



IZA DP No. 2923

Unscheduled School Closings and Student Performance

Dave E. Marcotte
Steven W. Hemelt

July 2007

Unscheduled School Closings and Student Performance

Dave E. Marcotte

*University of Maryland Baltimore County
and IZA*

Steven W. Hemelt

University of Maryland Baltimore County

Discussion Paper No. 2923
July 2007

IZA

P.O. Box 7240
53072 Bonn
Germany

Phone: +49-228-3894-0
Fax: +49-228-3894-180
E-mail: iza@iza.org

Any opinions expressed here are those of the author(s) and not those of the institute. Research disseminated by IZA may include views on policy, but the institute itself takes no institutional policy positions.

The Institute for the Study of Labor (IZA) in Bonn is a local and virtual international research center and a place of communication between science, politics and business. IZA is an independent nonprofit company supported by Deutsche Post World Net. The center is associated with the University of Bonn and offers a stimulating research environment through its research networks, research support, and visitors and doctoral programs. IZA engages in (i) original and internationally competitive research in all fields of labor economics, (ii) development of policy concepts, and (iii) dissemination of research results and concepts to the interested public.

IZA Discussion Papers often represent preliminary work and are circulated to encourage discussion. Citation of such a paper should account for its provisional character. A revised version may be available directly from the author.

ABSTRACT

Unscheduled School Closings and Student Performance^{*}

Do students perform better on statewide assessments in years in which they have more school days to prepare? We explore this question using data on math and reading assessments taken by students in the 3rd, 5th and 8th grades since 1994 in Maryland. Our identification strategy is rooted in the fact that tests are administered on the same day(s) statewide in late winter or early spring, and any unscheduled closings due to snow reduce instruction time, and are not made up until after the exams are over. We estimate that in academic years with an average number of unscheduled closures (5), the number of 3rd graders performing satisfactorily on state reading and math assessments within a school is nearly 3 percent lower than in years with no school closings. The impacts of closure are smaller for students in 5th and 8th grade. Combining our estimates with actual patterns of unscheduled closings in the last 3 years, we find that more than half of schools failing to make adequate yearly progress (AYP) in 3rd grade math or reading, required under No Child Left Behind, would have met AYP if schools had been open on all scheduled days.

JEL Classification: I2, I21

Keywords: education, accountability, testing, school resources

Corresponding author:

Dave E. Marcotte
Department of Public Policy
University of Maryland Baltimore County
1000 Hilltop Circle
Baltimore, MD 21250
USA
E-mail: marcotte@umbc.edu

* This research was supported by grants from the Spencer Foundation and the Smith Richardson Foundation. Thanks to Charlie Clotfelter, David Figlio, Raegen Miller, and Steve Pischke for helpful comments and suggestions. Of course, any errors and all opinions are our own.

The amount of research on the relationship between various aspects of schooling and student performance is vast. This includes work by economists on the impact of teachers and other inputs, and work on the effects of curricula or teaching methods. While questions about what happens inside of schools and classrooms and how these affect student learning and performance are surely important, almost no work has addressed the fundamental question: What is the impact of having no school at all? Each winter, administrators regularly are forced to cancel school days because of bad weather. Each day a school is subject to an unscheduled closure, teachers, curricula, and school resources, no matter how effective, can have no real impact on student learning.

In this paper, we examine the impact of school closures on student performance using data from Maryland public schools. We begin by describing the context within which unscheduled school closure decisions occur, and how these might affect student performance. We then describe relevant research by economists, and our empirical approach. Next we discuss our results, and finally consider their implications.

Background

Annually across the country students begin a school year that includes high-stakes tests in selected grades. The results of these tests are part of state efforts to improve accountability in public schools. They have been used to provide information to parents, to pressure administrators, and in some cases trigger re-constitution of individual schools. More recently, they are used to track adequate yearly progress for the federal No Child Left Behind Act.

In Maryland, students in the 3rd, 5th and 8th grades take math and reading assessments, and have done so since 1994.¹ Initially, the testing regime was called the Maryland State Performance Assessment Program (MSPAP), which was then replaced in 2003 by the Maryland

¹ The MSPAPs were first administered in 1993, but the MSDE has not released results for all grades in that year.

State Assessments (MSAs).² So that teachers and principals can accommodate the tests, the test dates are set going into the school year. For security reasons, the tests are administered on the same day(s) statewide. So, all over the state, teachers and administrators plan curricula and instruction in preparation for the exam date.

One important variable for which it is harder to plan is the number of unscheduled school closings before the test date arrives. Each year districts schedule days over and above the 180 day minimum mandated by the U.S. Department of Education, so that schools can be cancelled in the event of bad weather while still meeting the minimum. In the event of severe weather, if the number of days cancelled exceeds the excess number of days scheduled, the school year is extended. Or, if the number of days cancelled is less than the cushion, schools can be dismissed before the scheduled closing. Regardless of whether the year is extended or shortened, this occurs at the end of the year - after the MSAs are administered in March or the MSPAPs were administered in April.

The empirical question we explore in this paper is straightforward: Do students perform better on statewide assessments in years in which they have more school days to prepare for the tests? For obvious reasons, more days in school ought to help students better prepare for state assessments. Of course, it is possible that schools and teachers could alter curricula or forego activities less useful in preparing for assessments if days are unexpectedly lost to closure. But teachers may not be able to completely make up for days lost to closure. Indeed, the very notion of a mandatory 180 day school year rests on a premise that a certain amount of time is necessary for teachers to cover and students to comprehend material.

² Since 2002, students in 4th, 6th and 7th grades are also tested as part of the MSA. However, because the MSPAP tested only students in 3rd, 5th and 8th grades, we restrict our analysis to the smaller, common set.

Work by economists on the relationship between schooling inputs and student performance has focused largely on the impact of higher quality inputs – not on marginal changes in input quantity. This includes a substantial amount of work on changes in class size, and on increased expenditures on education.³ Little empirical work has been done on the impact of more or less school.

Recent work on teacher absences provides some insight, since days when teachers are absent provide one less day of exposure to the treatment as intended. Clotfelter, Ladd and Vigdor (2006) and Miller, Murnane and Willett (2006) have examined the impact of teacher absences on student performance. Both studies find evidence that students learn less when teachers are absent. Of course, when teachers are absent, students are supervised and likely even taught by substitute teachers. So, characterizing this as a quantity change is questionable.

