 4 points.

As for the results of the strikes in grade 3 and 2 it is worth asking whether a particular board is driving the results. As seen in table 1 there are three strikes the cohorts sampled experienced in grade 6, and seven strikes that they experienced in grade 5. Among the strikes experienced in grade 6 strikes (in 2002/03), two are over two weeks and one is one week long. The comparison of the full and long strike samples reveals it is the two longer strikes that are pivotal.

To discover the relative impact of the two longer strikes, in the third panel are the results when the Simcoe Muskoka Catholic School Board is deleted from the sample. In this case it is only the Toronto Catholic School Board identifying the impact of the long strike. The results suggest that while both of the longer strikes play a role in the results in the second panel, the impact of a grade 6 strike is smaller for the Toronto Catholic Board.

This result makes some sense given the timing of the two strikes within the school year. The Simcoe Muskoka strike occurred between November 15 and December 2 of 2002. After the strike students had just over 2 and half weeks of instruction leading into the two week winter holiday. In contrast the Toronto Catholic strike occurred between May 15 and June 3 2003, which is around the time the EQAO tests are written. Therefore a contemporaneous effect of this strike on grade 6 scores is more likely through a distraction than a reduction in material covered. That said, the number of days of instruction lost through this strike is almost three weeks, and upon return to school

[20] The writing results remains statistically significant if the socio-demographic characteristics are entered directly as explanatory variables in the regression.

there were just three and half weeks of instruction before the two month summer holiday. In each case the disruption caused by these strikes was enhanced by its interaction with school holidays.[21]

In the final panel of the table are the results of the board level regressions. Many of the estimates are larger than their cohort based counterparts. The standard errors are also substantially larger, but statistically significant impacts of long strikes in grade 6 are still observed for the math and reading test score growth.

How does a long strike impact different students? In table 6 the effect of a long strike in grades 5 or 6 is estimated for different metrics of student achievement.[22] In the first panel are the results for the provincial "pass rate", defined as a score of 3 or higher on the test. For math the impact is a reduction in the grade 3/grade 6 change in the pass rate of just over 5 percentage points. This is over 8½ percent of the average pass rate in the grade 6 math test among the no strike cohorts (table 2). There is also an impact of a strike in grade 5 of 3 percentage points. The impacts of a grade 6 strike on the change in pass rates in reading and writing are smaller reductions of 1-2 percentage points.

[21] Another factor is that each of these strikes was preceded by a work to rule campaign by teaching staff. During these actions teachers literally refuse to complete any duties that are not explicitly outlined in the work contract. This typically means that teachers do not perform any administrative functions or attend meetings and they will not supervise any co- or extra-curricular activities or field trips. The withdrawal of administrative functions can disrupt student feedback such as report cards. During these actions teachers are instructed to arrive at school just before classes commence and leave shortly afterwards. This would appear to compromise any extra academic assistance provided to students, for example those at risk of falling behind. While a work to rule campaign does not change the days of instruction children receive, it does change the context in which they are delivered and the general atmosphere at the school. In the case of Simcoe Muskoka the work to rule campaign preceding the strike lasted 19 instructional days, while in the case of Toronto Catholic it lasted 9 days.

[22] The results for grade 3, available on request, do not shed any additional light on the average results reported in table 4.

The remaining panels of table 6 show how the impact of a long strike plays out across the distribution of test scores. In math the impact of a strike is seen across the score distribution: clear negative impacts on the grade 3/grade 6 change in the proportions of students at the highest two scores categories and positive impacts on the change in proportions of students at the lowest two scores. Note also that there is an impact of a grade 5 strike at the bottom of the score distribution. For the other tests the shifts in the proportions are more modest. The impact in writing is largely a redistribution of the top students between scores 3 and 4.

One explanation of these different results across the tests is a math curriculum that is more structured and sequential than the curricula in reading and writing. For math a strike could prevent certain components of the curriculum from being covered, or require that they be introduced too quickly for many students. Achievement in writing and reading may be more a long term function of practice and experience, so a strike does not so sharply disrupt the course of instruction.

Finally, in table 7 are results by gender. There is a notable gender difference in the point estimates for the math score growth. Males experience the larger impact of a grade 6 strike, and there is no evidence of a lagged impact of a strike on the female math score growth.

Sensitivity Tests

Given that it not possible to follow students from grade 3 to grade 6, there is reason to wonder if there are significant changes in the composition of school cohorts between these grades that are correlated with the strike variable. These could occur, for

example, if boards/schools that experience strikes strategically excuse students from writing the tests to maintain their levels of achievement. Note the simple mean comparisons of the proportion of cohorts writing the tests reported in table 3 do not suggest this is an important issue, and the change in the proportion of the cohort that writes a given test is a control variable in the regressions. Alternatively students who experienced a strike might decide to switch schools or school boards. Others might decide to leave the public school system for a private school.[23] Again note that the regressions estimated at the board level, which captures inter school mobility within boards, confirm small, negative but statistically insignificant impacts of strikes in grades 2 and 3 and larger, negative statistically significant impacts of strikes in grades 5 and 6.

There are a number of strategies available to address these issues.[24] The first is to see if cohorts that exhibit large changes in the proportion of test takers between grades 3 and 6 have a substantive effect on the results. To do this a new sample is created that excludes cohorts that experienced changes in the proportion of test takers that fall outside the 5^{th} and 95^{th} percentiles of the distribution. The estimates from this new sample are reported in table 8.[25] The point estimates are very similar to the results reported in the second panels of tables 4 and 5, although not surprisingly the standard errors are larger.[26]

---

[23] Another check that was performed was whether the results are sensitive to weighting. The reported regressions are not weighted by cohort size. The mean outcome from a small school/cohort is given the same weight as the mean from a large school/cohort. Weighting the data by cohort size makes very little difference to the results. These estimates are available on request.

