THESIS

UNCONVENTIONAL OIL AND GAS DEVELOPMENT AND STUDENT PERFORMANCE ON STANDARDIZED TESTS

Submitted by

Gal Koss

Department of Agriculture and Resource Economics

In partial fulfillment of the requirements

For the Degree of Master of Science

Colorado State University

Fort Collins, Colorado

Spring 2021

Master's Committee:

Advisor: Jordan Suter
Co-Advisor: Dale Manning

Dan Zimmerle

ABSTRACT

UNCONVENTIONAL OIL AND GAS DEVELOPMENT AND

STUDENT PERFORMANCE ON STANDARDIZED TESTS

This paper evaluates the impacts of unconventional oil and gas extraction on academic achievement. The preparation, drilling, and fracturing of oil and gas wells has been found to create air and noise pollution—which can have negative effects on cognitive performance. Analyzing public school standardized test scores in Colorado, we find that additional drilling activity within 3 km of schools before tests decreases the portion of students who meet statewide standards by 0.75 percentage points, implying 1,857 fewer students met expectations statewide over the analysis period or 1.28% of all students who failed to meet expectations at treated schools. These findings impact how we view the scope of externalities from oil and gas development and informs the ongoing policy debate about minimum distance requirements between new wells and schools.

ACKNOWLEDGEMENTS

Funding for this research was provided in part by the Colorado State University Energy Institute. This paper benefited from conversations with Dan Zimmerle, Jude Bayham, Eric Kehoe, and Pete Maniloff.

TABLE OF CONTENTS

1 Introduction

Weak sustainability argues that drawing down natural capital can be sustainable if society's productive capacity is maintained through investment in other forms of capital stock, including human capital (Pearce and Atkinson, 1993). If depleting a natural capital also decreases human capital formation, however, it can lead to further decreases in an economy's productive capacity. Research suggests that effective education yields lifelong benefits. For example, better teachers can increase salaries and educational achievement for their students (Chetty et al., 2014) and early childhood intervention decreases the probability of experiencing poor health as an adult (Deming, 2009). From society's perspective, higher standardized test scores nationwide are associated with the formation of human capital and have been shown to increase national income and decrease infant mortality (Jamison et al., 2019). While rents from energy development increase local budgets which can be used to fund human capital formation (Bartik et al., 2016), the disruption caused by resource extraction can have negative local impacts (Ransom and Pope, 1992).

Beyond the trade-offs that arise from changing capital stocks, there are trade-offs in the activity of extracting resources. In the context of unconventional oil and gas development (UOGD, defined by the use of horizontal drilling and hydraulic fracturing), development can provide employment opportunities (Gittings and Roach, 2019; Agerton et al., 2017) and funding for public goods (Bartik et al., 2016), while the preparation of a well for production generates local externalities as operators produce particulate matter (PM), noise, truck traffic, and dust (Allshouse et al., 2019; Fedan, 2020; NYS DEC; Statewide Setbacks). The existing literature suggests that students' health and academic performance face risks from industrial activity and pollution (Ransom and Pope, 1992; Mejía et al., 2011) but research has not explicitly examined the connection between nearby UOGD and student academic achievement. Therefore, research is needed to evaluate whether there are costs to students' academic performance from UOGD activity.

This paper bridges the gap between the literature on the local effects of UOGD and the literature on air and noise pollution impacts on student performance. Research has found oil and gas development to impact student performance via school funding mechanisms (Marchand and Weber, 2015) and local economic changes to the job market (Zachary and Ratledge, 2017). While student performance has been shown to be impacted by air and noise pollution from nearby highways (Heissel et al., 2019) and by air pollution from nearby coal-fired power plants (Komisarow and Pakhtigian, 2020; Persico and Venator, 2018; Duque and Gilraine, 2020), there is a gap in the literature on the effect from nearby oil and gas activity.

This paper contributes by measuring the immediate impacts of UOGD activity on education outcomes, while controlling for longer term impacts. We measure the causal impact of UOGD activity on student academic test performance by examining the effect of well activity on the portion of students who meet Colorado Department of Education (CDE) expectations on standardized tests at all public schools in Colorado. The outcomes of this paper can inform the ongoing policy debate about the appropriate minimum distance from schools at which to allow drilling of new oil and gas wells.[1]

We use spatial and temporal variation in statewide UOGD from 2007-2019 to determine the impact of a well drilled near a public school in a time window around test dates on the portion of students from 3rd grade through high school who meet statewide expectations on standardized tests. Our rich dataset allows us to use fixed effects to control for many potentially confounding factors and identify a causal effect of UOGD activity on test performance. Further, by conditioning on school demographic composition, we control for the possibility that UOGD activity affects the characteristics of students in a school rather than the performance of a given student. During the time period of analysis, there was a sharp increase in oil and gas development that can be attributed to the cost effective use of both horizontal drilling and hydraulic fracturing to access otherwise inaccessible energy resources in shale formations (Weiner, 2014). We find that the wells drilled within 3 km of 145 schools in the 90 days leading up a typical test period decreased the portion of students who meet CDE expectations on each subject test by 0.75 percentage points. Applying

[1]For example, the Colorado Oil and Gas Conservation Commission voted in September 2020 for preliminary approval to increase the distance that wells must be set back from schools and homes from 1,000 to 2,000 feet (Kohler, 2020).

our estimates to the exams taken at treated schools, this implies 1,857 fewer students met expectations on a subject test, or 1.28% of all failures to meet expectations at treated schools. If a well were to be placed within 3 km of an average Colorado school, we would expect to see the average rate of students failing to meet expectations at that school increase from 10.3% to 11.0% — 5.81 additional fails across all subjects.

The rest of this paper is laid out as follows. The next section reviews relevant literature on education and on oil and gas impacts, followed by background on UOGD activity in Section 3 and an overview of the data used in Section 4. Section 5 describes the empirical specifications used to identify the impact of drilling activity on student performance, and Section 6 presents the results of those specifications, including a variety of robustness checks. Finally, section 7 discusses the results and concludes.

2 Literature Review

The research on education and natural resource use is largely focused on the connections between resource extraction and education at the national level (Gylfason, 2001; Cockx and Francken, 2016). By comparison, little research investigates how resource extraction itself impacts educational outcomes. This paper builds on existing literature that shows that student outcomes respond to shocks from noise, air pollution, and distraction in their immediate environment. Shield and Dockrell (2008) and Dockrell and Shield (2006) show that both acute and prolonged exposure to elevated noise levels lower performance on test scores. Similarly, several papers have shown that exposure to particulate matter levels can affect performance in umpires' decision making (Archsmith et al., 2018), on cognitive tests (Zhang et al., 2017), and on students' standardized tests (Roth, 2016; Marcotte, 2017; Ebenstein et al., 2016). Student absentee rates have also been linked to the mean PM levels of the trailing 28 days, and changes in PM levels are attributed to both seasonality and nearby industrial production (Ransom and Pope, 1992).

