OPEN

A Quasi-Experimental Analysis of Elementary School Absences and Fine Particulate Air Pollution

Nicholas M. Hales, Caleb C. Barton, BS, Michael R. Ransom, PhD,
Ryan T. Allen, and C. Arden Pope III, PhD

Abstract: Fine particulate air pollution ($PM_{2.5}$) has been associated with many adverse health outcomes including school absences. Specifically, a previous study in the Utah Valley area, conducted during a time with relatively high air pollution exposure, found significant positive correlations between school absences and air pollution.

We examined the hypothesis that ambient $PM_{2.5}$ exposures are associated with elementary school absences using a quasi-natural experiment to help control for observed and unobserved structural factors that influence school absences. The Alpine, Provo, and Salt Lake City school districts are located in valleys subject to daily mean $PM_{2.5}$ concentrations almost twice as high as those in the Park City School District. We used seminonparametric generalized additive Poisson regression models to evaluate associations between absences and daily $PM_{2.5}$ levels in the 3 districts that were exposed to the most pollution while using Park City absences as a quasi-control. The study covered 3 school years (2011/12-2013/14).

School absences were most strongly associated with observed structural factors such as seasonal trends across school years, day-of-week effects, holiday effects, weather, etc. However, after controlling for these structural factors directly and using a control district, a 10 µg/ m^3 increase in $PM_{2.5}$ was associated with an approximately 1.7% increase in daily elementary school absences.

Exposure to ambient air pollution can contribute to elementary school absences, although this effect is difficult to disentangle from various other factors.

(*Medicine* 95(9):e2916)

Abbreviations: O_3 = Ozone, ACS = American Community Survey, AIC = Akaike's information criterion, AQS = Air Quality System, CO = carbon monoxide, EPA = U.S. Environmental Protection Agency, PM_{10} = suspended particles with an aerodynamic diameter less than or equal to 10-µm cut point, $PM_{2.5}$ = suspended particles with an aerodynamic diameter less than or equal to 2.5-µm cut point.

INTRODUCTION

Fine particulate matter air pollution ($PM_{2.5}$) has been associated with various adverse cardiopulmonary health outcomes.[1–6] Several studies have found positive associations between air pollution and school absences among children. School absences have been associated with high carbon monoxide (CO) levels.[7] Short-term changes in ozone (O_3) have also been associated with increased respiratory illness among children, leading to increases in school absences.[8] Other studies have found that exposure to air pollution places children at risk for asthma exacerbations and hospitalizations, as well as deficits in lung growth. Some researchers have even suggested that pediatricians should educate children and their families about how to avoid the harmful effects of air pollution by rescheduling outdoor activities.[9]

In the Utah Valley, positive associations have been found between particulate air pollution (PM_{10}) and absenteeism among elementary schoolchildren.[10] Increased absenteeism is negatively correlated to many measures of student performance, including pass rates on standardized tests[11] and teacher-assigned reading grades.[12] Moreover, in many cases such absences cause parents to miss work[13] and thus add to the overall economic burden caused by particulate air pollution. The present study evaluates associations between pollution and elementary school absenteeism in 3 school districts (Alpine, Provo, and Salt Lake City) located in two moderately polluted Wasatch Front valleys of Utah using a neighboring control district (Park City) in a quasi-experimental approach[14] that capitalizes on Park City's much lower pollution levels.

A primary difficulty of evaluating associations between air pollution and school absences is controlling for structural factors that influence absences. Such structural factors can be conceptualized using 2 categories: observed structural factors and unobserved structural factors. Observed structural factors are those factors that can be measured and explicitly controlled for in statistical models, for example, seasonal trends across the school year, holidays, day-of-week effects, weather effects, etc. Unobserved structural factors are those factors that are difficult or impossible to measure and model within a statistical framework, such as one-time events that cause parents to pull their kids out of school (changing concerns regarding school safety, local events, etc.). In an idealized experimental framework, an identical school district without exposure to air pollution, but that experienced all other common factors that influence absences, would be used as a control. Such a framework would control for both observed structural factors and unobserved factors; pollution would be the only factor to vary across districts.

The present study takes advantage of conditions that approximate an idealized experimental framework. Park City absences are influenced by observed structural factors that are nearly identical to those that affect the other districts and presumably by unobserved factors that are similar to those in

the other districts. The Park City School District, however, is located away from the larger metropolitan areas, sits at an elevation above the typical stable temperature inversion layer, and has much lower concentrations of air pollution. Thus, Park City school absences should not be affected by pollutants from the Utah and Salt Lake valleys. Variations in Park City absences theoretically represent, at least in part, the impact of both observed and unobserved factors while not being affected by the fluctuations in pollution that occur in the more urban Wasatch valleys. Controlling for absences from Park City provides a reasonable quasi-experimental approach to help control for unobserved factors. Moreover, if structural controls for observed factors of absences are sufficient to prevent confounding, Park City serves to verify the effectiveness of these controls.

The primary hypothesis of this analysis is that elementary school absences from the Alpine, Provo, and Salt Lake City school districts are positively associated with $PM_{2.5}$ concentrations, even while controlling for changes in absences in the Park City School District and while controlling directly for various structural and weather-related relationships (day-of-week, holidays, seasonality, temperature, and snowfall). To more confidently test this hypothesis, we investigate the plausibility that variation in Park City absences represents variation in absences attributable to all factors, observed and unobserved, common across the 4 districts under consideration.

METHODS

Study Areas

The geographic locations of the 4 school districts used in this study are presented in Figure 1. The Alpine and Provo school districts are both located on the valley floor of Utah Valley at an elevation of approximately 1400 m. The Salt Lake City School District is in the Salt Lake Valley (just north of Utah Valley) with an elevation of approximately 1300 m. These three districts are in urban areas and share common weather patterns. During low-level temperature inversion episodes, $PM_{2.5}$ concentrations are elevated, with local emissions becoming trapped in a stagnant air mass near the valley floor. In contrast, the Park City School District is situated east of Salt Lake Valley at an elevation of approximately 2070 m, in a higher Wasatch Mountains area that includes ski resorts. The highway (I-80) from Salt Lake City to Park City is useful in considering elevation differences between Park City and the other three study areas: elevation along this road rises from 1300 m to a maximum of approximately 2135 m, and then decreases until reaching approximately 2070 m in the vicinity of Park City. The mountain range between Park City and the Alpine and Provo areas has peaks reaching as high as 3850 m. The stable temperature inversion layer is typically at an elevation of approximately 1800 m.[15–17] Although the Park City School District is located only approximately 20 miles away from the metropolitan area of Salt Lake City, it is at an elevation that is typically above the temperature inversion level. Thus, Park City has much lower average air pollution levels than the other 3 school districts and serves as a quasi-control school district.

Absences and Enrollment Data

Daily counts of elementary school absences for 3 school years (2011/12-2013/14) and information regarding school days and vacation days were obtained from each school district for all elementary schools in the 4 school districts. In each school district, daily counts for all schools were added to get a daily absence count for the district. Days in which not all elementary schools were in session or in which not all schools reported an absence count were recorded as missing observations and excluded from the analysis. Thus, the district-wide daily absence counts represent the sum of absences for a fixed set of schools for each day. A few schools were excluded from the analysis to increase the consistency of comparisons. These included 2 Kindergarten through 8th grade schools in the Salt Lake City School District and 1 school in the Alpine School District with incomplete absences data. It is noted that 29 schools in Alpine and 19 schools in Salt Lake have pre-kindergarten students, and three schools in Salt Lake and all of the schools in Park City do not have sixth grade students.