[24] Some of the strategies follow the analysis of Card et al. (2010) using these same data.

[25] These and the rest of the estimates reported in this section are from the matching estimator.

[26] As a more severe accommodation the regressions for grade 6 were re-estimated excluding any cohort that exhibits a change on the proportion of students writing the test that was greater than +/- 3 percent. This excludes more than two-thirds of the original

Next, it is possible to investigate changes in the size of school cohorts that are correlated with the strike variable. Using data from the Ontario Ministry of Education estimates of enrollment by grade, by school as of the fall of each school year are available. In table 9 are the results when the dependent variable is defined as the percentage change in cohort enrollment between grade 3 and grade 6 based on these data. Additional lags of the strike variables are specified in these regressions because the decision to move schools in response to a strike may have a longer gestation period then the impact of strikes on student achievement.[27] There is very little evidence here of statistically significant or substantive impact of strikes in grades 1 through 6 on changes in cohort enrollment.

Another perspective on this same issue is provided by examining the choices of students in a school's FSA. Students in an FSA have some choice whether to enroll in the local English or French language school, or the English options to enroll in a public or Catholic school, and this choice may be affected by teacher strikes. In the first row of table 10 the impacts of strikes in grades 4 through 6 on the grade 3 through grade 6 change in the proportion of the FSA's students that are enrolled in English language schools is investigated. Note that all these strikes over the period occurred in English language boards. There is no evidence here that students "flee" the English Language boards in response to the strikes over the period.

In the next row the dependent variable is the change in the proportion of the FSA students who are enrolled in English language *secular* schools. In this case I need to

---

sample. With the exception of results for the writing test, the point estimates are very similar to those reported in table 7 although the standard errors are larger.

[27] With the additional lags a 7 day strike in the Brandt Haldimand DSB in 1996 is included in the sample.

distinguish between strikes occurring in secular (public) and Catholic boards, because presumably the proportion would be increasing in a strike in a local Catholic Board and decreasing in strike in a local public board. Note that the impact of a grade 6 strike in a public board is not identified since all strikes between 2001 and 2005 were in Catholic boards. Here there is some weak evidence that a strike experienced in grade 5 leads to a small decrease (<1 percentage point) in the change in the proportion of eligible students enrolled in public school. Note also, however, that the estimated impact of strikes at Catholic schools are "wrong signed" as they suggest a resulting decrease in the change in enrollment in public schools.

The corresponding results for strikes in grade 1 through 3 are reported in the second panel. The estimates are generally very small and statistically insignificant, with the exception of the small, "wrong signed" impact of strikes in grades 1 and 2 at Catholic schools.

A final exercise is to examine data that matches students' test scores in grade 3 and grade 6. Starting in 2004 the EQAO uses unique student identifiers so that it is possible to match grade 3 and grade 6 test scores for a specific student. One year of this matched data is now available for students who wrote their grade 3 tests in 2004 and grade 6 tests in 2007. Using these data it is possible to estimate the proportion of grade 3 test takers who are in the same school or the same school board when they write the grade 6 tests.

An analysis of these data is presented in table 11. The sample is chosen to match the analysis sample used in the regressions.[28] Because strikes occur at the board level, the primary focus is the comparison between the full sample of students who write a test and the sample of students who are in the same school board in grade 3 and grade 6. Of the groups of students who wrote tests in either grade 3 or 6 83-86 percent were in the same school board in both years.[29] The average test scores of students who remain in the same board for grades 3 and 6 are marginally higher than for the full sample of students, although the difference is quite small. Similarly, the average test score change between grade 3 and grade 6 is very similar in the two samples, although there are systematically marginally better outcomes—by less than one half a percentage point—in the sample of same school board students.

The message of these different analyses is there is no large systematic changes in cohort composition between grade 3 and grade 6 that are correlated with the incidence of teacher strikes. This supports the working assumption of the analysis that any changes in the unobserved cohort/school factors as a result of student mobility are uncorrelated with the included regressors.

Cross Section Estimates

One of the distinguishing features of this study is the attempt to control for unobserved differences between students who experience a strike and those who do not through the use of cohort fixed effects. As noted in the Introduction the more common

[28] Schools with more than 10% missing student identifiers were also dropped to facilitate the matching of students between 2004 and 2007.
[29] The percentage of students in the same school is 70-72 percent. This is very similar to the estimates of Card et al. (2010) from these data for a sample of urban schools.

approach in the literature is to make cross section comparisons of students who do and do not experience a strike. It is worth asking therefore whether this difference in strategy makes a difference here.

In table 12 are the cross section estimates from the same sample of cohorts and strikes used for estimates in tables 4 and 5. The impact of strikes in grade 2 and grade 3 on grade 3 test results, and the impact of strikes in grade 5 and grade 6 on grade 6 test results, are estimated. For strikes in grade 3 the cross section results are largely similar to those in table 4. The estimates in each case are small and statistically insignificant. Where the two approaches disagree is for the effect of a strike in grade 2. The negative impact reported in table 4, which is driven by the Lambton-Kent DSB and is clear in figure 1, is missed in the cross section approach.