A different piece of evidence comes from Pischke (2003), who examines the impact of shortened school years in Germany. He exploits the fact that in the mid-1960s, West German states switched from the practice of beginning school years in January, to beginning in the autumn, with the exception of Bavaria, which already started in September. To accommodate this switch, the school year beginning in January was abbreviated, so that the next year could begin in September. Pischke finds evidence that students fared more poorly immediately following abbreviated school years, but that there were no persistent, longer-term effects on schooling or on labor market outcomes.

Both Card and Krueger (1992) and Grogger (1996) examine the relationship between length of school year across states within the U.S. and subsequent labor market earnings. This approach is much less direct, and relies on fairly small levels of variation in length of the

³ See Card and Krueger (1992, 1996) and Hanushek (2002) for summaries.

academic year. Indeed, during the past several decades there has been no real variation in term length within the U.S., as districts widely adhere to a 180 day calendar.⁴

Eren and Millimet (2007) examine the effect a length of school year on student performance using data from the National Education Longitudinal Study: 1988. They use a dichotomous measure of length of school year (180 days or fewer versus more than 180 days) and find that high performing students do better with longer school years, while low performing students fare worse.

Marcotte (2007) estimates the impact of snowfall on student performance in Maryland. He finds that students who took exams in years with heavy snowfall performed significantly worse on the MSPAP assessments than did their peers in the same school who took the exams in other years. However, this work focuses solely on the reduced form relationship between snow and performance, not on the impact of the central policy variable, days of instruction.

Data and Methods

In order to examine the relationship between instructional time and performance, we have constructed a panel from school level data provided by the Maryland State Department of Education (MSDE). In 1993, the State of Maryland implemented a standardized testing program (MSPAP) for students in grades 3, 5 and 8, with results for all grades publicly available beginning in 1994. In 2003, the MSPAP was replaced by the MSA. The most recent MSA data available is 2005. Together, the MSA and MSPAP provide measures of students' performance in the areas of mathematics and reading for the years 1994 to 2005.⁵ In the models below, we use

⁴ This may be changing, as some districts and some schools are extending the school year to 200 days.

⁵ The MSPAP also included subject tests in science, social studies, writing and language usage, while the MSA includes only reading and math assessments.

measures of the percentage of students in a school who met MSDE guidelines for satisfactory performance on the MSPAP or MSA reading and math assessments.

The MSDE data also provide information about average student characteristics in the school and resources in the district. The measures of student characteristics include: the number of students with English as a second language (ESL) barriers; the number of students with special education needs; the number of students receiving free or reduced-price lunches, and; the number of students who are Title I eligible. We include these in the models estimated below to control for differences in student characteristics across schools that may be expected to affect performance.

To the panel constructed from MSDE data, we merge in data on the number of unscheduled closing days in an academic year, collected directly from school districts. Not all districts were able to provide data on closures for each year back to the 1993-1994 school year, so the panel is a bit unbalanced with districts entering the estimation panel in the first year these data are available. On average, we have just more than 7 years of data for each district.

We have data on total snowfall for all academic years leading up to the test dates. Data on snow accumulation are provided by the National Oceanic and Atmospheric Administration's National Climatic Data Center (NCDC). From NCDC, we obtain the accumulation recorded at the principal weather-reporting station within each county during a winter. In Maryland each county constitutes its own local education authority (district), as does Baltimore City.

Using this panel, we estimate models of the relationship between unscheduled closures in a year and schools' performance on the math and reading tests for students. Separately, for each grade the basic set up is:

$$P_{ost} = \alpha + \beta_1 X_{st} + \beta_2 C_{st} + \beta_3 t + \alpha_s + \alpha_d + (t^* \alpha_d) + \beta_4 MSA_t + \varepsilon_{ost}$$

Where, P_{ost} is the percentage of students in the relevant grade at school s at time t who perform at least satisfactorily on subject test o (math or reading). MSDE did not provide scores for individual students or the means for students within schools/grades. Nonetheless, the percentage of students performing satisfactorily is the fundamental policy variable of interest, since it is the yardstick with which performance under NCLB is measured.

X_{st} is a vector of resources and student characteristics for school s at time t known to affect achievement, including: the percentage of children receiving free or reduced price meals, the percentage of students in the school who are Title I, the percentage of English as a Second Language (ESL) students, and the percentage of students in special education.

C_{st} is a measure of the number of scheduled school days that school s was closed in year t . We include unscheduled closing days directly, and in a specification to check for possible non-linear effects. We also estimate models in which we examine whether unscheduled closing days have different effects at schools with more poor students (as defined by the percentage of students receiving free/reduced price meals).

We include school fixed-effects so that the variation of interest is within-school differences in test performance. We also include district-specific linear trends to control for changes in unmeasured student characteristics within districts. The empirical question of interest is then: Do students in a given school perform better, on average, than their peers at the same school who, in different academic years, received fewer days of instruction prior to the state assessments?

We also include dummy variables measuring year effects for the two years prior to the change in the test regime to the current MSA assessments. It was widely known that the MSPAP

tests were being phased out, and that test results in these final two years were less consequential. We do not include fixed year effects for all years. More than 70% of the variation in school closings during the panel is across time, and is common across districts.⁶ So, year effects would absorb most of the variation of interest.

All models are estimated using weighted-least-squares, to adjust for heteroskedasticity in the error term, using each school's enrollment in the relevant grade as a weight. We account for the possibility of serial correlation in the error term by using Huber-White robust standard errors, clustering on school district, since superintendents make decisions about closure for all schools in the district.

We estimate separate models for performance on the reading and math assessments, in turn for students in 3rd grade, 5th grade and 8th grade. If instruction time improves test performance we expect that the proportion of students performing at a level the MSDE considers satisfactory declines with the number of unscheduled closings in a given year.