Data concerning baseline enrollment by grade level in the 2011 to 2012 school year were collected from the National Center for Education Statistics Common Core of Data.[18]

Pollution and Weather Data

Daily concentrations of fine particulate matter air pollution (suspended particles with an aerodynamic diameter less than or equal to 2.5- μm cut point, $PM_{2.5}$) for the 3 school years under consideration were available from central-site monitors located in the Alpine, Provo, and Salt Lake City school districts. The $PM_{2.5}$ data were compiled from the US Environmental Protection Agency (EPA) Air Quality System (AQS).[19] Limited $PM_{2.5}$ data, with less quality assurance, for the Park City School District were obtained from the Summit County Health Department. For each of the 4 districts, a central monitoring site was chosen based on central location and availability of a comprehensive collection of daily $PM_{2.5}$ concentrations. In addition to $PM_{2.5}$ measurements from central monitoring sites, data were also collected from monitors located at 4 other nearby sites (Bountiful, Spanish Fork, North Salt Lake, and Kamas). Daily $PM_{2.5}$ concentration data from the central sites were regressed on the data collected from surrounding sites. $PM_{2.5}$ concentrations across nearby sites were highly correlated (R^2 values 0.80–0.95). Missing daily $PM_{2.5}$ concentrations at the central site were imputed using the regression results and the $PM_{2.5}$ concentrations from the $PM_{2.5}$ monitor whose regression yielded the highest R^2 values. The number of imputations for each area ranged from 4 (Provo) to 111 (Park City). Additionally, PM_{10} data were gathered and imputed comparable with $PM_{2.5}$ to evaluate the sensitivity of results.

Weather data documenting mean daily temperature (Fahrenheit) and daily snowfall (inches) at the Salt Lake City International Airport (see Figure 1 for location) as well as temperature and daily snowfall for Park City (from the Thayne's canyon and Park City monitoring stations, respectively) were collected from the National Climatic Data Center.[20]

Demographic and Socioeconomic Data

Data regarding the education level, median income, and percent of families below the poverty line for each school district were collected from the 2013 American Community Survey (ACS).[21] Additionally, data indicating the percent of the population that was black, Asian, and Hispanic or Latino were collected from the 2010 Total Population Census.[22] Data from both the ACS and the 2010 census were collected at the zip code level. Estimates of these variables at the school district level were calculated using population, household, and family weighted averages of those zip codes that primarily comprised each of the 4 school districts.
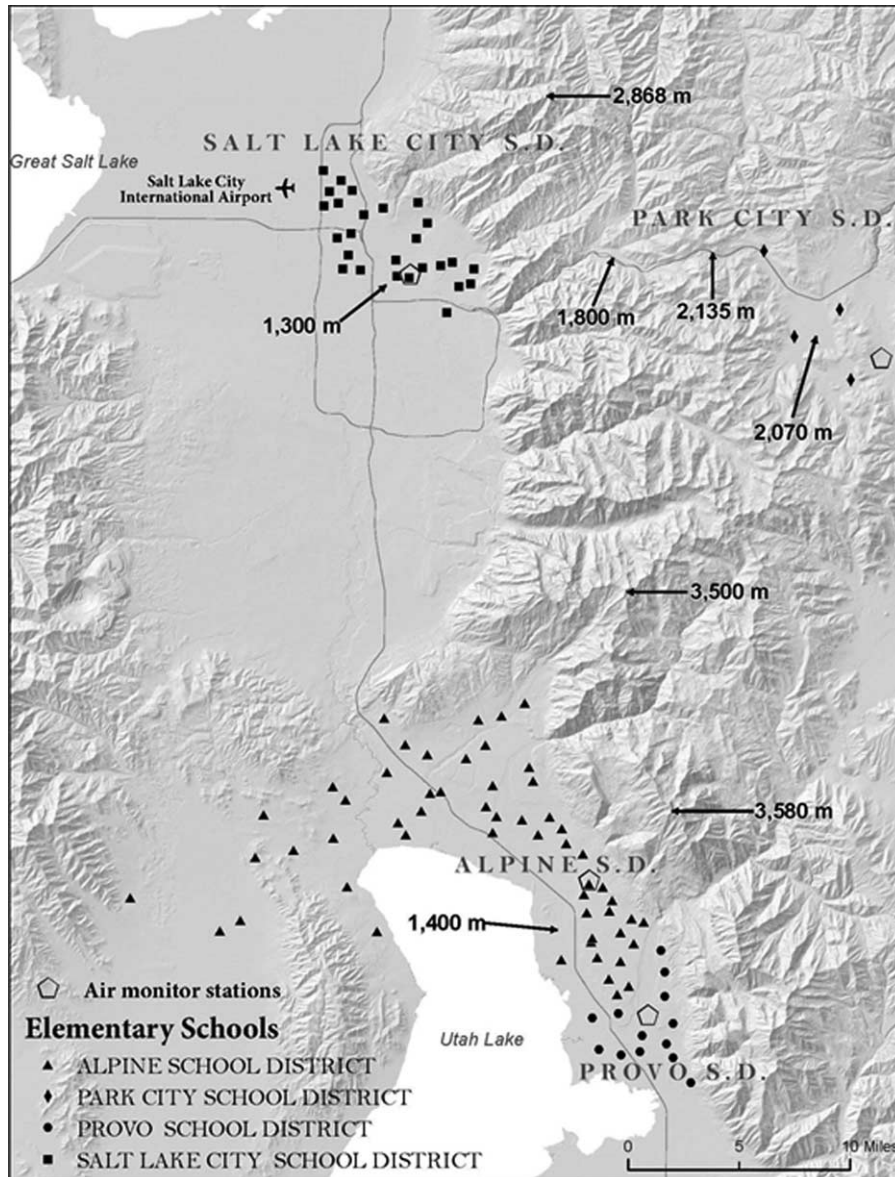
FIGURE 1. Map of the 4 school districts under consideration, including elevations of $PM_{2.5}$ monitors, study areas, and nearby mountain peaks.

Statistical Analysis

We used seminonparametric generalized additive Poisson regression models to estimate the relationship between $PM_{2.5}$ and absenteeism while controlling for the observed structural factors that influence elementary school absences. This analysis was conducted for the Alpine, Provo, and Salt Lake City school districts. The dependent variable in all models was the total count of students absent per day for each district.

Five models were used to investigate the air pollution effect. Each model was estimated using $PM_{2.5}$ from the previous day as well as 3, 5, and 7-day lagged moving averages. Model 1 included $PM_{2.5}$ as a linear independent variable. Model 2 included $PM_{2.5}$ and daily absence counts in Park City as linear independent variables. As explained previously, absences in

Park City were used to control for both observed and unobserved structural determinants of absenteeism. Because of Park City's higher elevation and lower pollution levels, absences in Park City are less likely to be pollution induced and serve as a relevant control for shared structural factors that influence attendance. The estimation of model 2 serves as a benchmark by which to measure subsequent models that control explicitly for observed structural factors.