There is a growing body of research that uses exogenous changes to students' environments to identify the impact of these changes on student outcomes. Ransom and Pope (1992) used the changes in recorded PM levels from the closing (and re-opening) of a steel mill to identify the impact of PM on nearby school absences. Similarly, Clark et al. (2005) measure the aircraft noise at schools in a matched sampling approach, and find that increased noise exposure impairs reading comprehension. Heissel et al. (2019) show that attending a school downwind of a highway decreases student performance by comparing test scores of students who change schools. In Metcalfe et al. (2011) the researchers use the degree of overlap between test dates and English football tournaments in a difference in difference strategy for tournament and non-tournament years; they find a decrease of 0.12 standard deviations in test scores during tournament years, attributed to the increased level of distraction in those years. In this paper we use an identification strategy most similar to that in Persico and Venator (2018), in which the authors find that students at schools within one mile of industrial plants experience a decrease of 0.024 standard deviations in their test scores after a plant opens. The research in this paper is a novel contribution to the education literature as we attribute changes in student performance to the activity from UOGD.

A large part of the oil and gas literature in economics is dedicated to the impacts that UOGD has on nearby residents. In, Hill (2018), living within 2.5 km of a well is found to negatively impact infant health at birth. Muehlenbachs et al. (2015) find that development negatively affects homes that depend on groundwater in Pennsylvania, and Boslett et al. (2015) demonstrate that a moratorium on drilling in New York also decreased property values. Similarly, shale gas exploration activity has been found to have a short-term negative impact on nearby property values (Gopalakrishnan and Klaiber, 2014).

An additional line of research explores the positive economic impacts of UOGD on employment. The effects are primarily in the oil and gas industry, with negligible indirect or induced income impacts (Parades et al., 2015). Allcott and Keniston (2018) find that resource booms increase welfare, wages, and employment without long-run crowding out effects of resource booms on manufacturing. In Maniloff and Mastromonaco (2017), 550,000 jobs nationwide are attributed to the shale activity. While the boom was found to increase local wage income (Feyrer et al., 2017; Smith, 2020), in Alaska it was found to increase non-resident employment without affecting total resident employment (Guettabi and James, 2020).

There is also research that connects UOGD to certain educational outcomes, though clear nationwide trends do not emerge. For instance, development increased student populations in North Dakota, decreased student populations in Pennsylvania and West Virginia, and had no effect in Colorado (Zachary and Ratledge, 2017). School funding is generally found to respond positively to increases in oil and gas extraction, if at all (Bartik et al., 2016). In Colorado, local oil and gas property tax generates the majority of government revenue from oil and gas production– more than royalties from state and federal lands and severance taxes (Raimi and Newell, 2016). Nearly half of the revenue collected goes to local schools, and additional portions are put into school trust funds to insulate against future price swings. Changes to property taxes collected may not change total school funding, however, due to flexibility in the mill levy rate and the displacement of federal and state funding (Raimi and Newell, 2016). In the districts and years when UOGD might increase school revenue, it may also lead to decreases in student performance when the funds are not allocated to teacher salary or other pupil-related spending. In this case, rising wages in other industries can lead to increases in student absenteeism and teacher turnover (Marchand and Weber, 2015). Where other papers have used oil and gas availability in a region to identify regional education effects, we use UOGD activity at each well to identify changes in school-level education outcomes. This approach allows us to observe the localized and immediate effects that come from the preparation, drilling, and hydraulic fracturing of a well.

3 Institutional Background

In this section, we describe the steps involved in oil and gas development in detail to provide context for interpreting the impacts of UOGD activity that happens before, during, and after standardized test periods. Generally, preparing an oil and gas well for production requires three distinct (though potentially overlapping) steps. The focus of this research is on wells that are fractured, which account for the overwhelming majority of wells drilled during our study period. First, the site must be prepared, second, operators drill the well, and finally, fracturing service providers fracture the well. A more detailed discussion of the activities occurring during each of these phases provides motivation for when we expect to observe impacts on nearby schools.

Before an oil or gas well can be drilled, operators must prepare the site by constructing access roads and a well pad (the pad has a footprint of 2.2-5.5 acres during drilling) (NYS DEC). With the well pad built, 20-30 trucks (Pearson, 2020) arrive to bring one or more drill rigs over the course of 5-30 days; the rigs can be 40-150 feet tall (NYS DEC).

With the well pad prepared, operators move on to drilling, the second step of UOGD activity. Multi-well pad drilling has become more common since 2006 (Thuot, 2014) and operators may drill up to 28 wells on a single pad (Laramie Energy). From 2005-2014 operators increased their usage of highly mobile rigs to drill multiple wells (Mason, 2014; Langley, 2011). Drilling begins on the spud date and can take 7-14 days (Figure 1), as operators drill up to 10,000 feet vertically and then typically an additional 4,000-10,000 feet horizontally (Dunn, 2016). Most wells drilled in Colorado over the last decade are horizontal, which are more efficient to drill and can reach shale at farther distances (Dunn, 2017). After the well is drilled, it is necessary to cement and case it before hydraulic fracturing can begin. The total rig work of drilling, cementing, and casing typically lasts three to four weeks per well before the rig is moved off site (NYS DEC).

Figure 1: Timeline of a new multi-well pad.

The first day of drilling is often referred to as the spud date. Day 0 is the spud date of the first well drilled on the pad.

For many wells in Colorado, the process of hydraulic fracturing — injecting a fracking fluid composed of water, sand, and chemicals to release the oil and gas trapped in rock (Pearson, 2020) — is what makes them profitable. After wells are drilled, fracking service providers transport truckloads of sand, water, and chemicals to the well pad, in addition to the trucks needed to blend them together and to pump the fluid into the well bore (Dunn, 2016). The removal of drilling equipment and arrival of fracking equipment can take 30-60 days (Figure 1), and typically begins during the late stages of horizontal drilling (NYS DEC). The fracking process can take two to five days and up to 7.8 million gallons of water per well (NYS DEC). This is the equivalent of 24 acre feet– the same as the amount needed by 50 typical Colorado households annually (Waskom and Neibauer, 2014) or the average amount of water applied to 15 irrigated acres of farmland in Colorado in 2018 (USDA, 2019). Finally, over a period of a few days to weeks, the fracturing fluid flows back and the production of a well can begin (Goodman et al., 2016). Due to the size and quantity of the fracturing equipment, operators often use a single fracking rig to fracture multiple nearby wells sequentially — either on the same well pad or multiple well pads. This practice has become increasingly common since 2013. From the construction of a well pad to removal of flow-back fluid, a single-six-well pad requires as many as 6,500 truck visits (Goodman et al., 2016). Compared to lighter vehicles, the heavy duty trucks used with well drilling and fracking are associated with greater noise and PM emissions, in addition to disproportionate disturbances of traffic congestion (Goodman et al., 2016).

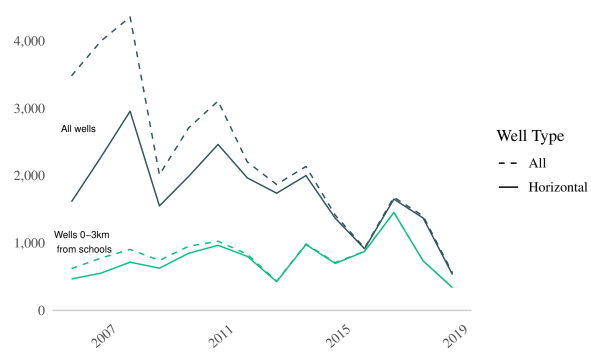
The Colorado Oil and Gas Conservation Commission (COGCC) implemented regulations in 2014 to set maximum daytime and nighttime noise levels at well sites, based on the types of land use in the surrounding area (Aesthetic and Noise Control Regulations). To comply with these limits, operators use sound barriers: commonly acoustic panels placed around the well pad, the drill rig, and/or other equipment at the well pad (Behrans and Associates). This change is intended to limit the externalities at a well, and might mitigate the treatment effect of well activity on student test performance.