Results

Descriptives:

We begin by considering descriptive statistics on test performance, unscheduled closing days, district resources and student characteristics for our sample, presented in Table 1. Over the course of the panel, between about a third and one-half of students performed satisfactorily on reading and math tests in various grades. For example, on average, districts saw 43.5 percent of third graders perform satisfactorily on reading assessments, and 45.4 percent perform satisfactorily on math assessments. However, the range for both of these scores is large. The percentage of students whose scores were satisfactory was comparable among 5th grade students, but performance was relatively poor among students in the 8th grade, with 37.6 percent and 43.5

⁶ The R^2 from a regression of total closing days on year fixed effects only is .711

percent performing satisfactorily on the reading and math assessments, respectively. On average, districts reported that about 42 percent of 3rd grade students were eligible for free or reduced-price meals, less than 5 percent were ESL students, and about 12 percent were in special education programs.

To illustrate the potential impact of unscheduled closures, in Figure 1 we report variation in the number of unscheduled closing days over the period, by district. The diamond within each box represents the median number of unscheduled closures over the period within a district and the top and bottom of each box the 75th and 25th percentiles, respectively. Of importance, the minimum value for many districts was zero. That is during at least one school year in the period there were no unscheduled closings. Nonetheless, even in such districts, there were other years with many unscheduled closings: For example, Baltimore County also had a year in which 12 days were lost to unscheduled closings. An interesting outlier is the 21 unscheduled closing days in Prince George's County. This was during the Washington area sniper's reign of terror, which included the shooting of a student at a Prince George's middle school. This outlier notwithstanding, over the course of the panel, many districts had some years in which there were no unscheduled closures and some years when more than two weeks of instruction time was lost to snow.

As a way to begin understanding the relationship between unscheduled school closures and performance on state assessments, in Figure 2 we present time series of the average number of school closures in the state, and de-trended average performance on 3rd grade reading and math tests, net of school fixed effects. The solid line is the time series of a student-weighted average number of unscheduled closing days across the state of Maryland between 1994 and 2005. The dashed and dotted lines report the average number of 3rd grade students passing state

assessments in math and reading (respectively) net of trends and school fixed effects. So, in a year when students in a school perform better on math/reading tests than average over the panel the series is positive, in years below average, it is negative.

The series are interesting and suggestive: Both math and reading scores vary inversely with unscheduled school closings. Years in which students score above their school means were years with relatively few unscheduled closures, and years of below mean performance were years with frequent closures.

To more fully consider the relationship between days lost to unscheduled school closures and the percent of students performing satisfactorily on the reading and math assessments, we present results on key variables from our main models in Table 2. In the first panel we present results for performance on the reading and math assessments for 3rd grade students. In the second and third panels, we present results for 5th and 8th grade, respectively. These models include controls for student characteristics, school fixed effects, and district specific time trends. In the first column in panel 1, we find each day lost to unscheduled closure reduces the percent of 3rd grade children in an elementary school that pass the reading exam by 0.508. The pass rates on the math exam falls by 0.527 percent with each day 3rd graders lose to unscheduled closing. Since on average there are a bit more than 5 unscheduled closings each winter, nearly 3 percent fewer third graders pass reading and math tests in a typical winter than would have if there were no unscheduled closings at all. In winters with 10 unscheduled closings, more than 5 percent fewer 3rd graders would pass reading and math assessments.

Of course, it may be that each day lost to unscheduled closure may not have the same impact on performance. Losing one or two days over the course of a winter may not have much impact on student performance, and many teachers may even find it restorative. But, losing

many days may be more seriously disruptive perhaps because lessons are forgotten and time needs to be spent re-teaching material covered before a long unplanned intermission. To test this, we include a series of indicators for the total number of days lost in a school year, ranging from 1-2 days to more than 12. These results, presented in columns 2 and 5, do suggest that test scores fall as unscheduled closures mount. There is some evidence that the impact of closures is non-linear, but not substantially. For reading, a bit more than 0.40% fewer students will perform satisfactorily for each day lost in winters with 11-12 closures, while this daily loss rises higher in the heaviest winters. For math, in years with 3 to 5 days of closure, about 0.25% fewer students will pass state assessments for each day lost – while in years with 8-10 closings, about 0.33% fewer students will pass for each day lost. This rises to a bit more than 0.50% per day in winters with more than 12 unscheduled closings. These patterns suggest that long lay-offs may be especially disruptive to learning or teaching mathematics.

In the next columns, we allow the impact of unscheduled closings on performance to vary with the percent of students eligible for free/reduced price meals. We distinguish between schools where the percent of students eligible was less than 15%; between 15 and 35%; between 35 and 67%; and more than 67 percent. These are approximate cut-points for statewide unconditional quartiles.

We estimate that days lost to unscheduled closure decrease reading pass rates by about one-third of a percent (.369) at schools with the least poor students and by more than three-fifths of a percent (.614) at schools with the most poor students. These patterns are consistent with the interpretation that lost instruction time has a larger negative impact on reading performance for the marginal student in schools with more poor students. This pattern does not hold for

performance on the math assessments. Indeed, we estimate that the impact of days lost on the percent of students passing math tests is at least as large at wealthier schools as poor schools.

In the next panels, we present results for 5th and 8th grades. In both cases, unscheduled school closings negatively affect performance on math and reading assessments, but the impact is smaller than observed for 3rd graders. For 5th grade students, the impact on math performance is smaller, though comparable to the 3rd grade results, but the effect on reading performance is quite a bit smaller. For 8th graders, the unscheduled school closings have similar effects on the percentage of students passing the reading and math tests, with the effect about half as large as that observed among 3rd graders.

That the largest effects are observed for students in lower grades is not surprising. Since districts can and do make up for snow days at the end of all school years, unscheduled closings during the year result in a relatively larger loss in *cumulative* instruction at the time of the assessments for students in lower grades. That is, going into the third grade, students had already received at least 360 instructional days in 1st and 2nd grades. Going into the 8th grade, students have received a minimum of 1,260 instructional days. So, any closures in the testing year would result in a larger fraction of lost instructional time for students in 3rd grade.

Specification Checks:

The results in Table 2 provide evidence that days lost to unscheduled closure have negative and non-trivial consequences for student performance on state assessments. Whether we can interpret these effects as causal depends on whether unscheduled closing days provide exogenous information on variation in school instruction and resources. The decision to close schools in an LEA is made by the district superintendent. Naturally, superintendents are also aware of past and expected performance on state assessments within their districts. One concern

is that in districts where schools are underperforming, superintendents may be less willing to close school in the face of bad weather than in districts where schools are expected to be safely performing above state standards.