Model 3 directly controlled for observed factors influencing absences using structural controls. Specifically, model 3 dropped the number of absences in Park City and included the following variables: indicator variables for each of the 3 school years; indicator variables for school days immediately before or after a school holiday; functions of day-of-week

TABLE 1. Yearly and Total Number of School Days, as well as Mean Daily Absences and Standard Deviations for the 4 School Districts Under consideration

School Year	Alpine (53 Schools) 2011–12 Enrollment: 41,039 Avg. % Absent*: 4.19%			Provo (13 Schools) 2011–12 Enrollment: 7,934 Avg. % Absent: 4.44%			Salt Lake City (27 Schools) 2011–12 Enrollment: 13,850 Avg. % Absent: 9.79%			Park City (4 Schools) 2011–12 Enrollment: 1,945 Avg. % Absent: 4.32%		
	School Days†	Mean Daily Absences	SD	School Days	Mean Daily Absences	SD	School Days	Mean Daily Absences	SD	School Days	Mean Daily Absences	SD
2011/12	176	1720	455	173	352	94	177	1356	292	178	84	36
2012/13	177	1959	618	178	400	134	177	1445	493	176	99	41
2013/14	177	1900	458	178	380	90	177	2106	413	176	88	35
Total	530	1860	525	529	378	110	531	1636	527	530	90	38

* Average absences as a percentage of enrollment in the 2011–12 school year.
† Number of days includes only days in which all schools in each district were in session and reported a daily absence count.

(Monday–Friday corresponding to days of week 2–5, respectively); a smooth function of the number of days from the start of each school year; and smooth functions of the average daily temperature (measured in Fahrenheit) and the daily snow fall (measured in inches). The smooth functions were penalized cubic spline functions calculated with 3 degrees of freedom. Thus, rather than controlling for structural factors implicitly by using Park City, model 3 explicitly controls for a variety of observed structural factors hypothesized to influence daily absences.

Model 4 combines the explicit structural controls of model 3 with the implicit control of Park City; it included the same controls as Model 3 as well as daily absences from Park City. If, as hypothesized, variation in Park City absences represents the impact of observed and unobserved structural determinants of absences, then model 4 doubly controls for observed structural factors and appropriately controls for unobserved structural factors. Model 5 is similar to model 4, but is intended to remove this double controlling. Model 5 includes the explicit controls of model 3 and an implicit control calculated from Park City. This model was the same as model 4, but rather than using daily absences from Park City, this model used the residuals calculated from estimating the number of Park City absences using observed structural factors (a regression similar to model 3 but using Park City absences as the dependent variable, excluding pollution, and using temperature and snowfall from Park City rather than from Salt Lake City). Because these residuals are calculated from a model that excludes pollution, they theoretically represent the impact of all the unobserved structural factors and pollution (whose impact in Park City is hypothesized to be zero). Thus, model 5 explicitly controls for observed structural factors and implicitly controls for unobserved structural factors.

All models were also estimated using PM_{10} instead of $PM_{2.5}$, although only a selection of these results is presented here. The smooth functions in models 3, 4, and 5 were calculated separately for each district and lag structure. Models were estimated using SAS, version 9.4, PROC GAM (SAS Institute, Cary, NC).

Ethics Review

School absences data were compiled from legally accessed administrative data provided by the public school districts that did not include any individual student information. The baseline enrollment, pollution, weather, demographic, and socioeconomic data came from other publically and legally accessible data sources. Therefore, review by an institutional review board for human subjects or ethics committee was not necessary.

RESULTS

Table 1 presents the means and standard deviations for the daily absences in all 4 school districts for each of the 3 school years. Table 1 also gives baseline enrollment for the 2011 to 2012 school year and the average daily percent of students absent that year. Compared with a similar analysis performed during the period from 1985/86 school year to the 1990/91 school year,[10] the average percent of students absent has decreased slightly. From 1985/86 to 1990/91, the average weekly percent of 1st to 6th grade students absent in the Provo district was 5.11%. During the 2011/12 school year, average daily percent absent in the Provo district was 4.44%, slightly lower the 1985/86 to 1990/91 period. Average daily percent absent at Northridge Elementary (in Alpine School District) was

TABLE 2. Total Number of Daily PM$_{2.5}$ Measurements, Including Means, Standard Deviations, and Low and High Values for Daily PM$_{2.5}$ Measurements for Each School District

School District	Central-site Monitor	Central Monitor Only					Central Monitor and Imputed Values				
		N	Mean	SD	Low	High	N	Mean	SD	Low	High
Alpine PM$_{2.5}$	Lindon	990	9.52	12.15	0	123.3	1076	9.45	11.77	0	123.3
Provo PM$_{2.5}$	North Provo	1072	9.01	11.10	0	124.5	1076	9.04	11.17	0	124.5
Salt Lake City PM$_{2.5}$	Hawthorne	1030	9.66	10.78	0	69.2	1073	9.55	10.61	0	69.2
Park City PM$_{2.5}$	Park City	303	4.33	2.26	0	11.8	414	4.32	2.29	0	13.7

4.54% from 1985/86 to 1990/91. Although the data do not allow identification of Northridge Elementary specifically, average daily percent absent in the Alpine district in 2011/12 was 4.19%. These numbers, although not directly comparable, do indicate that on average the percent of students absent during the time period of our study was slightly lower than during the 1985/86 to 1990/91 period.

Means and standard deviations for daily PM$_{2.5}$ concentrations from monitors near the Alpine, Provo, Salt Lake City, and Park City areas are presented in Table 2. The mean concentrations (and SD) of PM$_{2.5}$ for central site values were nearly the same as those for the central site plus imputed values. Although the PM$_{2.5}$ data are limited for Park City, it is noted that the average pollution levels are much lower.

Table 3 gives demographic and socioeconomic variables aggregated across the zip codes that primarily comprise each school district. Additionally, Table 3 provides a breakdown of 2011–12 enrollment into 3 buckets: percent prekindergarten and kindergarten students, percent 1st to 3rd grade students, and percent 4th to 6th grade students.

Table 4 gives simple Pearson correlation coefficients between absences in all 4 districts, pollution from the Lindon, North Provo, and Hawthorne PM$_{2.5}$ monitors, and temperature in Salt Lake City and Park City. Daily school absences and PM$_{2.5}$ concentrations were highly correlated between the 2 school districts in the Utah Valley (Alpine and Provo) but less so with Salt Lake City and Park City. Across the entire year, temperatures between Salt Lake City and Park City were highly correlated (0.92, $P < 0.001$), reflecting seasonal correlations in temperatures; however, the correlation between these 2 temperature measurements was also computed for only winter months and for inversion days (Salt Lake City PM$_{2.5} > 30\,\mu g/m^3$). These

correlations were much lower than the correlation for the entire year, and were calculated as 0.29 ($P < 0.001$) for winter months and 0.24 ($P = .06$) for inversion days.