For strikes in grade 5 and grade 6 there is more general disagreement between the two approaches. First, the cross section approach does not capture the negative statistically significant impact of grade 6 strikes reported in table 5. Second, while all the cross section estimates are statistically insignificant, the point estimates are generally larger for a strike in grade 5 than a strike in grade 6, just the opposite of the result in the second panel of table 5. Finally, in the cross section results the point estimates are not generally larger for longer strikes as was seen in table 5.

Because the there is no need to follow cohorts over time in the cross section approach it is possible to construct estimates using all the strikes and years of data that are available. This does not add any new strikes to the analysis of strikes in grade 2 and grade 3, but does add the 1998/99 strikes to the analysis of strikes in grade 5 and grade 6.

Table 13 contains the cross section results using the larger samples. These results are very similar to those in the previous table, documenting the robustness of the cross section results. There is no evidence that strikes have an effect on student test scores.[30]

Discussion

The results show a statistically significant, negative effect of grade 6 teacher strikes in excess of 10 instructional days on students' test score growth between grade 3 and grade 6. The largest impact is on the math scores. A long strike, which averaged 12 instructional days in this sample, decreased math score growth by 3.8 percent of the mean. The proportion of students achieving the provincial standard on this test decreased by almost 9 percent. There is also evidence of a smaller effect of strikes in grade 5 on grade 3 through grade 6 test score growth in math. These results suggest that the primary impact of a teacher strike is in the year they occur. One year later there is no impact for 2 of the 3 subject areas.

The estimated impact of strikes in grades 2 and grades 3 are smaller, statistically insignificant, and in the board level regressions consistently negative. Given the standard errors it is possible to rule out negative effects in excess of 5-10 percent of the standard deviation in test scores across cohorts.

Heterogeneous impacts of changes in educational inputs across primary grades have been observed in other contexts (e.g., Angrist and Levy 1999, Rivkin et al. 2005,

[30] Johnson (2009) who also uses EQAO data and analyzes many of the strikes considered here uses a cross section estimator. He reports small and statistically insignificant impacts for grade 3 test scores, but statistically significant results for grade 6 scores. The estimated impact is much smaller than the results here: a 10 day strike reduced the provincial pass rate by about 1 percentage point.

Ding and Lehrer 2010). The difference of the impact of a strike in grades 3 and 6 observed here could result from a number of factors. One consideration is the pace and the complexity of the curriculum in different grades. The timing of the tests in grade 3 and 6 marks the completion of the "primary" and "junior" division curriculums respectively. The results are consistent with a system in which the primary division curriculum is slower and involves repetition to establish the foundations of each subject. As a result interruptions in the school year are less consequential. At higher, junior division, grades increased pace and complexity mean interruptions lead to reduced or much higher paced coverage of certain subject areas.

This is certainly consistent with parents' intuition that taking your child out of school for discretionary reasons such as a vacation is less costly at earlier grades, although there does not appear to be any research supporting this view. What research does indicate is that student absences decline monotonically starting in kindergarten, leveling out at about grade 5 (e.g., Easton and George Jr. 1982, Romero and Lee 2007).

There is also a heterogeneous impact of strikes across subject areas. This again is observed for educational inputs in other contexts (e.g., Dee and Jacob 2010). Interpreting the impact of a teacher strike as the impact of a withdrawal of classroom instruction, the estimates indicate that progress in math is more dependent of this type of instruction than progress in reading or writing. This is consistent, for example, with educational campaigns that emphasize the importance of reading at home to children's command of this skill.

Conclusions

Many previous studies of the impact of teacher strikes indicate that they have little impact on student achievement. This is surprising, because evidence from related literatures indicates that interruptions such as teacher absences affect student outcomes. However, much of the previous evidence on teacher strikes is potentially affected by omitted variables bias.

The estimates in this paper, which control for unobserved fixed student characteristics at the school/cohort level, indicate that teacher strikes have a sizable negative effect on test scores at the late primary level (grade 6 scores). The impact is largest for math at 29 percent of the standard deviation of test scores across student cohorts. It is experienced by students across the skill distribution and there is a lagged impact that affects weaker students. Smaller impacts are estimated for scores in reading and writing tests.

The estimated impact on scores at earlier grades (grade 3) is on average close to zero and statistically insignificant. In this case there is evidence of some heterogeneity of the impact across school boards. For some school boards the impact is similar to the estimates for later grades, while for others there is a zero or even small positive impact.

A comparison to the previous literature is drawn by re-estimating the impact of teacher strikes using a cross section identification strategy. For either grade the estimated impact using this approach is small and statistically insignificant, demonstrating the value of accounting for unobserved effects.

References

Angrist, Joshua D. and Victor Levy, (1999), "Using Maimonides' Rule to Estimate the Effect of Class Size on Scholastic Achievement", *Quarterly Journal of Economics*, 114(2), 533-575.

Belot, Michele and Dinand Webbink, (2006), "The Lost Generation: The Effect of Teachers Strikes on students Evidence from Belgium", University of Essex.

Caldwell, W. E., and M.D. Moskalski, (1981), "The effect of school district strikes on student achievement", *Government Union Review*, 2(4), 3-14.

Caldwell, W. E., and L. Jeffreys, (1983), "The Effect of Teacher Strikes on Student Achievement: New Evidence", *Government Union Review*, 4(1), 40-58.

Card, David, Martin D. Dooley, and A. Abigail Payne, (2010), "School Competition and Efficiency with Publicly Funded Catholic Schools", *American Economic Journal: Applied Economics*, 2(4): 150–76.