Of course, the weather has no regard for superintendents' worries about test scores, and it seems likely that the decisions about last minute closures are highly constrained. One way to shed light on whether closing decisions are the product of natural fluctuations in weather, or something else, is to examine variation in school closings and snow within districts. In Figure 3, we plot average snowfall and closing days for the state (in the 1st panel), and then in the three other panels, we present closings and snowfall for a district in the western part of the state (Garrett), a central district (Baltimore County), and one on the Eastern Shore (Talbot). Each panel presents the snowfall (in inches) on the primary y-axis, and the number of unscheduled closing days in an LEA on the secondary y-axis. Both on average, and within counties, unscheduled school closings substantially track snowfall. In each case, closing days track above snowfall in some years and below in others. This pattern is to be expected due to variation in the timing of snow⁷ or in how it is distributed across storms,⁸ but it does not appear consistent with a systematic over- or under-use of unscheduled closings. A better way to assess the possibility that unscheduled closures are endogenous is to make use of data on snowfall as an instrument for unscheduled closings. In Table 3 we present a summary of the results of 2SLS estimates of the impact of total closings on the percentage of 3rd, 5th, and 8th graders passing state reading and math assessments. These results are from models that include the same specification used in the

⁷ Snow on a school day would have a different effect than during a holiday.

⁸ Winters with several average storms may have a larger effect than winters with one large storm – even if snow totals are similar. For example, in 1996, much of the state faced a storm that dumped between 20 and 30 inches of snow in one day.

odd numbered columns in Table 2. Here, however, we use total snow accumulation during the winter in a district to instrument for the number of unscheduled closing days.⁹

As expected, snow substantially predicts the number of unscheduled closings in a school year. For each grade, the coefficient on snowfall in the first stage model to predict the number of closings is about 0.125: On average, schools close a day for each 8 inches of additional snow accumulation during the winter. The smaller t-statistic for the 8th grade is the result of a larger standard error, since there are fewer middle schools, not the result of a smaller first stage point estimate.

To test whether the number of closings in a district is endogenous, we include the residual from the first stage model of total closings in models of reading and math scores. This sets up a test of endogeneity as a test of significance on this coefficient (Hausman, 1978). Any evidence that unmeasured factors associated with a superintendent's propensity to use unscheduled closing days are positively related to test scores is modest, at best. In four of six cases (and then only at the 10% level of significance) is there some evidence consistent with the concern described above, that superintendents are more likely to close school for the day if they are less worried about poor test performance.

In the last two columns of Table 3, we present the 2SLS estimates of the effects of unscheduled closings on reading and math scores. While these estimates are on the order of twice as large (in absolute value) as those reported in Table 2, the same patterns emerge: Unscheduled school closings appear to negatively affect test performance, especially in lower grades, and more so for mathematics.

⁹ There is some chance that snowfall, though purely a natural event, may not satisfy the exclusion restriction. If snowfall affects kids' ability to focus on work while in school, or spend time on homework outside of school we may be attributing some of the direct effect of snow to the policy variable, school closure. Unfortunately, we have no way of testing this.

A second way to check the specification is to examine whether exam performance suffers more if school is interrupted closer to the exam date – and not at all if it is interrupted after the exam date. Unfortunately, data on the timing of school closures are available only for an unrepresentative subset of LEAs. Only 13 of the 24 LEAs were able to provide at least some historical data on the dates of unscheduled closings. Worse, these were largely rural districts on the Eastern Shore and in western and southern Maryland. Only two populous counties in central Maryland provided these data.¹⁰ However, while we do not have data on the timing of unscheduled closures themselves in all LEAs, we do have data on snowfall by month for all LEAs.¹¹

We use the incomplete information on the timing of unscheduled closures, along with the complete information on the timing of snowfall to carry out two separate checks of the relationship between timing of interruptions and test performance. First, we estimate models for the limited sample that provided information on the timing of school closures to identify timing effects. Second, for the full sample we estimate reduced form models of the impact of snowfall on test performance.

In both cases, we estimate the impact of the interruption (closing/snowfall) by month. We include separate measures of snow late in the season (March and April) for MSA versus MSPAP years. The MSA has been administered in the first week of March. Closures and snow after the test date should have no effect on MSA performance. But, closures and snow in March and April should have an effect on performance on the MSPAP, which was administered in late April.

¹⁰ The LEAs providing data on the timing of closure are: Baltimore County, Caroline, Charles, Dorchester, Frederick, Garrett, Howard, Kent, Queen Anne's, St. Mary's, Talbot, Washington, and Worcester.

¹¹ We were not able to acquire information on the day of the week on which snowfall occurred, though this would obviously be helpful for identifying effects on school closures.

In Table 4, we present results on the effects of unscheduled closings, by month, for the sub-sample providing these data. By way of assessing the characteristics of this sub-sample, we first present the effects of unscheduled closings on test performance, regardless of the timing. This specification is identical to the first columns in each subject/grade in Table 2. We then differentiate between unscheduled closures in the fall (in or before November), December, January, February, and then March and April during MSA years separate from March and April during MSPAP years.

It is clear from the patterns across grades that unscheduled closures in the early part of the academic year do not have a consistent, significant impact on test scores. Closures in the fall, December and January have somewhat mixed effects on performance. However, unscheduled closures as the exams loom have a negative and significant effect on test performance. In February, unscheduled closures have a consistent, substantial and significant negative effect on performance. Closings in March and April are also negatively related to test performance in MSPAP years, when the exams were given in mid to late April. However, closings late in the year have no negative impact on performance in MSA years, when the exams were administered earlier.¹² For this sub-sample, at least, the relationship between timing of closures and test scores is consistent with a causal link between unscheduled closures and performance.

In Table 5, we present results from the reduced form approach for the whole sample. Snow early in the season has a negative effect only in the case of performance on 3rd grade reading exams. Snow in the December, but not in January, has a consistently negative effect on pass rates across subjects and grades. As was the case for the models of unscheduled closure, it is disruption in February that appears to have the most consistent, negative impact on test

¹² This should be interpreted with some caution: Only one district (Frederick) reported a closing this late in the year during the three years of MSA testing.

performance. For both subject tests and for all grades, snow in February is substantially and negatively related to test performance. Also like the results in Table 4, snow after March is generally negatively associated with MSPAP scores, but not with MSA scores. As with the model of timing on test performance on the limited sample, this reduced form evidence from the full sample is largely consistent with the interpretation that interruptions to class schedules negatively affect test performance.