Figures 2–5 present school absences for the Alpine, Provo, Salt Lake City, and Park City school districts and daily PM$_{2.5}$ concentrations plotted over time for each of the 3 school years. Several structural patterns in school absences are easily observed, including: day-of-week effects, with the lowest absenteeism on Wednesdays and higher absences nearer weekends; higher absences just before breaks and holiday weekends; and a nonlinear school year time trend with absences being low at the beginning of the school year, rising for a time, and then falling toward the end of the school year. It is noted that absences were highest during the 2012 to 2013 school year, which was also the year of highest influenza illness of the 3 years under consideration.[23]

Table 5 presents regression results from each of the 5 models for Alpine, Provo, and Salt Lake City. Model 1 yields significant positive correlations between PM$_{2.5}$ and school absences for all lag structures for each of the 3 school districts. Model 2, which includes Park City absences as a control variable, gives slightly smaller coefficients than Model 1, but estimated associations are all positive and statistically significant. Model 3, which drops Park City absences but includes indicator variables for school year and holidays as well as flexible spline smoothers to control for day-of-week and seasonal and weather patterns, finds positive, statistically significant associations for PM$_{2.5}$ in both the Provo and Alpine school districts. Estimated associations from Model 3 for the Salt Lake City School District are positive and statistically significant for 3, 5, and 7-day lagged moving averages of pollution, but negative and statistically significant using pollution exposure from the previous day. Estimated associations from Model 4 were all positive and significant except for Salt Lake City using pollution

TABLE 3. Demographic and Socioeconomic Characteristics for Each School District, as well as a Breakdown of 2011–2012 Enrollment by Grade Ranges

	2013 Demographic and Socioeconomic Variables						2011–12 Enrollment by Grade Ranges[†]		
	% High School Graduates[*]	Median income	% Below Poverty Line	% Black	% Asian	% Hispanic or Latino	% Prekindergarten and Kindergarten	% 1–3 Grade	% 4–6 Grade
Alpine	95%	$67,274	8%	1%	1%	9%	18%	43%	40%
Provo	91%	$40,714	21%	1%	2%	15%	17%	44%	39%
Salt Lake City	86%	$48,073	14%	3%	4%	22%	20%	43%	37%
Park City	89%	$83,202	18%	1%	2%	24%	16%	50%	35%

[*] Percent of the population age 25 or over with a high school diploma or equivalent.
[†] Note that numbers may not add to 100 because of rounding.

TABLE 4. Pearson Correlation Coefficients Between Daily Absences and Daily PM$_{2.5}$ Measurements

	Alpine Absences	Provo Absences	Salt Lake Absences	Park City Absences	Alpine PM$_{2.5}$	Provo PM$_{2.5}$	Salt Lake PM$_{2.5}$
Alpine Absences	1	0.93[*]	0.68[*]	0.66[*]	0.32[*]	0.31[*]	0.29[*]
Provo Absences	0.93[*]	1	0.66[*]	0.65[*]	0.34[*]	0.33[*]	0.29[*]
Salt Lake City Absences	0.68[*]	0.66[*]	1	0.47[*]	0.15[*]	0.13[*]	0.16[*]
Park City Absences	0.66[*]	0.65[*]	0.47[*]	1	0.12[*]	0.11[*]	0.06
Alpine PM$_{2.5}$	0.32[*]	0.34[*]	0.15[*]	0.12[*]	1	0.97[*]	0.86[*]
Provo PM$_{2.5}$	0.31[*]	0.33[*]	0.13[*]	0.11[*]	0.97[*]	1	0.82[*]
Salt Lake City PM$_{2.5}$	0.29[*]	0.29[*]	0.16[*]	0.06	0.86[*]	0.82[*]	1

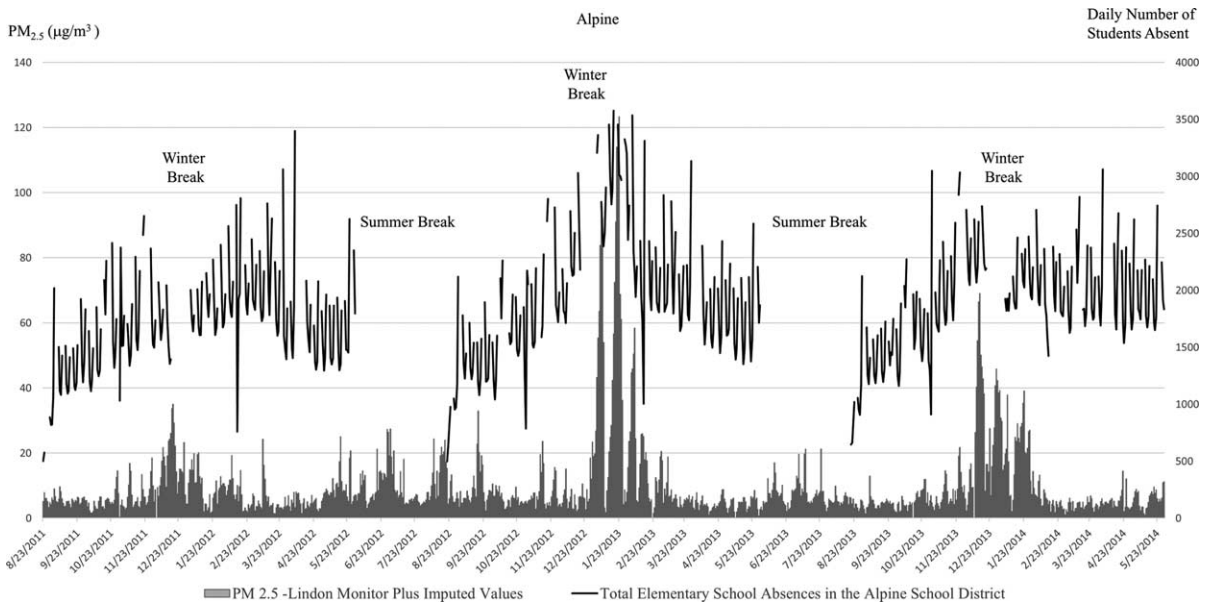[*] Indicates statistical significance at $P < 0.01$.

FIGURE 2. Alpine elementary school absences and daily PM$_{2.5}$ levels from the Lindon monitor plus imputed values.
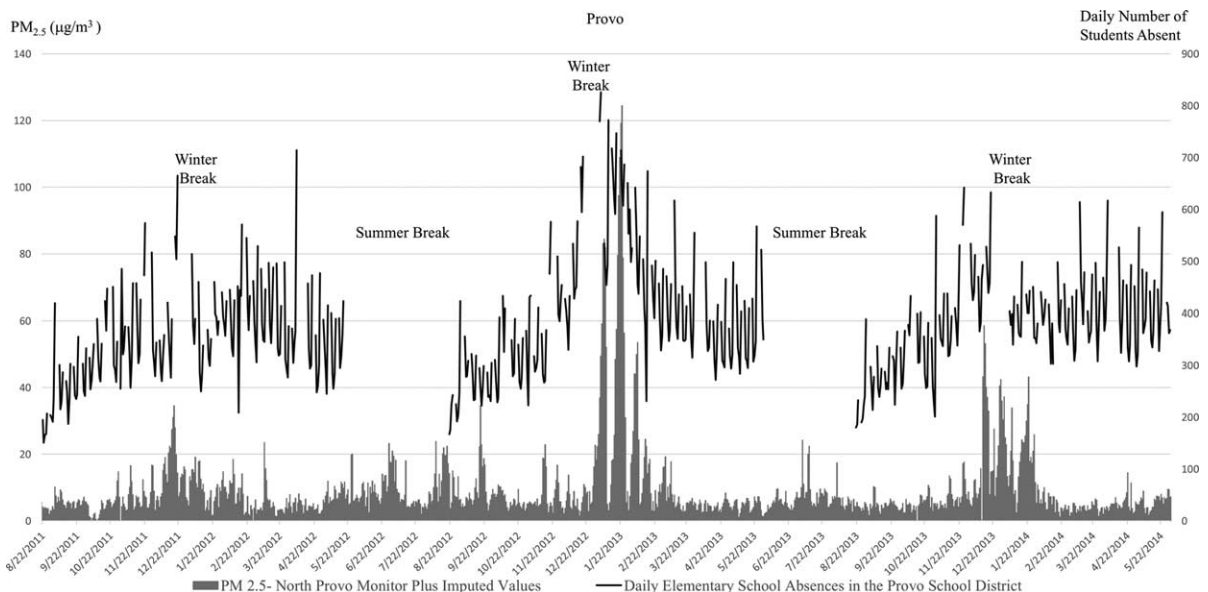