Clotfelter, Charles T., Ladd, Helen F. and Jacob L. Vigdor, (2009), "Are Teacher Absences Worth Worrying About in the United States?" *Education Finance and Policy*, 4(2) 115-149.

Colasanti, Michael, (2008), *State Collective Bargaining Policies for Teachers*, Denver: Education Commission of the States.

Dee, Thomas S., and Brian Jacob, (2010), "The Impact of No Child Left Behind", Mimeo, University of Virginia and University of Michigan.

Ding, Weili and Steven F. Lehrer, (2010), "Estimating Treatment Effects from Contaminated Multiperiod Education Experiments; The Dynamic Impacts of Class Size Reductions", *Review of Economics and Statistics*, 92(1): 31–42.

Easton, John Q. and Engelhard George Jr., (1982),"A Longitudinal Record of Elementary School Absence and Its Relationship to Reading Achievement", *The Journal of Educational Research*, 75(5), 269-274.

Eberts, Randall W. and Joe A. Stone, (1987), "Teacher Unions and the Productivity of Public Schools", *Industrial and Labor Relations Review* 40(3), 354-363.

Herrmann, Mariesa and Jonah Rockoff, (2009), "Work Disruption, Worker Health and Productivity: Evidence from Teaching" Columbia University.

Hoxby, Caroline M., (1996), "How Teachers' Unions Affect Education Production", *Quarterly Journal of Economics*, 111(3), 671-718.

Imbens, Guido W., (2004), "Nonparametric estimation of Average Treatment Effect

Under Exogeneity: A Review", *Review of Economics and Statistics*, 86(1), 4-29.

Johnson, David, (2009), "How do Work Stoppages and Work-to-Rule Campaigns Change Elementary School Assessment Results?" Manuscript, Wilfred Laurier University.

Lawson, Alexandra, Randall Penfield and Philip Nagy, (1999), Relating Attitudes, Gender and Student Achievement in Grades 3 and 6", EQAO Research Series No. 1.

Lovenheim, Michael F., (2009), "The Effect of Teachers' Unions on Education Production: Evidence from Union Election Certifications in Three Midwestern States", *Journal of Labor Economics*, 27(4), 525-588.

Miller, Raegen T. , Richard J. Murnane, and John B. Willett, (2008), "Do Teacher Absences Impact Student Achievement? Longitudinal Evidence from One Urban School District", *Educational Evaluation and Policy Analysis*, 30(2), 181-200.

Rivkin, Steven G., Eric A. Hanushek and John F. Kain, (2005), "Teachers, Schools, and Academic Achievement" *Econometrica*, 73(2), pp. 417-458.

Romero, Mariajose and Young-Sun Lee, (2007), "A National Portrait of Chronic Absenteeism in the Early Grades", National Center for Children in Poverty.

Rose, Joseph B., (2002), "The Assault on School Teacher Bargaining in Ontario" *Relations Industrielles/Industrial Relations* 57(1): 100-128.

Thornicroft, Kenneth, (1994), "Teachers Strikes and Student Achievement: Evidence from Ohio" *Journal of Collective Negotiations* 23(1), 27-40.

Zirkel, P., (1992), "The academic effects of teacher strikes", *Journal of Collective Negotiations in the Public Sector*, 21(2), 123-138.

Zwerling, Harris L., (2008), "Pennsylvania Teachers' Strikes and Academic Performance" *Journal of Collective Negotiations* 32(2): 151-172.

Table 1: Strikes in Ontario Elementary Schools 1997/98 through 2005/06

Year	School Board	Strike Days	Make up Days	Preceded by WTR (days)	Students Affected
1998/99	Simcoe County DSB	9		No	6160
1998/99	York Catholic SB	15	7	No	31011
1998/99	York Region DSB	5		No	54,412
1998/99	Waterloo Region DSB	5		No	36,395
2000/01	Hamilton-Wentworth DSB	17	5	No	40,587
2000/01	Sudbury Catholic SB	3		7	5,952
2000/01	Keewatin-Patricia DSB	10		No	4,606
2000/01	Lambton-Kent DSB	14		19	19,037
2002/03	Simcoe Muskoka Catholic SB	11		19	15,475
2002/03	Toronto Catholic SB	13		9	68,554
2002/03	Windsor-Essex Catholic SB	5		10	18,411

Notes: A general strike of all boards, October 27-November 7 1997, is excluded from this list. The 1998/99 strikes exclude rotating strikes of one day in the York Catholic School Board and Halton District School Board, of 6 days in the York Region District School Board and a one day full strike in Thunder Bay District School Board before the school year.

Table 2: Mean Values of Standardized Test Scores

Grade 3 Results	Strike in Grades 2 or 3	No Strikes
Number of Cohorts	2335	8903
Math Score	2.69	2.68
	(0.34)	(0.35)
Reading Score	2.51	2.49
	(0.29)	(0.29)
Writing Score	2.66	2.63
	(0.22)	(0.22)
Math Pass Rate	0.63	0.63
	(0.19)	(0.20)
Reading Pass Rate	0.56	0.55
	(0.17)	(0.17)
Writing Pass Rate	0.60	0.58
	(0.15)	(0.16)
Grade 6 Results	Strike in Grades 5 or 6	No Strikes
Number of Cohorts	1054	8903
Math Score	2.63	2.66
	(0.38)	(0.35)
Reading Score	2.63	2.65
	(0.28)	(0.27)
Writing Score	2.63	2.63
	(0.24)	(0.24)
Math Pass Rate	0.59	0.61
	(0.19)	(0.19)
Reading Pass Rate	0.62	0.64
	(0.16)	(0.16)
Writing Pass Rate	0.58	0.59
	(0.15)	(0.16)

Notes: The tests are marked on a four point scale. A "pass" is defined by achieving the Provincial standard, which is a score of 3 or 4. The reported means are the average of the school/grade cohort means in the indicated school boards.