Conclusions and Implications:

The experiences of school districts in Maryland provide evidence that losing school days to unscheduled closings has negative effects on performance on state assessments. We find that school closings have larger effects on performance for students in lower grades. We estimate that the pass rate for 3rd grade math and reading assessments will fall by more than a half percent for each school day lost to an unscheduled closure. This means that in years with a high level of unscheduled closure (10 days is a common value in heavy snow winters), more than 5 percent fewer students will pass 3rd grade reading and math tests than in winters with no unscheduled closures.

That unscheduled closings affect the percent of students meeting state standards raises troubling questions about whether such schools can make adequate yearly progress (AYP) in meeting mandates for 100 percent proficiency by 2013-14 under No Child Left Behind. Between 2003 and 2006, Maryland AYP was set so that the percentage of students performing at least satisfactorily (Annual Measurable Objectives, or AMOs) increased at an average rate of 5% for reading and 6% for math assessments. For schools at the margin of meeting AMO for a year, unscheduled closures could impose a substantial handicap. Our estimates imply that in academic years with an average number of unscheduled closures, about 3 percent *fewer* 3rd graders will

perform satisfactorily on these tests. It is easy to see that in a bad winter, a real increase in performance necessary to make AMO could be swamped by the consequences of school closures.

In Table 6 we illustrate the potential importance of unscheduled closures on the likelihood schools fail to make AYP. We present the actual number of elementary schools in the state failing to make AYP in 3rd grade reading and math tests for 2003-2005.¹³ We then estimate the number of these schools where performance was sufficiently close to the AYP standard that if there had been no unscheduled closings, the school would have met the standard, assuming average increments to pass rates for each additional day of instruction. We estimate that in 2003, 30 of the 52 failing elementary schools would have surpassed the AYP threshold if there had been no unscheduled closures. Similarly we estimate that 34 of the 56 schools failing to meet AYP in math would have exceeded the standard had it not been for unscheduled closings. Only in the final year, with relatively few unscheduled closings, does the number of failing schools far enough below their AMO that they would not have met AYP even if school had been in session for all scheduled days exceed the number of schools that would have met AYP with the extra days.

We find some evidence that superintendents may be aware of the marginal impact of a lost school day, and that those with lower average test scores are somewhat less likely to make use of an unscheduled closing when faced with the same weather conditions. This under-utilization of unscheduled closings in poor performing districts (or excessive use in districts scoring better than expected) biases the impact of closings on performance towards zero. So, the

¹³ AYP was first measured in Maryland in 2003.

causal effect of unscheduled closings on student performance on MSA math and reading assessments are, if anything, larger than the estimates presented here.

A further implication of this work bears on the question of whether the academic year should be extended. At least as far back as the 1983 report to the President, “A Nation at Risk,” extending the school year has received some attention as a mechanism to improve performance. Currently, state or district superintendents in Massachusetts, Indiana, and Minnesota are pushing for longer school years, and public charter schools are experimenting with lengthening school days and school years. For example, the Knowledge is Power Program (KIPP) association of charter schools has extended the length of the school day and year. Our results suggest that additional days of instruction do improve achievement on standardized tests. However, whether lost school days in the middle of the academic year have the same effect on achievement as additional days in late spring or early summer is an open question. Besides the distractions inherent with warm weather, the possibility of diminishing returns to the marginal day beyond 180 makes it difficult to use the current results to forecast achievement gains that might result from adding days to the school calendar. Credible evidence on this question can best be had by evaluating experiences of districts that have made these changes.

References

- Card, David, and Alan Krueger. 1992. "Does School Quality Matter? Returns to Education and the Characteristics of Public Schools in the United States," *Journal of Political Economy*, v. 100, pp. 1-40.
- Card, David, and Alan Krueger. 1996. "School Resources and Student Outcomes: An Overview of the Literature and New Evidence from North and South Carolina," *Journal of Economic Perspectives*, v. 10(4), pp. 31-50.
- Clotfelter, Charles, Helen Ladd, and Jacob Vigdor. 2006. "Teacher absences: Importance, incidence, and consequences. Working Paper. Duke University.
- Eren, Ozkan, and Daniel L. Millimet. 2007. "Time to learn: The organizational structure of schools and student achievement. *Empirical Economics*. 32: 301-322.
- Grogger, Jeff. 1996. "Does School Quality Explain the Recent Black/White Wage Trend?" *Journal of Labor Economics*, v 14, pp. 231-53.
- Hanushek, Eric A. 2002. "Publicly Provided Education," in Auerbach and Feldstein eds. *Handbook of Public Economics*, vol. 4. (Amsterdam: Elsevier Science).
- Hausman, Jerry A. 1978. "Specification Tests in Econometrics," *Econometrica*. v. 46, pp. 1251-71.
- Marcotte, Dave E. 2007. "Schooling and Test Scores: A Mother-Natural Experiment," *Economics of Education Review*. Forthcoming.
- Miller, Raegen, T., Richard J. Murnane, and John T. Willett. 2006. "The Impact of Teacher Absences on Student Achievement," Working Paper. Harvard University, Graduate School of Education.
- Pischke, Jorn-Steffan. 2003. "The Impact of Length of the School Year on Student Performance and Earnings: Evidence from the German Short School Years," (Cambridge, MA: NBER) Working Paper #9964.
- U.S. Department of Education, National Center for Education Statistics (2005). *The Condition of Education 2005*. NCES 2005-094. (Washington DC: U.S. Government Printing Office).