Given what we know about the nature and timing of activities at well pads, we hypothesize that the biggest impacts on student performance will come from the activity after a well is drilled: hydraulic fracturing or the preparation for fracturing (Figure 1). Therefore, the model we use to identify the effects from well activity separates the activity, estimating distinct treatment effects for the activity before, during, and after well drilling.

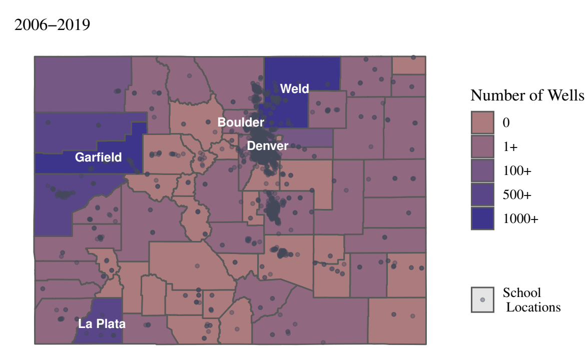
4 Data

To empirically explore the connection between UOGD and student performance, we use spatially and temporally explicit data on oil and gas well activity, annual test scores, school characteristics, and weather controls. The summary statistics for the data used are found in Tables 1 and 2.

Data on wells' spud dates (the first day of drilling) and locations are from the COGCC data (COGCC, 2020). We limit the data to the wells classified as active and to horizontal wells, using a secondary dataset (COGCC, 2020) for directional lines. Of the 31,827 wells drilled from 2006-2019, 25,311 are horizontal (Figure 2a), and 4,759 of these horizontal wells are within 3 km of 155 schools. Figure 2a shows the temporal distribution of all wells and wells near schools, and figure 2b shows the spatial distribution of schools and all horizontal wells. We use 3 km as the primary treatment distance following the findings of health impacts from both oil and gas well pollution at 2.5 km (Hill, 2018) and well noise at 2 km (Hays et al., 2016).



(a) Colorado wells drilled over time



(b) Colorado horizontal wells by county

Figure 2: Colorado wells drilled, by year and by county.

To account for operator behavior in drilling and fracturing multi-well pads, we cluster horizontal wells by their locations and spud dates. By clustering wells we account for the shared well-pad preparation and drilling preparation. As the drill rig runs in quick succession during the drilling of each well head at the well pad, we consider this to be continuous drill rig operation. Similarly, clustering wells allows us to account for the fracturing preparation and continuous fracturing rig operation that occurs during the fracturing of adjacent well-pads. Since fracturing happens after the well cluster (not individual well) is drilled, clustering provides a better indication for when different stages of production occur relative to observed spud dates. Similarly, it allows our analysis to accurately parse which UOGD activities affect student performance.

The clustering method creates clusters such that all wells in a cluster are drilled by the same operator and drilled within 2 km and 40 days of at least one other well in that cluster. Further, the method ensures that all wells in a cluster are not separated by a primary road. For each operators' wells, we create pairwise adjacency matrices of all horizontal wells' distances, days between spud dates, and primary roads between them. Filtering these matrices with critical values of 2 km, 40 days, and no primary roads (respectively), we construct a graph for each operator in which two wells share a cluster if all three criteria of distance, time, and roads between them are met (Jones et al., 2020). Clusters contain an unbounded number of such connections such that all wells in a cluster meet the criteria with at least one other well in that cluster. More information on the clustering methods can be found in Appendix B.

To merge the well cluster data with the school data, we intersect the layer of wells (points) with a layer of 3 km buffers around the centroid of each school (polygons). A cluster is considered to treat a school-year if one well is within 3 km of the school and the last well drilled in that cluster was drilled within one of our three defined impact windows: the 90 days leading up to, during, or the 90 days following the test period (more information on the clustering assignments can be found in Appendix B). We use the last well's spud date for treatment assignments since it is the date in the data closest to the start of fracking activity. Figure 3 shows how schools can have both treatment and non-treatment wells within 3 km. Using the institutional background (Figure 1), when a well cluster is drilled 0-90 days before the test, the test takers are most likely treated with horizontal drilling or fracturing activity; when it is drilled during the test, they are likely treated with vertical drilling activity, and when it is drilled 0-90 days after the test, they are likely treated with the activity of preparing the well pad.

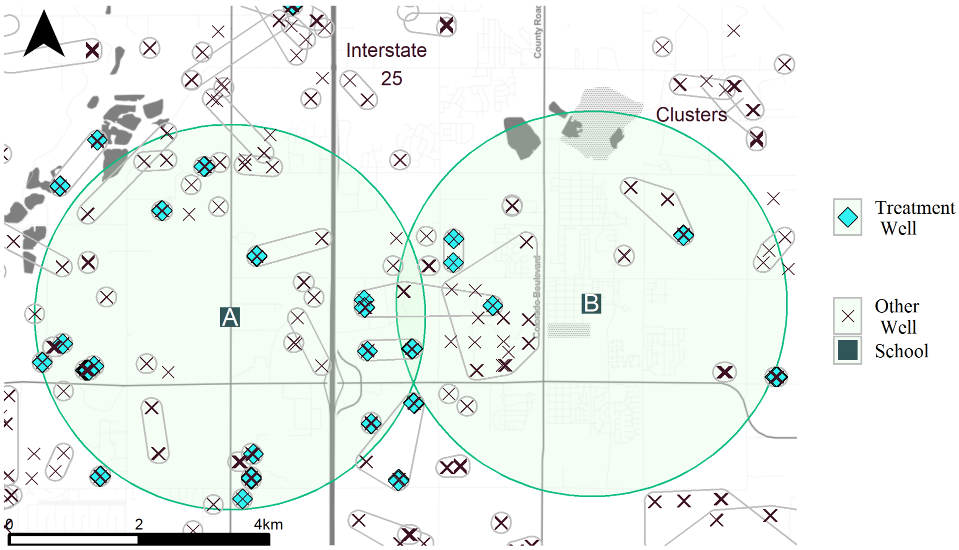


Figure 3: Wells and schools

This figure shows wells drilled near two schools ('A' and 'B') in Weld County. Each school is shown as the center of a green circle with a 3 km radius. Each well is marked by an 'X' and in a circle of its assigned cluster, with the wells drilled in the 90 days before, during, or the 90 days after a test period placed in a blue diamond. For a well to be a treatment of a school, it must be both temporally relevant (in a blue diamond) and spatially relevant (in a green circle).

There were between 24 and 68 schools treated each year with wells drilled within 3 km before, during, or after the test period (Table 1). From 2007-2019, the average school was treated with 0.27 wells and 0.05 clusters, with one school in Garfield County being treated in one year with 80 wells in 5 clusters in 2014 (Table 1).