Table 3: Average Characteristics of the Students Writing the Standardized Tests

Grade 3 Characteristics	Strike in Grades 2 or 3	No Strikes
Female	0.49	0.49
Special Support	0.11	0.10
Gifted	0.01	0.00
English as a Second Language	0.03	0.03
French Immersion	0.02	0.05
Catholic	0.38	0.31
Share of Cohort Completing Math Test	0.91	0.92
Share of Cohort Completing Reading Test	0.88	0.88
Share of Cohort Completing Writing Test	0.91	0.91
Grade 6 Characteristics	Strike in Grades 5 or 6	No Strikes
Female	0.49	0.49
Special Support	0.09	0.10
Gifted	0.02	0.01
English as a Second Language	0.03	0.03
French Immersion	0.03	0.05
Catholic	0.84	0.31
Share of Cohort Completing Math Test	0.92	0.92
Share of Cohort Completing Reading Test	0.91	0.91
Share of Cohort Completing Writing Test	0.88	0.91

Notes: The reported means are the average of the school grade cohort means in the indicated school boards.

Table 4: Estimates of the Impact of Teacher Strikes in Grade 2 and Grade 3 on the Grade 3 through Grade 6 Change in Student Test Scores

	All Strikes		Long Strikes	
	Strike in Grade 3	Strike in Grade 2	Strike in Grade 3	Strike in Grade 2
Base Results				
Math	0.005	-0.061*	0.011	-0.102**
	(0.024)	(0.032)	(0.030)	(0.044)
Reading	0.006	-0.025	-0.004	-0.053**
	(0.020)	(0.021)	(0.027)	(0.026)
Writing	0.001	-0.039	0.004	-0.082**
	(0.015)	(0.025)	(0.020)	(0.020)
Matching Results				
Math	0.005	-0.056*	0.001	-0.106*
	(0.022)	(0.032)	(0.030)	(0.053)
Reading	-0.000	-0.020	-0.019	-0.053*
	(0.019)	(0.020)	(0.027)	(0.031)
Writing	0.000	-0.037	-0.013	-0.082**
	(0.014)	(0.024)	(0.024)	(0.034)
Matching and Drop Lambton-Kent District School Board				
Math	0.011	-0.031	0.018	-0.047
	(0.024)	(0.031)	(0.034)	(0.053)
Reading	0.007	-0.009	0.006	-0.012
	(0.021)	(0.021)	(0.027)	(0.035)
Writing	0.008	-0.022	0.016	0.043
	(0.013)	(0.025)	(0.016)	(0.047)
Matching Board Level Regressions				
Math	-0.029	-0.025	-0.026	-0.099
	(0.039)	(0.043)	(0.057)	(0.070)
Reading	-0.003	-0.004	-0.035	-0.041
	(0.025)	(0.024)	(0.029)	(0.032)
Writing	-0.013	-0.048**	-0.019	-0.072
	(0.013)	(0.024)	(0.019)	(0.045)

Notes: The reported statistics are the estimated parameter on the strike variables in equation (2). Standard errors in parentheses. "Long strikes" are 10 instructional days or longer. *, **, *** denote statistically significant at the 10, 5, 1 percent level, respectively.

Table 5: Estimates of the Impact of Teacher Strikes in Grade 5 and Grade 6 on the Grade 3 through Grade 6 Change in Student Test Scores.

	All Strikes		Long Strikes	
	Strike in Grade 6	Strike in Grade 5	Strike in Grade 6	Strike in Grade 5
Base Results				
Math	-0.049	-0.008	-0.101***	-0.023
	(0.063)	(0.023)	(0.033)	(0.023)
Reading	-0.023	0.000	-0.048**	0.015
	(0.036)	(0.025)	(0.023)	(0.014)
Writing	-0.031**	-0.017	-0.032**	-0.019*
	(0.014)	(0.011)	(0.016)	(0.010)
Matching Results				
Math	-0.055	-0.031	-0.109**	-0.045**
	(0.075)	(0.023)	(0.045)	(0.021)
Reading	-0.028	-0.011	-0.052*	-0.005
	(0.044)	(0.021)	(0.027)	(0.018)
Writing	-0.035*	-0.017	-0.033	-0.020
	(0.020)	(0.013)	(0.024)	(0.013)
Matching and Drop Simcoe Muskoka Catholic School Board				
Math	-0.007	-0.028	-0.069***	-0.044**
	(0.091)	(0.024)	(0.016)	(0.022)
Reading	-0.001	-0.017	-0.029**	-0.011
	(0.055)	(0.022)	(0.011)	(0.019)
Writing	-0.019	-0.016	-0.011	-0.017
	(0.013)	(0.014)	(0.007)	(0.014)
Matching Board Level Regressions				
Math	-0.026	-0.032	-0.163**	-0.060
	(0.155)	(0.047)	(0.081)	(0.050)
Reading	-0.024	-0.023	-0.081*	-0.027
	(0.082)	(0.031)	(0.047)	(0.032)
Writing	-0.031	-0.006	-0.055	-0.011
	(0.043)	(0.032)	(0.040)	(0.038)

Notes: The reported statistics are the estimated parameter on the strike variables in equation (2). Standard errors in parentheses. “Long strikes” are 10 instructional days or longer. *, **, *** denote statistically significant at the 10, 5, 1 percent level, respectively.