Figure 1

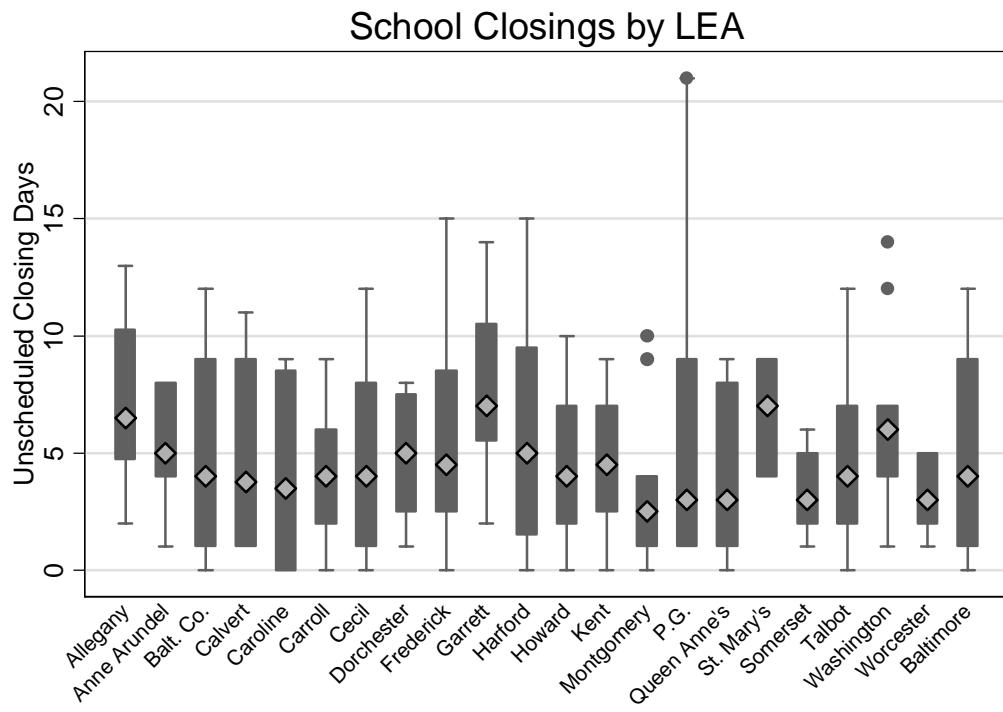


Figure 2

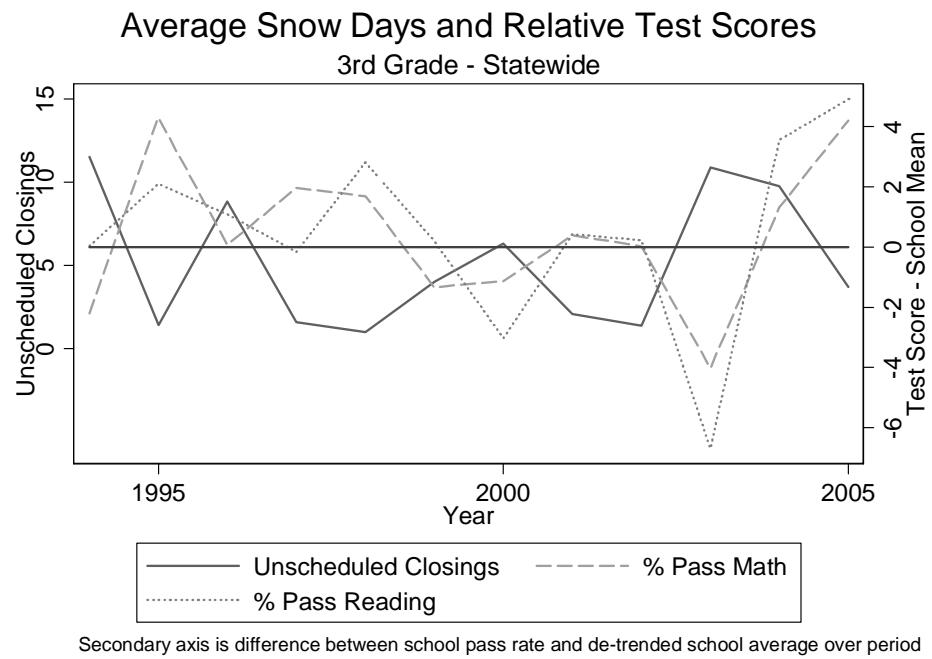


Figure 3

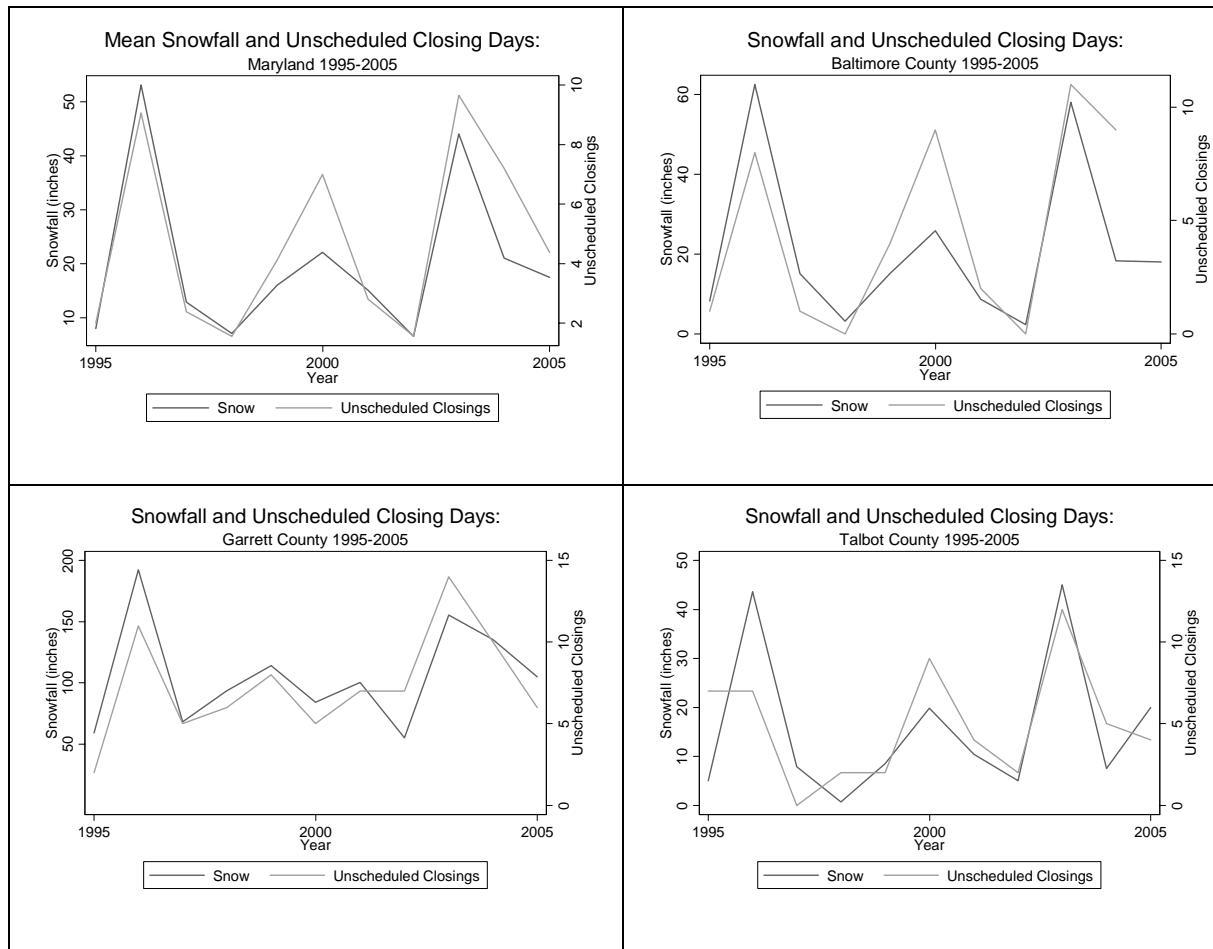


Table 1 Descriptive statistics

Variable	N	Mean	Standard Deviation
3rd Grade			
Percent satisfactory - reading	6173	43.5	23.88
Percent satisfactory -math	6173	45.4	26.10
Title I students (as % of grade)	6173	27.2	41.06
Students receiving free/reduced lunch (as % of grade)	6173	42.1	30.95
ESL students (as % of grade)	6173	4.6	8.59
Special ed. students (as % of grade)	6173	12.6	6.36
Unscheduled Closing Days	6173	5.3	4.67
5th Grade			
Percent satisfactory - reading	6143	46.3	23.60
Percent satisfactory - math	6143	47.9	24.08
Title I students (as % of grade)	6143	26.0	40.34
Students receiving free/reduced lunch (as % of grade)	6143	41.0	31.09
ESL students (as % of grade)	6143	4.4	8.18
Special ed. students (as % of grade)	6143	12.5	6.32
Unscheduled Closing Days	6143	5.3	4.68
8th Grade			
Percent satisfactory - reading	1834	37.6	23.72
Percent satisfactory - math	1834	43.5	23.03
Title I students (as % of grade)	1834	8.5	26.48
Students receiving free/reduced lunch (as % of grade)	1834	34.2	24.96
ESL students (as % of grade)	1834	2.1	3.48
Special ed. students (as % of grade)	1834	13.3	7.62
Unscheduled Closing Days	1834	5.4	4.49

Table 2
**Effects of Unscheduled Closings on
MSPAP/MSA School Test Performance**

3rd Grade						
	Reading			Math		
Unscheduled Closings (Days)	-0.508 (0.126)***			-0.527 (0.055)***		
1-2 Unsch. Closing Days		-1.164 (0.407)***			-0.364 (1.071)	
3-5 Unsch. Closing Days		0.863 (0.798)			-0.882 (0.933)	
6-7 Unsch. Closing Days		-2.635 (1.614)			-1.750 (1.338)	
8-10 Unsch. Closing Days		-1.818 (1.186)			-2.388 (0.658)***	
11-12 Unsch. Closing Days		-4.571 (1.116)***			-6.128 (1.627)**	
>12 Unsch. Closing Days		-11.009 (1.986)***			-9.950 (1.654)***	
Unscheduled Closings (Days) at Schools with:						
- 15% or less FARM students				-0.369 (0.122)***		-0.638 (0.115)***
- 15% to 35% FARM students				-0.441 (0.109)***		-0.498 (0.111)***
- 35% to 67% FARM students				-0.487 (0.104)***		-0.444 (0.077)***