Information on student test scores comes from the Colorado Department of Education based on the results of required standardized test scores taken annually in all public schools (CDE, a). While the content of the tests change each year, the CDE tests grades 3-8 in math and English and high school math each spring. The test score data used in this research are from all public schools (n=1,711) between 2007 and 2019. We use information on the test periods from the CDE, which are 10-32[2] day ranges provided each year, in which all schools are expected to administer their tests. Before the tests are administered, the CDE assigns scores to

[2] In the years 2007-2014, the 3^{rd} grade reading test windows were 10-11 days long and all other test windows were 32 days long. In 2015, the test window for all tests was 25 days long, and from 2016-2018 all test windows were 18 days long.

performance levels that reflect students' abilities to meet the subject standards for their grade. Our primary outcome of interest is the portion of students with test scores classified as meeting expectations, who can be thought of as able to answer "some" or "most of the test questions correctly, including the most challenging questions," compared to their peers who have "little success with the challenging content" (CDE Interpretive Guide; CDE Achievement Levels). Across grades, subjects, schools, and years, the mean of the portion of students that meet expectations is 0.9, with a standard deviation of 0.12 (Table 1).

Additionally, the CDE publishes school locations (CDPHE) and annual education statistics that describe the school demographics and overall classroom characteristics. These include the student mobility rate, the demographic composition, the annual student teacher ratio, and the portion of students eligible for free and reduced lunch at each school. Student mobility rate considers the portion of students who transfer schools, are expelled, drop out, are absent for extended periods, or are seriously ill or deceased (including 10 consecutive days of unexcused absences). In later years this portion is restricted to students who transfer schools, are expelled, or drop out. We also use the population density of schools' census tracts. This variable is the number of enrolled students in grades 1-8 in a given school's census tract (US Census, a) per tract area (US Census, b). Other CDE statistics of interest are annual district teacher turnover rates and the portion of students enrolled in a district from outside of that district (CDE, b).

Finally, we use estimates for total precipitation, mean daily wind speed, days with maximum temperatures below 10° Celsius, days with maximum temperatures above 25° Celsius, and days with dangerous relative humidity (greater than 85 or less than 5%) from GridMET (Climatology Lab). The estimates are daily gridded on a 4 km x 4 km grid — we pair each school with its nearest grid centroid and summarize over the days of each test period (summing precipitation, cold days, hot days, and dangerous humidity days and averaging mean daily wind speed).

5 Empirical Specification

5.1 Main Specification

The primary empirical specification identifies the effect of UOGD activity before, during, and after a test on the portion of students who meet CDE expectations for a math or English spring standardized test. The dependent variable is the portion of students who meet CDE expectations, and we include controls for existing wells along with demographic controls, X_{1it} and weather controls, X_{2it} at school i in year t,

$$\begin{aligned} PortionMet_{ijkt} = &\beta_{W_1} W_{1it} + \beta_{W_2} W_{2it} + \beta_{W_3} W_{3it} + \beta_{WC_1} WC_{1it} + \beta_{WC_2} WC_{2it} + \\ &\beta_{X_1} X_{1it} + \beta_{X_2} X_{2it} + \alpha_i + \gamma_{jkt} + \varepsilon_{ijkt}. \end{aligned} \tag{1}$$

The treatment variables are the well clusters drilled 0-90 days before (W_1), during (W_2), and 0-90 days after (W_3) the test period, within 3 km of school i in year t. In addition to fixed effects for school i (α_i) and grade j and test subject k in year t (γ_{jkt}), we include two controls for existing wells within 3 km of schools in year t: a variable for the wells drilled in the ten years leading up to the first day of school in a given year (WC_1), and a variable for the wells drilled from the start of the school year to 90 days before the test period (WC_2). The three treatment variables and two well controls all use the same clustering definitions for horizontal wells within 3 km, and contain non-overlapping counts of horizontal wells drilled from 10 years before the school year to 90 days after the test period for each observation. We consider two measures of treatments by wells and well controls. In the first ('dummy') specification, each treatment variable and well control receives a value of 1 if there is a well cluster within 3 km during the time period, and a 0 if not; in the second ('count') specification the treatment variables and well controls are counts of the well clusters drilled before, during, and after the subject k test period for grade j in year t within 3 km of school i.

The school FE controls for district and school-level characteristics that are assumed time-invariant, such as school demographics, funding, and teacher and facility quality. This complements the time-variant school demographic controls explicitly included in X_{1it}. We include controls for the annual school-level student mobility rate, portion of non-white students, student teacher ratio, and portion of students eligible for free or reduced lunch, as well as the annual district-level teacher turnover rate and portion of students from out of district. The grade-subject-year FE controls for the test characteristics, such as the specific content and scoring for each grade and subject test annually. Given that the cut scores that define performance levels are determined annually for each test by grade and subject, this FE controls for the variation in the portion of students who meet expectations due to changes in scoring. The yearly component of this FE also accounts for state-wide time-variant characteristics such as regional trends in curricula, culture, large-scale weather events, or changes in the broader economy that could influence scores. Finally, we include weather controls for the test periods' total precipitation, mean daily wind speed, days with maximum temperatures below 10° Celsius, days with maximum temperatures above 25° Celsius, and days with dangerous relative humidity in X_{2it}.

The identification of β_{W_1}, β_{W_2}, and β_{W_3} in specification (1) relies on conditioning test score outcomes from well activity on the FE and the assumption of exogenous assignment of treatments. Figure 3 shows how the study design allows schools not treated in a certain year to act as controls for the years in which it is treated. For instance, school *A* in figure 3 is treated with well cluster activity in 2006 and 2008, and a baseline is found in 2007; school *B* is untreated in 2011, acting as a control for the treated years 2010 and 2012.

We assume both that residential sorting is uncommon and that a portion of the sorting effects are captured in the demographic controls. Households likely do not sort out of an area in the months between anticipating well development and the student test period affected by the activity; however, they may relocate in the long run in response to repeated UOGD activity or changes in school test performance. The demographic controls (student-teacher ratio, portion of students from out of district, student mobility rate, and portions of students not white and eligible for free or reduced lunch) capture much of the variation in school demographics that would result from residential sorting. Finally, our identification strategy is not threatened by households sorting out of the area due generally to fracking activity in that area over long time periods.

To more fully investigate the spatial impacts of UOGD activity, we include an additional set of specifications in which we change the treatment variables. Specifically, we map wells to schools with integer radii that vary from 2-7 km, and include the same controls. Similarly, in another model we split each treatment variable into 7 concentric 1 km wide radii of wells within 7 km of school i during the administering of subject test k to grade j in year t.

5.2 Heterogeneous Effects

We estimate two specifications designed to measure heterogeneity of impact based on the density of the population at the school and the school demographics. For school demographics heterogeneity, we are interested in comparing the well activity impact on performance at schools with varying portions of students non-white and eligible for free and reduced lunch. Examining the effects by population density is designed to provide evidence on whether UOGD activity is primarily treating students at school or at home. We expect that at schools in high density areas the effects would be larger if UOGD activity impacts students both at school and at home. An insignificant interaction would suggest that effects exist at all densities and that the effect is driven by where students attend school, not where they live.

$$\begin{aligned}
PortionMet_{ijkt} = &\sum_{d=1}^{3} \beta_{W_d} W_{dit} + \sum_{d=1}^{2} \beta_{WC_d} WC_{dit} + \\
&\sum_{d=1}^{3} \beta_{W_{5+d}} W_{dit} * Z_{it} + \sum_{d=1}^{2} \beta_{WC_{8+d}} WC_{dit} * Z_{it} + \\
&\beta_{X_1} X_{1it} + \beta_{X_2} X_{2it} + \alpha_i + \gamma_{jkt} + \varepsilon_{ijkt}.
\end{aligned} \tag{2}$$

Using the dummy version of well treatment variables $W_{1,2,3}$ and control variables $WC_{1,2}$ for wells within 3 km, $\beta_{W_{6,7,8}}$ and $\beta_{WC_{9,10}}$ show the marginal change in the impacts of well cluster activity with increases in Z_{it}, the population of students in grades 1-8 per square km in school i's census tract in year t.