FIGURE 3. Provo elementary school absences and daily PM$_{2.5}$ levels from the North Provo monitor plus imputed values.
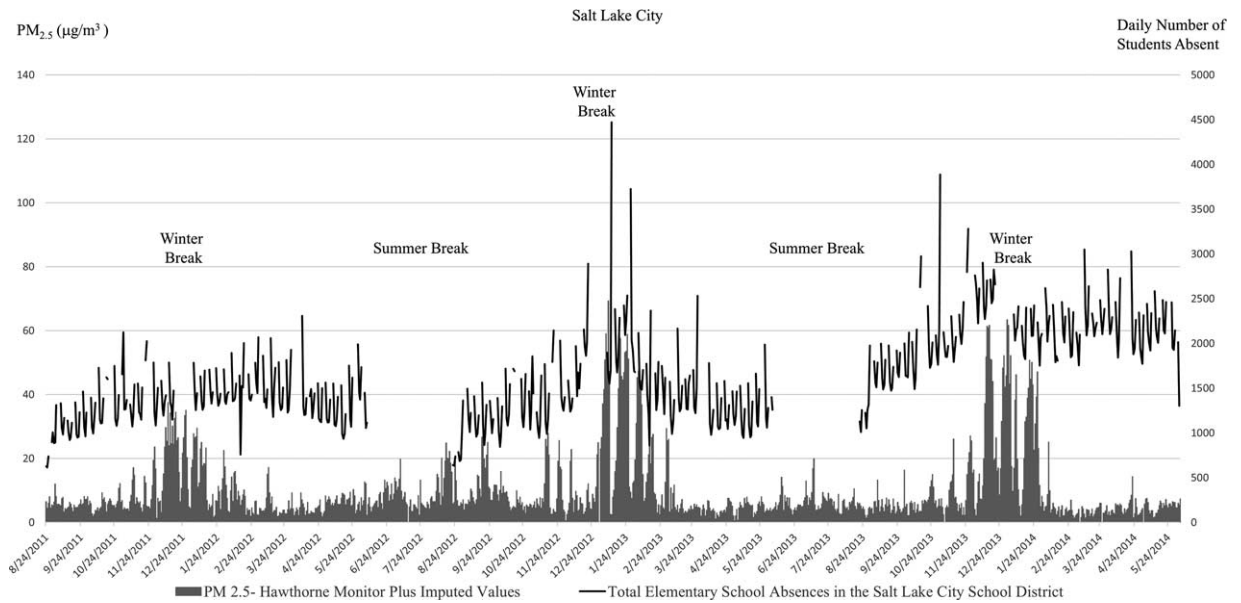
FIGURE 4. Salt Lake City elementary school absences and daily $PM_{2.5}$ levels from the Hawthorne monitor plus imputed values.

from the previous day, which yielded a coefficient that was negative but statistically insignificant.

Model 5 yielded similar results to model 4 in terms of significance and direction, although estimated associations were slightly smaller. Again, the relationship between absences and pollution was significant and negative only for 1 day lagged exposures in Salt Lake. For each study area and lag structure, we calculated the Akaike information criterion (AIC) for each of the 5 models. AIC values were uniformly lowest for model 5 for all lag structures and study areas, indicating that model 5 had the best fit relative to all models considered for estimating the relationship between $PM_{2.5}$ and school absences.

As hypothesized, school absences from the Alpine, Provo, and Salt Lake City school districts were associated with absences from the control district, Park City. Using a 5-day lagged moving average of $PM_{2.5}$, model 2 indicates an approximately 4% increase in absences associated with a 10-person (or approximately 12%) increase in absences in Park City, with the relationship being highly statistically significant ($P < 0.001$). These results suggest that much of the structural variability, from observed and potentially from unobserved factors, is at least partially controlled for with the inclusion of Park City absences. When variables that directly control for the observed structural factors are also included in the model (Model 4), the
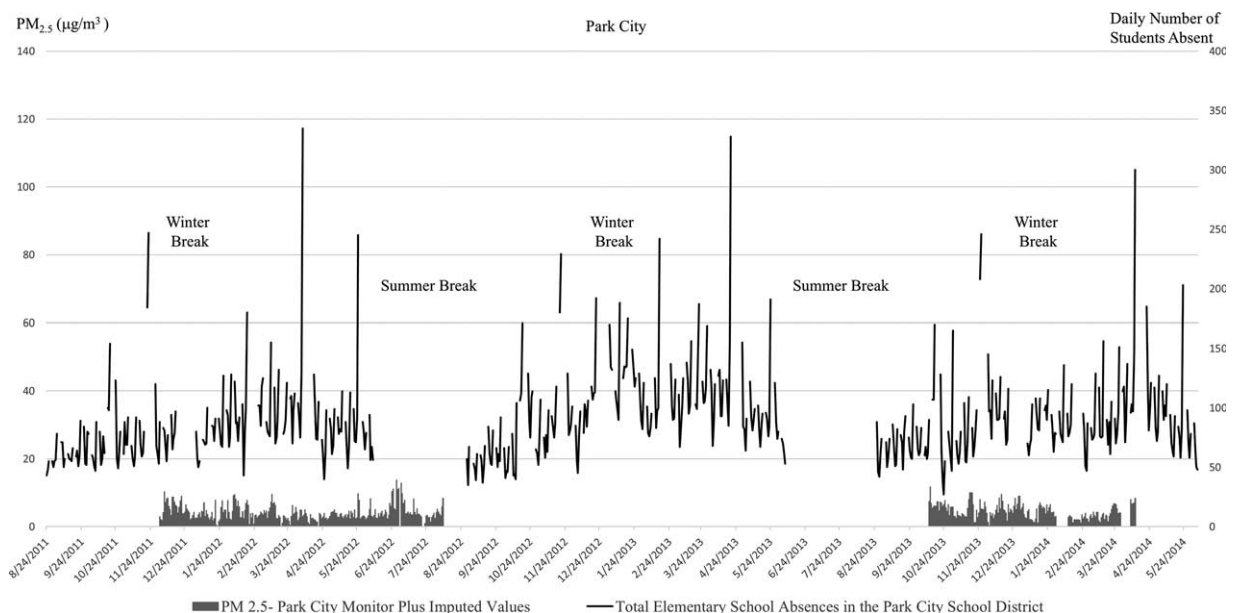

FIGURE 5. Park City elementary school absences and daily $PM_{2.5}$ levels from the Park City monitor plus imputed values.