- 67% or more FARM students				-0.614 (0.229)**		-0.559 (0.081)***
R ²	0.8853	0.8866	0.8811	0.881	0.8807	0.8811
n = 6,173						
843 schools						
5th Grade						
	Reading			Math		
Unscheduled Closings (Days)	-0.136 (0.076)*			-0.403 (0.050)***		
1-2 Unsch. Closing Days		-0.472 (0.856)			-0.729 (0.945)	
3-5 Unsch. Closing Days		0.129 (0.838)			-1.037 (1.13)	
6-7 Unsch. Closing Days		-2.103 (1.183)*			-4.073 (1.72)**	
8-10 Unsch. Closing Days		-0.729 (0.954)			-2.726 (1.106)**	
11-12 Unsch. Closing Days		-0.285 (1.271)			-7.444 (2.529)***	
>12 Unsch. Closing Days		-2.841 (1.465)*			-7.065 (1.239)***	
Unscheduled Closings (Days) at Schools with:						
- 15% or less FARM students				-0.145 (0.068)**		-0.467 (0.064)***
- 15% to 35% FARM students				-0.021 (0.059)		-0.379 (0.082)***
- 35% to 67% FARM students				-0.139 (0.075)*		-0.435 (0.06)***
- 67% or more FARM students				-0.200 (0.177)		-0.366 (0.112)***
R ²	0.9017	0.902	0.9019	0.8644	0.8656	0.8644
n = 6,143						
846 schools						
8th Grade						
	Reading			Math		
Unscheduled Closings (Days)	-0.286 (0.091)***			-0.271 (0.066)***		
1-2 Unsch. Closing Days		0.520 (0.709)			0.068 (0.791)	
3-5 Unsch. Closing Days		-0.374 (0.878)			-0.225 (1.089)	
6-7 Unsch. Closing Days		0.386 (1.287)			-2.221 (1.348)	
8-10 Unsch. Closing Days		1.031 (0.872)			-1.492 (0.957)	
11-12 Unsch. Closing Days		-5.473 (1.234)***			-3.605 (1.388)**	
>12 Unsch. Closing Days		-5.766 (1.533)***			-4.410 (0.731)***	
Unscheduled Closings (Days) at Schools with:						
- 15% or less FARM students				0.131 (0.121)		-0.289 (0.16)*
- 15% to 35% FARM students				-0.073 (0.093)		-0.213 (0.075)***
- 35% to 67% FARM students				-0.590 (0.097)***		-0.373 (0.067)***
- 67% or more FARM students				-0.503 (0.081)***		-0.069 (0.201)
R ²	0.946	0.948	0.9493	0.9426	0.9427	0.943
n = 1,834						
289 schools						

Robust standard errors in parentheses

* significant at 10% ** significant at 5%; *** significant at 1%

All models also control for school and student characteristics, district level weather and resources, a linear trend and dummies to measure change in test regime.