For the effect by school demographics, we interact each treatment variable (dummies for well treatment and control clusters within 3 km) with the index variable we construct for heterogeneous impacts. Specifically, Z_{it} is an index of the portion of nonwhite students and portion of students eligible for free and reduced lunch. These highly correlated variables (0.81) reflecting distinct but related sources of performance gaps can be condensed into a single socioeconomic index (Shogren et al., 2018). The variable is a score gap index constructed as the arithmetic mean of the respective school-level indices for the portion of nonwhite students and portion of students eligible for free and reduced lunch.

We estimate three additional models to measure the heterogeneity of UOGD activity by grade, by subject test, and for top performers. By interacting each independent variable in the main specification with a dummy variable for elementary school, we see how drilling activity affects students in different grades. Similarly, by interacting with a dummy variable for math, our results show the extent to which the impacts of UOGD vary by subject test. Finally, to estimate the results of UOGD activity on top performers we use the portion of students in the top performance level by school, grade, subject, and year as the dependent variable.[3]

5.3 Robustness

By changing the clustering criteria, we are able to see the robustness of our results to the clustering choices. We estimate additional models with treatment variables for wells un-clustered and clustered at 200m and 15 days (i.e. using 200m and 15 days as the critical values to filter the well adjacency matrix inputs for the graph), in addition to the main clustering choice of 2 km and 40 days.

[3]Note that the portion of students in the top performance level, like the portion of students who meet CDE expectations, is not consistent since the cut scores that define performance levels are determined annually for each test by grade and subject. As before, this is controlled for with a subject-grade-year FE.

We also estimate a model in which the only treatment clusters are those not impacted by changes to clustering choices. Comparing the primary well clustering (critical values of 40 days and 2 km) to alternate well clustering (all other combinations of 15, 40, and 65 days and 0.2, 1, and 2 km), we keep the wells in clusters where, over the 9 different clustering methods, the number of wells did not vary. This reduced the total horizontal wells from 1995-2019 in Colorado from 14,338 to 4,780 with a mean cluster size of down from 25.65 to 6.93 wells per cluster.

Similarly, we determine the temporal robustness of the results by splitting the cluster impact time windows into 45 day intervals and extending the range from 90 to 180 days before the start and after the end of the test period,

$$\begin{aligned} PortionMet_{ijkt} = &\sum_{r=1}^{4} \beta_{W_r} W_{1itr} + \beta_{W_5} W_{5it} + \sum_{r=1}^{4} \beta_{W_{5+r}} W_{3itr} + \\ &\beta_{WC_1} WC_{1it} + \beta_{WC_2} WC_{2it} + \beta_{X_1} X_{1it} + \beta_{X_2} X_{2it} + \alpha_i + \gamma_{jkt} + \varepsilon_{ijkt} \end{aligned} \tag{3}$$

where W_{1itr} and W_{3itr} are the wells drilled within 3 km of school i, from $45(r-1)+1$ to $45(r)$ days before the start and after the end of the test period, respectively. To accommodate the extended treatment windows, the variable for wells drilled during the school year is adjusted to include only the wells drilled from the first day of school until 180 days before the test period. The variable for wells drilled during the test period, W_{5it}, is unchanged, as are the controls for wells drilled before the school year, demographics, and weather.

6 Results

6.1 Main Specification

Using OLS to estimate our main specification, having horizontal wells drilled before the test period is estimated to decrease the portion of students who meet CDE expectations on a test by 0.0075 for the sample that includes all schools, and 0.0066 for a restricted sample that only includes treated schools (Table 3, dummy specification). The similarity suggests that the never-treated schools used to estimate other model coefficients do not change the treatment effect in a meaningful way. Each point estimate presented can be interpreted as the independent variable's impact on the portion of students who meet expectations: for instance having a well drilled in the 90 days before the test period reduces the portion of students who met expectations by 0.0075. The marginal well cluster (Table 3, count specification) drilled before the test period decreases the portion of students who meet expectations by 0.0045 and 0.0045 in treated schools.

Since it is the coefficient on the indicator for the 90-day window before the test period that is consistently significant, these results imply that students are affected by UOGD activity occurring in the 90 days after the last well in a cluster is drilled. This is likely during horizontal drilling or while the well is undergoing preparation for fracturing and fracturing activity. Anecdotally, we know that some teachers dedicate the weeks leading up to the test period to practicing for the test. The effect we find may also be driven by distraction during this preparation period from drilling activity.

Since our dependent variable for the portion of students who meet CDE expectations is constrained between 0 and 1, a fractional logit regression is also appropriate. The results from estimating our primary model specification with fractional logit are qualitatively similar to the results from estimating with OLS (Appendix A Table A1). All coefficient estimates have identical directions and are highly similar in magnitude and statistical significance. Due to our identification strategy's reliance on FEs, the comparative stability of OLS estimation with FEs, and the similarity of our main results estimated in OLS and fractional logit, the remaining results are estimated with OLS.

Applying our estimates to the exams taken at treated school-years, this implies post-drilling UOGD activity caused 1,857 fewer tests to meet expectations on a subject test over the study period, or 1.28% of tests that do not meet CDE expectations at treated schools. From 2007 to 2019, there were 247,651 subject tests that were treated with drilling and fracking activity and 144,599 failed to meet CDE expectations. With a 0.75 percentage point decrease in expectations met from UOGD activity, 1,857 of the subject tests were failed due to drilling and fracking (1.28% of all fails). If a well were to be placed within 3 km of an average Colorado school, we would expect to see the average rate of tests not meeting expectations at that school increase from 10.3% to 11.0% — 5.81 additional subject tests on which students do not meet expectations. The mean of total annual subject tests taken per school is 775, with an average failure rate of 10.3%, or 79.4 tests. Introducing drilling and fracking activity to this mean school-year would increase the failure rate to 11.0%, or 85.2 tests.

In the same specifications (Table 3) we see no significant effect from wells drilled during or after the test period. We interpret these findings to mean that students are not affected by drilling or pre-drilling activities. Similarly, looking at the insignificance of the coefficients for wells drilled that school year and in the previous ten years (*During School Year* and *Before School Year*), activities related to on-going oil and gas production have limited impacts on performance. Since wells drilled outside the test window could affect the learning process, a corollary of these results is that our findings primarily reflect a negative effect on student performance on standardized tests — more so than an effect on students' learning. The insignificant coefficients reflect that when controlling for UOGD activity contemporaneous with a test period, neither variation in having older wells or in the count of older wells explains the variation in the portion of students who meet CDE expectations on a subject test.

The included controls have intuitive results (Appendix A Table A2) consistent with the existing literature: schools having higher portions of students non-white and eligible for free and reduced lunch is correlated with fewer students meeting expectations on tests (Shogren et al., 2018). Similarly, we find that the portion of students from out of district is positively correlated with performance, which we interpret as an indicator of school desirability. This is in line with the findings of Brummet (2014) that students' performance can improve when they transfer to a better school. Finally, schools have worse performance with increases in teacher turnover (Ronfeldt et al., 2013) and student mobility (Wright, 1999), and our weather controls' results align with the findings in Cook and Heyes (2020), Zivin et al. (2018), and Park (2016): both cold and hot outside temperatures can negatively affect performance.