Table 6: Estimates of the Impact of Long Teacher Strikes in Grade 5 and Grade 6 on the Grade 3 through Grade 6 Change in Student Test Scores by Markers of Student Achievement.

	Strike in Grade 6	Strike in Grade 5
	Provincial Pass Rate	
Math	-0.057**	-0.030**
	(0.027)	(0.013)
Reading	-0.023*	-0.002
	(0.012)	(0.012)
Writing	-0.009	-0.011
	(0.009)	(0.008)
	Score=1	
Math	0.026***	0.023***
	(0.009)	(0.007)
Reading	0.012	0.005
	(0.013)	(0.007)
Writing	0.004	0.009**
	(0.011)	(0.004)
	Score=2	
Math	0.031	0.007
	(0.019)	(0.011)
Reading	0.009	-0.002
	(0.007)	(0.013)
Writing	0.005	0.002
	(0.007)	(0.005)
	Score=3	
Math	-0.032*	-0.037***
	(0.018)	(0.013)
Reading	-0.006	-0.005
	(0.010)	(0.009)
Writing	0.012	-0.011
	(0.007)	(0.011)
	Score=4	
Math	-0.025**	0.008
	(0.011)	(0.006)
Reading	-0.015***	0.002
	(0.005)	(0.003)
Writing	-0.020***	0.000
	(0.005)	(0.007)

Notes: The reported statistics are the estimated parameter on the strike variables in equation (2). Standard errors in parentheses. “Long strikes” are 10 instructional days or longer. All results based on the matching estimator. *, **, *** denote statistically significant at the 10, 5, 1 percent level, respectively.

Table 7: Estimates of the Impact of Long Teacher Strikes in Grade 5 and Grade 6 on the Grade 3 through Grade 6 Change in Student Test Scores by Gender.

	Strike in Grade 6	Strike in Grade 5	Strike in Grade 6	Strike in Grade 5
	Males		Females	
Math	-0.139***	-0.081***	-0.084*	0.010
	(0.050)	(0.026)	(0.046)	(0.019)
Reading	-0.043	-0.015	-0.048**	0.020
	(0.044)	(0.020)	(0.022)	(0.023)
Writing	-0.033	-0.054***	-0.025	0.004
	(0.027)	(0.019)	(0.027)	(0.013)

Notes: The reported statistics are the estimated parameter on the strike variables in equation (2). Standard errors in parentheses. "Long strikes" are 10 instructional days or longer. All results based on the matching estimator. *, **, *** denote statistically significant at the 10, 5, 1 percent level, respectively.

Table 8: Estimates of the Impact of Teacher Strikes on the Grade 3 through Grade 6 Change in Student Test Scores, Dropping Cohorts with Large Changes in the Proportion of Eligible Students Writing the Test

	All Strikes		Long Strikes	
	Strike in Grade 3	Strike in Grade 2	Strike in Grade 3	Strike in Grade 2
Math	0.021	-0.052	0.014	-0.101*
	(0.022)	(0.032)	(0.022)	(0.057)
Reading	0.005	-0.016	-0.016	-0.040
	(0.018)	(0.017)	(0.026)	(0.029)
Writing	0.001	-0.034	0.020	-0.085**
	(0.014)	(0.025)	(0.024)	(0.037)
	Strike in Grade 6	Strike in Grade 5	Strike in Grade 6	Strike in Grade 5
Math	-0.052	0.000	-0.114**	-0.013
	(0.081)	(0.022)	(0.051)	(0.022)
Reading	-0.018	-0.011	-0.044	-0.005
	(0.047)	(0.020)	(0.030)	(0.030)
Writing	-0.033	-0.009	-0.034	-0.011
	(0.022)	(0.009)	(0.026)	(0.009)

Notes: The reported statistics are the estimated parameter on the strike variable in equation (2). Standard errors in parentheses. All results based on the matching estimator. "Long strikes" are 10 instructional days or longer. Large changes in the proportion of a cohort writing fall outside the 5^{th} and 95^{th} percentiles of the distribution of this variable across cohorts. *, **, *** denote statistically significant at the 10, 5, 1 percent level, respectively.

Table 9: Estimates of the Impact of Teacher strikes on the Grade 3 through Grade 6 Percentage Change in Cohort Enrollment

All Strikes			Long Strikes		
Strike in Grade 6	Strike in Grade 5	Strike in Grade 4	Strike in Grade 6	Strike in Grade 5	Strike in Grade 4
0.021	0.010	0.025	-0.006	0.011	0.030
(0.025)	(0.017)	(0.025)	(0.016)	(0.018)	(0.027)
Strike in Grade 3	Strike in Grade 2	Strike in Grade 1	Strike in Grade 3	Strike in Grade 2	Strike in Grade 1
0.002	-0.002	-0.008	-0.004	-0.037	-0.044**
(0.013)	(0.020)	(0.015)	(0.016)	(0.036)	(0.017)

Notes: The reported statistics are the estimated parameter on the strike variables in equation (2) plus an additional lag. Standard errors in parentheses. All results based on the matching estimator. Enrolment is measured as of the fall of the school year. "Long strikes" are 10 instructional days or longer. *, **, *** denote statistically significant at the 10, 5, 1 percent level, respectively.