TABLE 5. Percent Increases in Elementary School Absences Associated With Changes in PM$_{2.5}$ for the Alpine, Provo, and Salt Lake City School Districts. Note That s () Indicates a Penalized Cubic Spline Function Calculated With 3 Degrees of Freedom

PM$_{2.5}$	District	Model 1 % Increase in Absences With a 10 μg/m³ Increase in PM$_{2.5}$ (95% CI)	Model 2 % Increase in Absences With a 10 μg/m³ Increase in PM$_{2.5}$ (95% CI)	Model 3 % Increase in Absences With a 10 μg/m³ Increase in PM$_{2.5}$ (95% CI)	Model 4 % Increase in Absences With a 10 μg/m³ Increase in PM$_{2.5}$ (95% CI)	Model 5 % Increase in Absences With a 10 μg/m³ Increase in PM$_{2.5}$ (95% CI)
Previous Day	Alpine	5.67 (5.5, 5.8)	4.37 (4.2, 4.5)	2.08 (1.9, 2.2)	1.92 (1.8, 2.1)	1.49 (1.3, 1.6)
	Provo	6.05 (5.8, 6.3)	4.58 (4.3, 4.9)	2.30 (2.0, 2.6)	2.08 (1.7, 2.4)	1.72 (1.4, 2.1)
	Salt Lake	4.42 (4.2, 4.6)	3.40 (3.2, 3.6)	−0.47 (−0.7, −0.3)	−0.19 (−0.4, 0.0)	−0.62 (−0.8, −0.4)
3-Day Moving Average	Alpine	6.16 (6.0, 6.3)	4.87 (4.7, 5.0)	1.93 (1.8, 2.1)	1.88 (1.7, 2.0)	1.60 (1.4, 1.8)
	Provo	6.82 (6.5, 7.1)	5.30 (5.0, 5.6)	2.78 (2.4, 3.1)	2.50 (2.1, 2.9)	2.25 (1.9, 2.6)
	Salt Lake	6.20 (6.0, 6.4)	5.33 (5.1, 5.5)	0.83 (0.6, 1.1)	1.44 (1.2, 1.7)	1.15 (0.9, 1.4)
5-Day Moving Average	Alpine	7.04 (6.9, 7.2)	5.71 (5.6, 5.8)	2.09 (1.9, 2.3)	2.08 (1.9, 2.3)	1.82 (1.6, 2.0)
	Provo	7.89 (7.6, 8.2)	6.18 (5.9, 6.5)	3.16 (2.8, 3.6)	2.79 (2.4, 3.2)	2.56 (2.2, 3.0)
	Salt Lake	7.07 (6.9, 7.3)	6.26 (6.1, 6.5)	0.95 (0.7, 1.2)	1.72 (1.5, 2.0)	1.48 (1.2, 1.7)
7-Day Moving Average	Alpine	8.29 (8.1, 8.5)	6.85 (6.7, 7.0)	2.66 (2.4, 2.9)	2.73 (2.5, 2.9)	2.42 (2.2, 2.6)
	Provo	9.26 (8.9, 9.6)	7.34 (7.0, 7.7)	3.91 (3.4, 4.4)	3.45 (3.0, 3.9)	3.16 (2.7, 3.6)
	Salt Lake	7.80 (7.6, 8.0)	6.99 (6.8, 7.2)	0.59 (0.3, 0.9)	1.59 (1.3, 1.9)	1.32 (1.0, 1.6)

CI = confidence interval.
Model 1 $Absences = PM2.5 + \varepsilon$.
Model 2 $Absences = PM2.5 + AbsencesPC + \varepsilon$.
Model 3 $Absences = PM2.5 + Yeari + Holiday + s(Dayofweek) + s(DaysFromStart) + s(Temp) + s(Snow) + \varepsilon$.
Model 4 $Absences = PM2.5 + AbsencesPC + Yeari + Holiday + s(Dayofweek) + s(DaysFromStart) + s(Temp) + s(Snow) + \varepsilon$.
Model 5 $Absences = PM2.5 + ResidualsPC + Yearsi + Holiday + s(Dayofweek) + s(DaysFromStart)s(Temp) + s(Snow) + \varepsilon$.
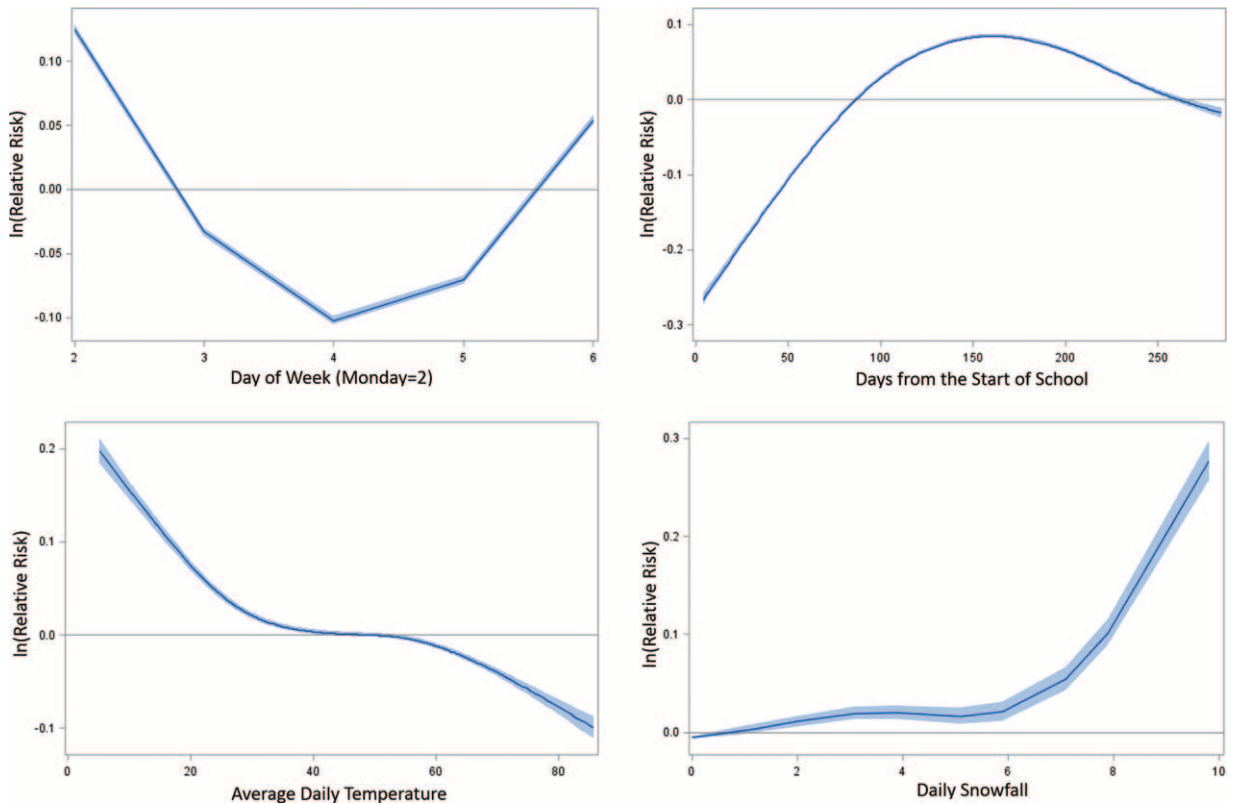
FIGURE 6. Cubic spline smoothers for day-of-week, days from the start of the school year, average daily temperature, and snow fall for the Alpine school district using 5-day lagged moving average values of $PM_{2.5}$ (Model 4: Fit and 95% confidence interval).

associations with Park City absences are attenuated (approximately 2% increase in absences associated with a 10-absence increase in Park City, $P < 0.001$), but remain significant. These results suggest that there remain some unobserved factors common across school districts that are at least partially controlled for by including the control district. Model 5 further investigates the impact of such unobserved factors. Estimated relationships are slightly smaller than model 4, and the associations with Park City residuals are positive and highly statistically significant ($P < 0.001$). Overall, however, in the most fully controlled models (models 4 and 5), the association between $PM_{2.5}$ and absences is quite small, with a 3 to 7-day lagged moving average increase of $10 \mu g/m^3$ $PM_{2.5}$ being associated with an approximately 1.5 to 2% increase in school absences.