We can make internal comparisons between the magnitudes of the negative impacts of UOGD activity before the test period and of demographic and weather controls. The teacher turnover rate and days during the test period with maximum temperatures above 25° C (77° F)— with point estimates of -0.039 and -0.0015 significant at the 0.01 level (Appendix A Table A2)— are conditionally correlated with 9,658 and 371 failures to meet expectations on standardized tests from 2007-2019 at treated schools. The 1,857 tests that did not meet CDE expectations over the study period due to oil and gas well drilling before the test periods indicate that being treated with post-drilling activity has a magnitude equivalent to a 0.19 percentage point increase in teacher turnover rate, or 5 additional hot days during a given test period.

The results presented in Table 4 provide coefficient estimates for the same model specification, but varies the distance at which treatment wells are defined. We find that at 2 km, 3 km, and 4 km there is a large and significant effect on the portion of students who fail to meet CDE expectations due to UOGD post-drilling activity (Table 4). Having wells within 2 km decreases the portion of subject tests that meet expectations at a school by 0.0107 (dummy specification, 0-2 km), with additional clusters having an effect of -0.0085 (count specification, 0-2 km). These effects decrease monotonically in economic and statistical significance as the radius is expanded: at 4 km the effect is -0.0046 and -0.0029 for having wells (dummy) and the marginal cluster (count). The trend continues, and when considering well clusters within 7 km there is no detectable effect on the portion of students who do not meet expectations. Using a model that defines treatments using concentric radii, we see that the effect on students can primarily be attributed to wells drilled 1-2 km from schools (Appendix A Table A3). Having a decrease in significance with greater radii aligns with intuition and our findings: wells closer to schools have a bigger impact on student performance than wells farther away. One might expect to see the largest effect for wells within 1 km but due to the geometric nature of radii and concentric radii there are too few treated schools to see an effect. In all models with radii from 2-7 km and the model with concentric radii there are persistent negative and insignificant coefficients for existing wells. This suggests that there may be a cumulative negative effect of activity or ongoing production, but such an effect is insignificant when accounting for the UOGD activity at the time of test periods.

6.2 Heterogeneous Effects

To understand the extent to which impacts of UOGD vary, we present the results of the model that explicitly accounts for population density. We do not see a significant difference in impacts of UOGD activity at schools with a higher density of enrolled students in the census tract in which the school is located (Figure 4), providing evidence that our main results represent the impact of UOGD activity on students at school, not at home.

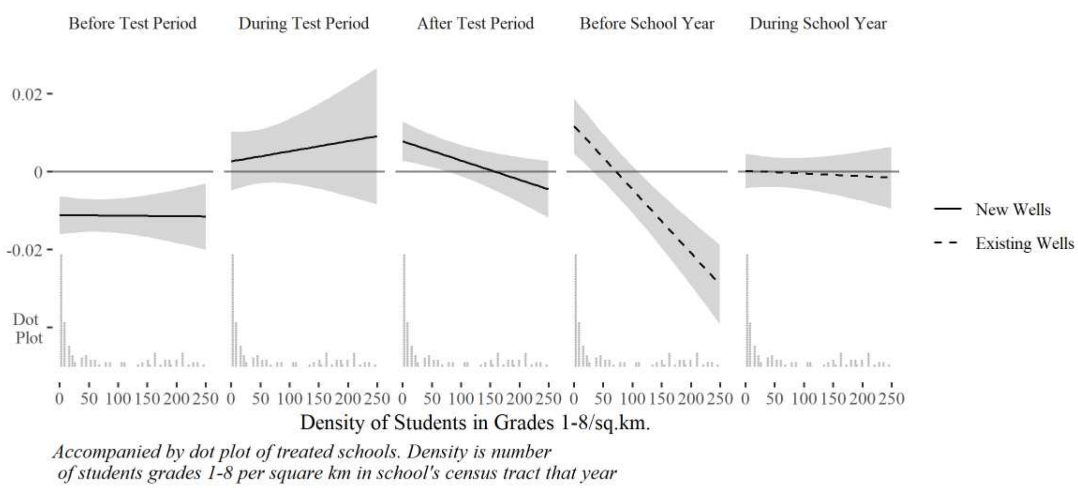


Figure 4: Heterogeneous impacts of well activity within 3 km of schools on portion students that meet expectations: by population density. (*Regression results in Appendix A Table A4*)

We also find that existing wells (those drilled in the ten years leading up to the first day of school) have an increasingly negative impact on test performance as student population density increases. This result suggests that ongoing well production may impact students at home. Similarly, it suggests that there may be positive funding effects: the benefit from tax revenue from existing wells may counteract the negative externalities near schools with low population densities.

Looking at the results of the model exploring the variation in treatment impacts by the score gap index — constructed from the indices of the portion of students not white and the portion eligible for free and reduced lunch — we see heterogeneous impacts by school demographics (the left panel in Figure 5). For schools in the 3rd quartile, being treated with post-drilling activity decreases the portion of subject tests that meet expectations by 0.0115. Compared to the students at schools in the 2nd quartile, who experience a 0.0068 decrease in the portion of tests meeting expectations (the median score gap index is 0.37). We bootstrap to find that the difference between the two marginal effects is significant ($p < 0.05$).

Over the study period, schools in the 2nd and 3rd quartiles of the score gap index reported 25,205 fails out of 76,696 tests (meeting expectations at a rate of 0.67) and 39,274 fails out of 62,014 (meeting expectations at a rate of 0.37), respectively. Applying our estimates for the post-drilling impact on performance in each quartile, we attribute 524.7 fails to drilling and fracking in the 2nd quartile, and 712.54 fails to drilling and fracking in the 3rd quartile. In the schools with a greater score gap index, 2.08% of test results that did not meet expectations were due to UOGD activity compared to 1.81% (i.e., 187.84 fewer UOGD-caused fails over the study period) in schools with a lower score gap index. We see a greater negative effect of drilling and fracking on student performance in schools with fewer students meeting CDE expectations on subject tests.

The heterogeneous impact found along the score gap index can be explained by two potential mechanisms. It could be the case that at schools with higher portions of students non-white and eligible for free and reduced lunch, the observed distribution of scores has a lower mean (but similar variance) and there are