Table 3 IV Estimates of Effects of Unscheduled Closings on MSPAP/MSA School Test Performance

Tests of Instrument Quality and Endogeneity			Dep. var: % Passing Subject: Reading Math	
3rd Grade	$H_0: \beta_{SNOW} = 0$	$t = 43.02$	-1.143 (0.327)***	-1.030 (0.288)***
	Endogeneity of Total Closings	$t = 1.75$ (Reading) $t = 1.68$ (Math)		
5th Grade	$H_0: \beta_{SNOW} = 0$	$t = 42.84$	-0.393 (0.078)***	-0.681 (0.240)***
	Endogeneity of Total Closings	$t = 2.80$ (Reading) $t = 0.84$ (Math)		
8th Grade	$H_0: \beta_{SNOW} = 0$	$t = 24.37$	-0.053 (0.131)	-0.668 (0.203)***
	Endogeneity of Total Closings	$t = -1.38$ (Reading) $t = 2.29$ (Math)		

Robust standard errors in parentheses

** significant at 5%; *** significant at 1%

All models also control for school and student characteristics, school fixed effects, a district specific linear trend and dummies to measure change in test regime.

Table 4

**Effects of Unscheduled Closings on Test Performance
in sub-sample of LEAs: by Timing of Closure**

	3rd Grade				5th Grade				8th Grade			
	<i>Reading</i>	<i>Math</i>	<i>Reading</i>	<i>Math</i>	<i>Reading</i>	<i>Math</i>	<i>Reading</i>	<i>Math</i>	<i>Reading</i>	<i>Math</i>	<i>Reading</i>	<i>Math</i>
Unscheduled Closings (Days)	-0.509 (0.066)***	-0.721 (0.093)***	0.0155 (0.058)	-0.55 (0.135)***	-0.145 (0.082)	-0.111 (0.109)						
Closings in Fall (Sept-Nov.)	-1.606 (0.555)**	0.211 (0.769)	-0.878 (0.565)	-0.661 (0.686)	-0.52 (1.115)	-0.953 (0.55)						
Closings in December	-1.266 (0.621)*	-1.159 (0.524)**	-0.626 (0.391)	-1.28 (0.555)**	-0.709 (0.509)	-1.647 (0.674)**						
Closings in January	-0.208 (0.287)	-0.473 (0.107)***	0.363 (0.188)*	0.344 (0.404)	0.823 (0.171)***	0.773 (0.225)***						
Closings in February	-1.176 (0.259)***	-1.322 (0.447)**	-0.026 (0.435)	-2.55 (0.267)***	-2.052 (0.347)***	-1.011 (0.334)**						
Closings in Mar./April in MSPAP years	-0.165 (0.574)	-1.416 (0.473)**	-0.446 (0.193)**	-2.105 (0.693)**	-1.766 (0.35)***	-1.262 (0.518)**						
Closings in Mar./April in MSA years	5.983 (1.945)**	5.249 (1.843)**	2.67 (1.138)**	1.776 (1.644)	2.135 (1.296)	6.023 (2.767)*						
R ²	0.8482	0.8554	0.8575	0.8562	0.8587	0.8628	0.8089	0.8106	0.9169	0.9256	0.8944	0.9002
	n = 2,638 277 schools				n = 2,505 274 schools				n = 866 103 schools			

Robust standard errors in parentheses

* significant at 10% ** significant at 5%; *** significant at 1% (all two-tailed tests)

All models also control for school and student characteristics, district level wealth and resources, a linear trend and dummies to measure change in test regime.

Table 5

Reduced Form Estimates of Snow on Test Performance By Timing of Snow

Snowfall (in inches)	3rd Grade		5th Grade		8th Grade	
	Reading	Math	Reading	Math	Reading	Math
Snow in Fall (Sept-Nov.)	-0.161 (.029)***	-0.101 (.043)**	-0.029 (.021)	-0.007 (.036)	-0.041 (.024)	0.021 (.047)
Snow in December	-0.225 (.041)***	-0.295 (.055)***	-0.112 (.041)**	-0.315 (.049)***	-0.097 (.044)**	-0.152 (.059)**
Snow in January	0.053 (.033)	0.022 (0.031)	0.019 (.024)	0.115 (.047)**	0.142 (.030)***	0.019 (.052)
Snow in February	-0.4 (.060)***	-0.258 (.065)***	-0.105 (.035)***	-0.287 (.054)***	-0.168 (.046)***	-0.207 (.074)**
Snow in Mar./April in MSPAP years	0.017 (.070)	-0.148 (.052)***	-0.013 (.054)	-0.037 (.052)	-0.059 (.024)**	-0.018 (.037)
Snow in Mar./April in MSA years	0.378 (.216)	0.181 (.157)	0.097 (.112)	0.173 (.109)	0.386 (.203)*	0.209 (.106)*
R ²	0.8888	0.8799	0.8976	0.8664	0.9481	0.9439

Robust standard errors in parentheses

** significant at 5%; *** significant at 1%

Models also control for school and student characteristics, school fixed effects, district specific trends and dummies to measure change in test regime.

Table 6**Estimated Impact of Unscheduled Closings on Schools Failing
to Make Adequate Yearly Progress on 3rd Grade Subject Tests**

	2003		2004		2005	
	Reading	Math	Reading	Math	Reading	Math
Total Number of Schools	828		830		833	
Schools Failing to Make AYP in Subject	52	56	21	24	35	42
Failing Schools within "Closings Effect" of AYP	30	34	11	15	12	9
Failing Schools further than "Closings Effect" from AYP	22	22	10	9	23	33