Results for PM_{10} (not presented) are highly comparable with results presented for $PM_{2.5}$ in direction and significance. The average ratios of $PM_{2.5}$ to PM_{10} at the Lindon and Hawthorne monitoring sites (those with the most available $PM_{2.5}$ and PM_{10} data) were 0.62 and 0.48, respectively, during the days for which school was in session. The magnitudes of the estimated PM_{10} effects reflect this relationship, with the mean ratio of PM_{10} effects to $PM_{2.5}$ effects for all models and lag structures equal to 0.59, 0.69, and 0.61 in Alpine, Provo, and Salt Lake City, respectively. Results from model 5 indicate that a $10-\mu g/m^3$ increase in 7-day lagged moving average PM_{10} is associated with an approximately 1 to 1.5% increase in absences.

Figure 6 presents the estimated associations between short-term $PM_{2.5}$ exposure and day-of-week, days from the start of the school year, average daily temperature, and daily snow fall for the Alpine School District using Model 4. These relationships were significantly nonlinear. Similar associations are observed in the other districts. As is seen in Figure 6, days closer to weekends are associated with higher absenteeism. Winter months are associated with higher numbers of absences, especially before and after Christmas breaks. Additionally, colder temperatures and larger amounts of snowfall are associated with more absences.

DISCUSSION

For the Alpine, Provo, and Salt Lake City school districts, elementary school absences are associated with daily $PM_{2.5}$ concentrations even while controlling for absences (or residuals) in the control district (Park City) and controlling for various observed structural and weather factors. These results are suggestive of an effect of air pollution on elementary school absences. Regardless of choice of model or lag structure, our estimates indicate a positive relationship for the Provo and Alpine school districts. In Salt Lake City, $PM_{2.5}$ effects on school absences were somewhat smaller and less consistent, perhaps because overall absenteeism is much higher in Salt Lake. The $PM_{2.5}$ effects were generally strongest for longer lag structures (3 to 7-day lagged moving average exposures).

This analysis has several strengths. An epidemiologic analysis that evaluates the relationship between absences and pollution by simply including observed structural factors in a regression model must assume that there are no unobserved factors excluded from the model that are correlated with

pollution. This approach is analogous to model 3 in our study. The inclusion of Park City as a control district introduces a quasi-experimental aspect that improves control for unobserved factors, but falls short of being a true randomized experiment. Although this study remains an observational study, it has one clear advantage to the traditional approach: a representation, even if imperfect, of the unobserved structural factors influencing absences. As mentioned, Park City experiences much lower levels of air pollution throughout the school year. Owing to its close proximity and socioeconomic similarity, we expect such structural factors to be common across all 4 districts. Thus variation in absences in Park City represents the variation attributable to shared observed and unobserved structural factors except for pollution. Therefore, although we neither explicitly control for all factors influencing absences nor randomly assign pollution exposure, we have the advantage of being able to check our estimates (model 3) against models that can more fully control for potential unobserved confounders (models 4 and 5).

Given that estimates remain relatively stable across models 3, 4, and 5 we have more confidence that the estimated relationship is not due to confounding by unobserved factors. Rather than assuming that there are no factors in the error term correlated with pollution and absences, one need to only assume that there are no factors unique to Salt Lake, Alpine, or Provo and not present in Park City, that are unobserved, uncontrolled for, and correlated with pollution. We are unaware of a potential confounder that is correlated to pollution, affects school absences, is not captured by the extensive structural controls included in model 3, and is not present in Park City.

In addition to the advantages of using Park City, the controls for observed structural factors are exhaustive. Fixed effects for school year allow us to control for year-to-year changes in overall number of students enrolled in each district as well as for all other factors influencing absences that are permanent throughout a given school year. Indicator variables for days before and after holidays control for the increases in absences attributable to travel and festivities. Similarly, day-of-week variables control for weekly trends in absences (increased absences on Mondays and Fridays), whereas smooth terms for days from the start of school and weather trends control for seasonal and weather effects on absences. The year fixed effects and nonlinear seasonal controls also serve to model influenza prevalence, which is highly seasonal and varies from year to year. This seasonal impact can be seen in Figure 6. The concave shape of the function on days from the start of the school year captures seasonal effects of influenza.

This analysis also has important limitations. The large majority of the variability in school absences is related to structural factors such as day-of-week, holidays, and trends across school years. Even after attempts to model these factors and the inclusion of the quasi-control district, the present study does not constitute a scientific experiment in the strictest sense. There are no guarantees that there is not some unobserved confounder that is associated with air pollution and is not present in Park City. Controlling for absences from Park City provides a reasonable quasi-experimental approach to help control for unobserved factors.

To gauge the value of Park City as a control, it is insightful to look beyond simple geographic proximity and observe socio-economic and demographic characteristics. Given that Park City is meant to capture unobserved factors common across districts, the more similar Park City is to those districts on observed factors, the more plausible it is to assume that the

districts share such unobserved factors. Specifically, Park City appears better suited to act as a control for Alpine and Provo than for Salt Lake City. Perhaps the largest evidence for this is that the average percent of enrolled students who are absent on a given day is quite similar for these 3 districts, whereas for Salt Lake City, it is almost twice as high. Given that the average percent of students absent is similar across Alpine, Provo, and Park City, it is more reasonable to assume that the factors driving absences are more similar within these 3 districts than between Salt Lake City and Park City.

A potential cause for concern is the reduced correlation in temperature between Salt Lake City and Park City during pollution episodes. However, although the temperatures in Salt Lake and Park City are much less correlated during inversions, model 5 explicitly controls for these temperature differences by including Park City temperature in the estimation of Park City residuals and by controlling directly for Salt Lake City temperature.

Regarding socioeconomic factors, Alpine most closely resembles Park City in terms of median income and Provo most closely resembles Park City in terms of percent of adults older than 25 years who are high school graduates. In terms of percent of families below the poverty line, Provo and Salt Lake are more similar to Park City than Alpine, although it is noted that the percent of families below the poverty line in Park City has risen sharply from 9% in 2011 to 18% in 2013. These observations also provide some evidence supporting the assumption that factors influencing absences are more similar between Alpine/Provo and Park City than between Park City and Salt Lake, where median income and high school gradu-ation rates are much lower than Alpine, Provo, and Park City (note that the low median income and high percent below the poverty line in Provo is partially attributable the sizeable population of university students). Additionally, the fact that there are no 6th grade or prekindergarten students in the Park City district is a limitation for which none of the models can compensate.

Considering the results in light of these comparisons is enlightening. Estimated relationships in Alpine are the most consistent across the most relevant models (3, 4, and 5), and results in Provo are similarly steady. Relationships in Salt Lake City are fairly constant, but exhibit some unexplainable beha-vior for 1-day lagged exposures. Although we know of no specific mechanism for why this would be the case, these results may be evidence that those districts for which Park City is better suited to serve as a control yield the most reliable results. Additionally, because more students are absent on average in Salt Lake, the group of students on the margin to be influenced by pollution may be different than in the other districts. That is, if the pollution effects are in fact larger for longer lag structures, the negative previous day associations in Salt Lake may be evidence that the students likely to be influenced by this smaller pollution effect are already absent. However, none of these factors adequately or fully explain the curious Salt Lake City behavior. Again, it is emphasized that although the quasi-experimental design improves on a purely observational approach, it does not constitute a scientific experi-ment and therefore residual confounding remains a possibility.