Table 10: Estimates of the Impact of Teacher strikes on the Grade 3 through Grade 6 Percentage Change in Schools' Enrollment of Local Students

	All Strikes			Long Strikes		
	Strike in Grade 6	Strike in Grade 5	Strike in Grade 4	Strike in Grade 6	Strike in Grade 5	Strike in Grade 4
Proportion of Students in English Language Schools						
	0.001	0.001	0.000	-0.000	0.000	-0.000
	(0.002)	(0.001)	(0.001)	(0.001)	(0.000)	(0.000)
Proportion of Students in English Language Public Schools						
Public Strikes	NA	-0.007*	-0.005	NA	-0.007*	-0.005
		(0.004)	(0.005)		(0.004)	(0.005)
Catholic Strikes	-0.003	-0.002	-0.000	-0.003	-0.004	-0.003
	(0.003)	(0.003)	(0.003)	(0.003)	(0.002)	(0.002)
	Strike in Grade 3	Strike in Grade 2	Strike in Grade 1	Strike in Grade 3	Strike in Grade 2	Strike in Grade 1
Proportion of Students in English Language Schools						
	-0.000	0.000	0.000	-0.000	0.000	-0.000
	(0.001)	(0.000)	(0.000)	(0.010)	(0.000)	(0.001)
Proportion of Students in English Language Public Schools						
Public Strikes	0.001	0.000	0.001	0.004	0.000	0.002
	(0.003)	(0.004)	(0.003)	(0.006)	(0.010)	(0.009)
Catholic Strikes	0.001	-0.006**	-0.005**	0.001	-0.005**	-0.005**
	(0.002)	(0.003)	(0.001)	(0.002)	(0.002)	(0.002)

Notes: The reported statistics are the estimated parameter on the strike variable in equation (2). Standard errors in parentheses. NA denotes that the indicated parameter is not identified. All results based on the matching estimator. Enrolment is measured as of the fall of the school year. "Local students" reside with a school's Forward Sortation Area (FSA). "Long strikes" are 10 instructional days or longer. *, **, *** denote statistically significant at the 10, 5, 1 percent level, respectively.

Table 11: A Comparison of Students who Remain in the Same School Board Between Grade 3 and 6 to the full Sample of Students based on Matched Data for 2004 and 2007.

		Grade 3		Grade 6	
		Full Sample	Same Board	Full Sample	Same Board
		102114	87837	104937	86891
		(100)	(0.86)	(100)	(0.83)
Math Test	Score	2.77	2.78	2.70	2.72
	% missing	0.053	0.047	0.045	0.032
Reading Test	Score	2.64	2.65	2.72	2.74
	% missing	0.081	0.074	0.047	0.032
Writing Test	Score	2.71	2.72	2.79	2.80
	% missing	0.070	0.063	0.042	0.028

	Full Sample	Same Board
Change in Math Score	-0.067	-0.063
Change in Reading Score	0.085	0.088
Change in Writing Score	0.074	0.078

Notes: Percentages in parentheses. The Full Sample includes all students writing the indicated test while the Same Board Sample includes students who were in the same school board for the grade 3 and grade 6 tests. % missing is the proportion of students in the cohort who have no score for the indicated test. Test score changes are between grades 3 and 6.

Table 12: Cross Section Estimates of the Impact of Teacher Strikes on Student Test Scores

	All Strikes		Long Strikes	
Grade 3 Test Scores	Strike in Grade 3	Strike in Grade 2	Strike in Grade 3	Strike in Grade 2
Math	-0.004	0.031	-0.006	0.053
	(0.035)	(0.053)	(0.024)	(0.044)
Reading	0.013	0.013	0.024	0.033
	(0.025)	(0.039)	(0.017)	(0.034)
Writing	0.008	0.022	0.019	0.041
	(0.026)	(0.032)	(0.033)	(0.038)
Grade 6 Test Scores	Strike in Grade 6	Strike in Grade 5	Strike in Grade 6	Strike in Grade 5
Math	0.016	-0.037	-0.029	-0.040
	(0.048)	(0.035)	(0.033)	(0.035)
Reading	-0.005	-0.020	0.000	-0.012
	(0.030)	(0.024)	(0.034)	(0.026)
Writing	-0.017	-0.018	-0.012	-0.011
	(0.043)	(0.038)	(0.046)	(0.042)

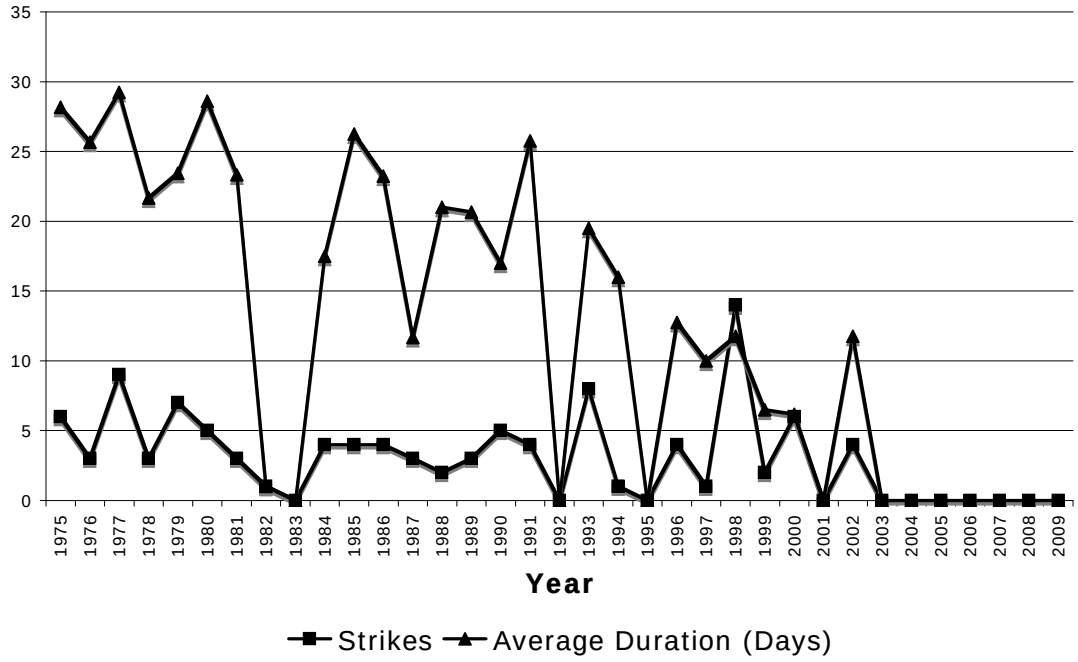
Notes: The reported statistics are the estimated parameter on the indicated strike variable on grade 3 test results (upper panel) or grade 6 test results (lower panel). Standard errors in parentheses. All results from the cohort sample using the matching estimator. "Long strikes" are 10 instructional days or longer. *, **, *** denote statistically significant at the 10, 5, 1 percent level, respectively.