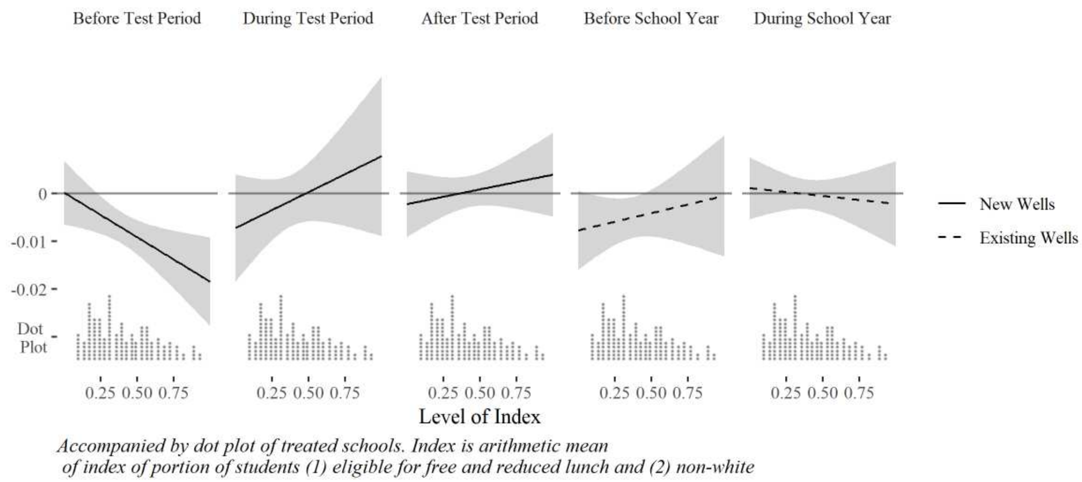


Figure 5: Heterogeneous impacts of well activity within 3 km of schools on portion students that meet expectations: by score gap index. (*Regression results in Appendix A Table A5*)

more students close to falling into the bottom statewide performance level. When these schools are treated with drilling and fracking activity, the students who were close to not meeting expectations are more likely to be pushed into that category. As a result, the non-attainment rate increases more than at schools with a lower score gap index.

Alternatively, students' performance at schools with a higher score gap index may be more susceptible to drilling externalities: for instance parental compensatory and avoidance behavior may be a function of race and income. Given the data, we cannot determine which of these two mechanisms drives the heterogeneity in impacts across score gaps.

The results from the models that investigate heterogeneity in impacts by grade and subject show that post-drilling UOGD activity has a negative impact on older students and English tests, respectively (Tables 5, 6). The coefficients for the impact of wells drilled before the test period on students in 6^{th} grade and older and on elementary school students are significantly different ($p < 0.01$ in the dummy specification, $p < 0.1$ in the count specification). The subject test finding is consistent with the negative effect of air pollution that has been found to impact verbal test scores (Zhang et al., 2017) and of noise from individual external events — like passing trucks — on reading ability (Shield and Dockrell, 2008). Similarly, the age finding is consistent to prior research on the effect of noise on children's test scores: while younger children are susceptible to general background noise, older children have been found to be more aware of and affected by individual external noise events (Shield and Dockrell, 2008).

Finally, UOGD activity within 3 km of schools is not found to have significant effects on students in the top performance level (Table 7). The results do reveal negative and insignificant treatment coefficients when treatment wells are defined by radii greater than 3km. This is suggestive of an intuitive negative effect on top performance, but more research is needed to more precisely identify such an effect.

6.3 Robustness

The results presented so far have used clustered wells as treatment variables, created so that every well in a cluster is drilled within 40 days and 2 km (unobstructed by a primary road) of at least one other well in that cluster. We find that clustering wells is both natural to the facts and aids in interpreting the effects. Overall, there are negligible differences across the dummy specifications with respect to the impact of having UOGD before the test period (Table 8). Without clustering and with wells clustered with 2 km and 40 days as critical values, we see a non-intuitive positive result for post-drilling activity in the count specification (Table 8). Since we consider a cluster to treat a school only if the last well drilled in that cluster has a spud date during an impact window, the fracking-level clustering decision (used in all other models) removes wells from the treatments that led to positive coefficients at narrower clustering levels. The clustering approaches generate parameter estimates that are largely similar to each other, suggesting that while clustering the treatment variables is important, the clustering critical values are less so. Importantly, our qualitative results related to the impact of post-drilling UOGD activity remains both with and without the clustering of wells in the dummy specification.

We confirm this finding in the results from the model in which the only treatment well clusters are the clusters not affected by clustering critical values (Table 9). Compared to the primary well clustering, the effect of post-drilling activity at only stationary well clusters is similar in magnitude and significance. This provides further evidence that well clustering captures operator behavior, and the results are not sensitive to the critical values used in clustering.

By splitting the impact windows into 45-day periods and including the windows up to 180 days before and after the test period, we see that the effects for post-drilling activity are concentrated 46-90 days after the spud date of the last well in a cluster (Table 10). There are no other statistically significant effects before the test period — including from existing wells. These results add clarity to the timing and type of UOGD activities that impacts students' testing performance; given the days between the spud date and the effect on students, it appears that hydraulic fracturing activity decreases the portion of students who meet CDE expectations on subject tests. Hydraulic fracturing contains several activities that would disrupt and distract students, such as preparation for fracturing, fracturing, and flow-back. Therefore, these results are intuitive. Future research is needed to further identify which of these activities drives the negative effects on test performance.

7 Discussion and Conclusion

We find that UOGD activity occurring after the spud date has a negative effect on student performance on standardized tests in Colorado public schools from 2007-2019. Our findings suggest that the horizontal drilling and the preparation for fracking, fracking, or flow-back activities occurring within 3 km of schools caused 1,857 additional failures to meet expectations on math and English standardized tests, accounting for 1.28% of all such fails at treated schools. This research contributes to the economics literature on both education and drilling in its novel attribution of test scores to nearby UOGD activity. The economics literature on education has linked industrial activity to student performance and the literature on oil and gas development has linked UOGD to student performance through funding, but this paper is the first to link well activity to student performance.

The estimated impacts we find are consistent with existing estimates in the literature for the effects of air and noise pollution and distraction on student performance on standardized tests. For the years 2015-2019 and class sizes greater than 16 students, the test score data include the mean scaled score, in which each observation is the mean of students' scaled scores (from 650-850) for each subject test by grade, school, and year. The mean scaled scores available are approximately normally distributed (Appendix A Figure A1) and at treated schools mean scaled scores have a mean of 736.14 and standard deviation of 16.37. At those schools in all years, the portion of students who did not meet CDE expectations is 0.1139. Had those schools not been treated with post-drilling UOGD activity, the econometric results imply that 1,857 fewer students would have failed and the portion of students who did not meet expectations in a counterfactual distribution would have been 0.1065. We can use the corresponding z-score for those points in the CDF if we assume normality in the distribution of students' individual scores (Appendix A Figure A2), similar trends at the student-level and state-level, and representativeness of observed test scores. The z-scores for the portion of students who fail to meet expectations are -1.21 in the data, and -1.25 in the counterfactual data in which schools are not treated.

The negative impact of 0.04 standard deviations from being treated with UOGD post-drilling activity is consistent with estimates in related literature. For example, Heissel et al. (2019) find that attending a school downwind of a highway is also associated with a decrease in test scores of 0.04 standard deviations. Similarly, Persico and Venator (2018) find that being exposed to air pollution from toxic release inventory sites is associated with a decrease in test scores of 0.024 standard deviations. Using the Chetty et al. (2014) finding that test scores increase by 0.1 standard deviations with an increase of one standard deviation in teacher quality, we estimate that schools treated with UOGD activity before the test window would need to improve teacher quality by 0.4 standard deviations to achieve the same test scores as an equivalent untreated school.

UOGD activity occurring 0-90 days after the spud date is marked by higher PM levels, noise levels (Allshouse et al., 2019), and traffic (Goodman et al., 2016). The impacts estimated in this research show that this activity causes fewer students to meet expectations, an effect that can be understood as having at least one of two parts. The air pollution, noise pollution, and traffic might affect all students, causing all students to perform worse on spring standardized tests. Given the negative but insignificant impacts on top performers, it is also possible that UOGD activity primarily impacts students who were in the lower distribution of test scores. This question can be answered with student-level data, and is an opportunity for future research.