Two other potential differences between the districts under consideration are the religious makeup of each district and the composition of $PM_{2.5}$ in each district. A majority of the people in Utah belong to the Church of Jesus Christ of Latter-day Saints; however, across counties, this percentage varies widely. The Alpine and Provo school districts are both located in Utah County,

where the percent of the population belonging to the Church of Jesus Christ is higher than in Salt Lake County, and much higher than in Park City, where only a minority are members of this church. However, we are unaware of any programs, activities, or related school functions for elementary aged children that differ by religious affiliation, and is unclear what effect, if any, these differences would have on absenteeism and how this might confound air pollution effects. Additionally, it is possible that the chemicals comprising $PM_{2.5}$ vary across the districts, and that the correlation between exposure and absenteeism is attributable only to some of these components; however, there is no reliable evidence that this would be the case.

Despite these limitations, the results of this article suggest that elevated levels of air pollution contribute to increased elementary school absences. These absences-pollution effects are relatively small compared to the effects of school year seasonal trends or day-of-week and holiday effects. These effects are also smaller than those previously observed in Utah County[10] when pollution levels were higher: they estimated that a 100 $\mu g/m^3$ increase in 28-day moving average PM_{10} was associated with a 40% increase in overall absences. Although the models are not directly comparable, our estimates indicate that a 100 $\mu g/m^3$ increase in 7-day moving average PM_{10} is associated with a 10% to 15% increase in absences. The smaller effect estimate may be the result of several factors: on average the percent of students absent has decreased somewhat since this earlier study; pollution exposures have decreased significantly over the period from 1985 to 2011, especially in Utah Valley; during the study period of the earlier study, the largest point source of $PM_{2.5}$ and PM_{10} was a large integrated steel mill that shut down in 2002. There is evidence the particulate pollution from the steel mill was relatively more toxic with potentially larger health effects.[24]

Given that some cities in the world experience $PM_{2.5}$ levels much higher than Utah's wintertime peaks,[25] the impact of pollution on school absences may be larger at higher levels of $PM_{2.5}$. Comparisons with this previous Utah county study also may indicate that absolute values of $PM_{2.5}$ matter more in determining school absences than do fluctuations from mean $PM_{2.5}$ levels. Nevertheless, associations between particulate air pollution and elementary school absences found in 1990 may still persist even at today's lower levels of pollution. The results also suggest that if pollution were reduced even more, perhaps to levels closer to those in Park City, the apparent effects of $PM_{2.5}$ on elementary school absences may completely vanish.

REFERENCES

1. Woodruff TJ, Grillo J, Schoendorf KC. The relationship between selected causes of postneonatal infant mortality and particulate air pollution in the United States. *Environ Health Perspect.* 1997;105:608–612.

2. Laden F, Schwartz J, Speizer FE, et al. Reduction in fine particulate air pollution and mortality: extended follow-up of the Harvard Six Cities study. *Am J Respir Crit Care Med.* 2006;173:667–672.

3. Dominici F, Peng RD, Bell ML, et al. Fine particulate air pollution and hospital admission for cardiovascular and respiratory diseases. *J Am Med Assoc.* 2006;295:1127–1134.

4. Brook RD, Rajagopalan S, Pope CA, et al. Particulate matter air pollution and cardiovascular disease: an update to the scientific statement from the American Heart Association. *Circulation.* 2010;121:2331–2378.

5. Hoek G, Krishnan RM, Beelen R, et al. Long-term pollution exposure and cardio-respiratory mortality: a review. *Environ Health.* 2013;12:43.

6. Gauderman WJ, Urman R, Avol E, et al. Association of improved air quality with lung development in children. *N Engl J Med.* 2015;372:905–913.

7. Currie J, Hanushek EA, Kahn EM, et al. Does pollution increase school absences? *Rev Econ Stat.* 2009;91:682–694.

8. Gilliland FD, Berhane K, Rappaport EB, et al. The effects of ambient air pollution on school absenteeism due to respiratory illnesses. *Epidemiology.* 2001;12:43–54.

9. Kim JJ. Ambient air pollution: health hazards to children. *Pediatr.* 2004;114:1699–1707.

10. Ransom MR, Pope CA. Elementary school absences and PM10 pollution in Utah Valley. *Environ Res.* 1992;58:204–219.

11. Ehrenberg RG, Ehrenberg RA, Rees DI, et al. School district leave policies, teacher absenteeism, and student achievement. *J Hum Resour.* 1991;26:72–105.

12. Easton JQ, Engelhard G. A longitudinal record of elementary school absence and its relationship to reading achievement. *J Educ Res.* 1982;75:269–274.

13. Neuzil KM, Hohlbein C, Zhu Y. Illness among schoolchildren during influenza season: effect on school absenteeism, parental absenteeism from work, and secondary illness in families. *Arch Pediatr Adolesc Med.* 2012;156:986–991.

14. Dominici F, Greenstone M, Sunstein CR. Particulate matter matters. *Science.* 2014;344:257–259.

15. Silcox GD, Kelly KE, Crosman ET, et al. Wintertime PM2.5 concentrations during persistent, multi-day cold-air pools in a mountain valley. *Atmos Environ.* 2012;46:17–24.

16. Whiteman CD, Hoch SW, Horel JD, et al. Relationship between particulate air pollution and meteorological variables in Utah's Salt Lake Valley. *Atmos Environ.* 2014;94:742–753.

17. Young JS, Whiteman CD. Laser ceilometer investigation of persistent wintertime cold-air pools in Utah's Salt Lake Valley. *J Appl Meteor.* 2015;54:752–765.

18. National Center for Education Statistics. 2015. Common Core of Data. Available at: https://nces.ed.gov/ccd/stnfis.asp. [Accessed July 2015].

19. United States Environmental Protection Agency. 2014. Download Daily Data. Available at: http://www.epa.gov/airquality/airdata/ad_data_daily.html. [Accessed May 2014].

20. National Climatic Data Center. 2015. Land-Based Station Data. Available at: http://www.ncdc.noaa.gov/data-access/land-based-station-data. [Accessed 10 April 2015].

21. American Community Survey 5-Year Estimates. 2009–2013. Available at: http://factfinder.census.gov/faces/nav/jsf/pages/index.xhtml. [Accessed October 8, 2015].

22. Total Population Survey. 2010. Available at: http://factfinder.census.gov/faces/nav/jsf/pages/index.xhtml. [Accessed October 8, 2015].

23. Utah Influenza Report. 2014–2015. Available at: http://health.utah.gov/epi/diseases/influenza/surveillance/2014_2015/influenza_season_summary.pdf. [Accessed January 15, 2016].

24. Ghio AJ. Biological effects of Utah Valley ambient air particles in humans: A review. *J Aeros Med.* 2004;17:157–164.

25. WHO's Ambient Air Pollution Database: Data summary of the AAP database. 2014. Available at: http://www.who.int/phe/health_topics/outdoorair/databases/AAP_database_results_2014.pdf?ua=1. [Accessed January 21, 2016].