Table 13: Cross Section Estimates of the Impact of Teacher Strikes on Student Test Scores (Extended Sample)

	All Strikes		Long Strikes	
Grade 3 Test Scores	Strike in Grade 3	Strike in Grade 2	Strike in Grade 3	Strike in Grade 2
Math	0.005	-0.021	-0.006	-0.038
	(0.036)	(0.038)	(0.022)	(0.032)
Reading	-0.005	-0.019	-0.021	-0.043
	(0.027)	(0.028)	(0.016)	(0.025)
Writing	-0.003	-0.017	-0.013	-0.030
	(0.027)	(0.027)	(0.033)	(0.035)
Grade 6 Test Scores	Strike in Grade 6	Strike in Grade 5	Strike in Grade 6	Strike in Grade 5
Math	-0.018	-0.001	-0.032	-0.029
	(0.044)	(0.043)	(0.031)	(0.035)
Reading	-0.003	0.006	-0.002	-0.009
	(0.027)	(0.032)	(0.022)	(0.028)
Writing	-0.009	0.016	-0.008	-0.005
	(0.031)	(0.035)	(0.039)	(0.037)

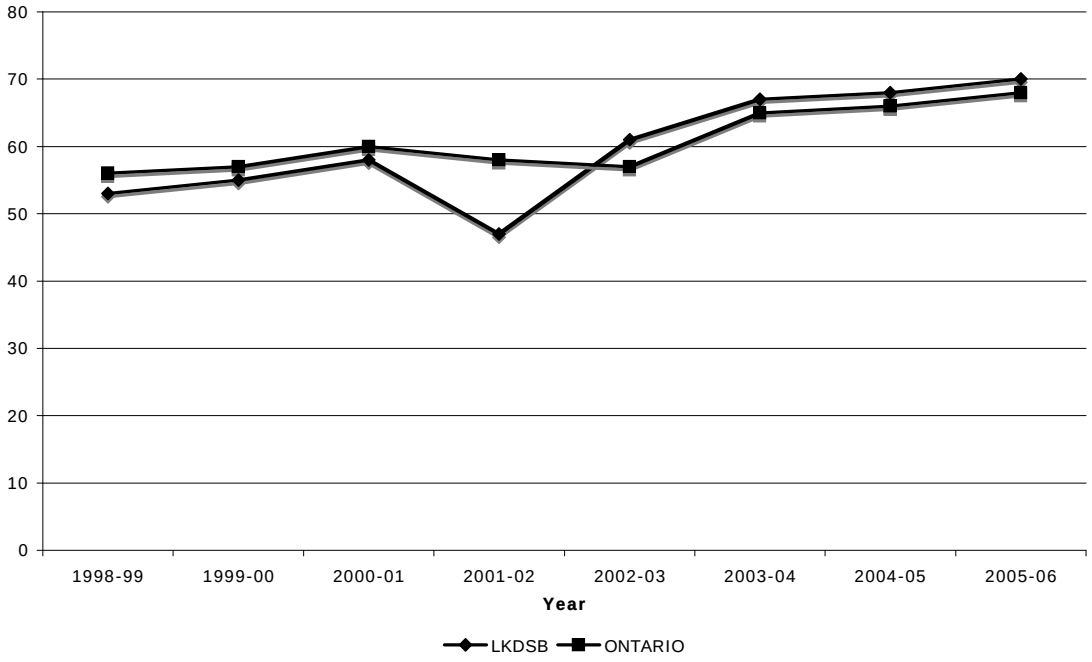
Notes: The reported statistics are the estimated parameter on the indicated strike variable on grade 3 test results (upper panel) or grade 6 test results (lower panel). Standard errors in parentheses. All results from the cohort sample using the matching estimator. "Long strikes" are 10 instructional days or longer. Extended sample contains all years (1998-2005) of school cohorts and teacher strikes. *, **, *** denote statistically significant at the 10, 5, 1 percent level, respectively.

Figure 1: The Incidence and Duration of Teacher Strikes in Ontario's Public Schools



Source: Author's calculation from data sources documented in the text.

Figure 2: Proportion of Students Achieving the Provincial Standard in the Grade 3 EQAO Math Test: Lambton Kent DSB and Province Wide



Notes: Source is data from the Lambton Kent DSB web site http://www.lkdsb.net/.