For students, having lower test scores can have lasting effects, even if those test scores do not reflect having learned less. In their analysis of teacher value-added, Chetty et al. (2014) point out that high achieving students can track into classes with higher value-added teachers in their findings that lower test scores can be correlated with decreases both in the probability that they go to college and in their lifetime earnings. In this research, we find that UOGD activity causes fewer students to meet expectations, which can lead to students not tracking into classrooms with higher value-added teachers. Further, we expect students at treated schools to be treated repeatedly by drilling and fracking, which suggests that their long-run learning outcomes can be negatively affected.

Schools are similarly impacted by lower test scores. We find that nearby UOGD activity causes lower test scores, adding to the known negative effects of nearby highways, industrial activity, noise, and pollution. Schools are not fully in control of student performance on standardized test scores— their locations also impact student performance and educational outcomes. In turn, worse school-wide performance on standardized tests may make them eligible for support funding or more drastic measures from the local board of education (CDE Assessment FAQ, 2021). That schools face financial repercussions due to their location has equity implications when considering our findings of heterogeneous impacts of drilling and fracking across the existing racial and socioeconomic score gap.

Inferences associated with the presented results are limited to the statewide institutions, mitigation measures, and drilling operations found in the study area — UOGD in other states that have different regulations or in oil and gas geology with other drilling procedures may have an effect that varies in magnitude or timing from the effect we find in Colorado. This research can inform future studies to examine the efficacy of mitigation policies. Additionally, future research could estimate the mechanisms of the impact related to funding, air pollution, noise, and traffic from UOGD activity. There are also teacher quality questions that can be studied, especially in teachers' school choices that are impacted by UOGD activity and resulting education outcomes.

To address the welfare implications of UOGD activity's negative impact on test scores, future research would need more granular data on how students, schools, and teachers are affected. Our econometric results suggest that there are negative impacts on test scores. Estimating the welfare costs of these negative impacts would benefit from data that allows for an understanding of how reduced test scores impact students' future earnings, schools' funding, and teach salaries. With student-level data we could apply the estimates from Chetty et al. (2014) to measure how lifetime earnings are affected from UOGD activity. With school and teacher data, we could estimate how enrollment, funding, and salaries respond to changes in test scores.

In addition to measures of social costs, the welfare implications would also depend on the social benefits to well operators from drilling wells near schools during the 90 days leading up to the test window. If we knew the costs of changing UOGD procedures to drill away from schools and to drill at a time away from the test window, we could estimate the lost benefits from policies that move wells farther from schools or restrict drilling new wells in the 90 days leading up to the test window.

The results of this paper add motivation to the policy debate to increase the minimum set-back distance new wells can be drilled from schools. The distance was increased from 350-1,000 feet (0.1-0.3 km) in 2014, and at time of writing, Colorado State Legislature is preparing to vote to increase the distance to 2,000 feet (0.6 km). We find a negative effect far beyond that boundary, adding to the known health impacts of drilling. Therefore, a social cost estimate of drilling that uses only health impacts underestimates the total costs. The decrease in test scores could impact school demand and nearby housing prices, and if lower test scores lead to worse educational outcomes, there may be an impact to students' lifetime earnings. Ultimately, the decisions on oil and gas well setbacks and mitigation measures depend on the net social benefits and costs of UOGD — this study finds an additional source of costs to be considered.

Tables

Table 1: Summary statistics; all schools

Variable	Min	Max	Mean	SD	N	Varies by
Students Met Expectations (Portion Students)	0	1	0.9	0.12	158281	Test & Year
Max. Temp. Below 10° C (days)	0	33	6.64	5.27	158281	Test Window & Year[4]
Max. Temp. Above 25 C (days)	0	22	1.16	1.84	158281	Test Window & Year
Total Precipitation (mm)	0	154.8	26.96	17.25	158281	Test Window & Year
Harmful Humidity (days)	0	30	7.12	4.46	158281	Test Window & Year
Mean Daily Wind Speed (m/s)	1.73	8.67	4.05	0.67	158281	Test Window & Year
Students Not White (Portion Students)	0.01	1	0.43	0.26	19084	School & Year
Mobility Rate (Portion Students)	0	1	0.22	0.13	19084	School & Year
Student to Teacher Ratio	2.26	628.33	17.38	11.87	19084	School & Year
Free and Reduced Lunch Eligible (Portion Students)	0	1	0.43	0.27	19084	School & Year
Student Density (Students/Sq.Km)	0.01	1252.02	124.44	135.39	13470	School & Year (2010-2018)
Teacher Turnover Rate (Portion of Teachers)	0	0.96	0.17	0.08	1941	District & Year
Out of District Students (Portion of Students)	0	1	0.13	0.18	1941	District & Year

Table 2: Summary statistics; treated schools

Variable	Min	Max	Mean	SD	N	Varies by
Yearly Schools Treated w. Wells 0-3 km	24	68	45.31	12.9	13	School & Year (Summed Yearly)[5]
Treatment Wells (Within 3 km)	0	80	0.27	2.25	158281	Test & Year[6]
Treatment Clusters (Within 3 km)	0	5	0.05	0.3	158281	Test & Year[7]
Students Met Expectations (Portion Students)	0.03	1	0.89	0.11	13119	Test & Year
Max. Temp. Below 10 C (days)	0	26	5.31	3.94	13119	Test Window & Year
Max. Temp. Above 25 C (days)	0	8	1.66	1.76	13119	Test Window & Year
Total Precipitation (mm)	0	80.6	23.44	14.48	13119	Test Window & Year
Harmful Humidity (days)	0	24	8.29	4.37	13119	Test Window & Year
Mean Daily Wind Speed (m/s)	2.27	6.74	3.68	0.57	13119	Test Window & Year
Students Not White (Portion Students)	0.02	0.97	0.42	0.22	1584	School & Year
Mobility Rate (Portion Students)	0	0.86	0.20	0.11	1584	School & Year
Student to Teacher Ratio	7.7	395	18.05	12.98	1584	School & Year
Free and Reduced Lunch Eligible (Portion Students)	0	0.98	0.42	0.23	1584	School & Year
Student Density (Students/Sq.Km)	0.07	683.28	80.38	106.6	1132	School & Year (2010-2018)
Teacher Turnover Rate (Portion of Teachers)	0.04	0.33	0.16	0.06	268	District & Year
Out of District Students (Portion of Students)	0	0.48	0.07	0.08	268	District & Year

[4]For all weather variables, we aggregate over only the days that fall within the years' test windows
[5]Schools treated with in the 90 days leading up to, during, or in the 90 days following each test window.
[6]Unique wells drilled in the 90 days leading up to, during, or in the 90 days following each test window.
[7]Unique well clusters drilled in the 90 days leading up to, during, or in the 90 days following each test window.

Results

Table 3: Primary specification

Dependent Variable:	Portion Students Met Expectations			
	All Schools		Treated Schools	
	Dummy	*Count*	*Dummy*	*Count*
New Wells Drilled				
Before Test Period	-0.0075*** (0.0018)	-0.0045*** (0.0013)	-0.0066*** (0.0018)	-0.0045*** (0.0013)
During Test Period	-0.0014 (0.0032)	1e-04 (0.0025)	-6e-04 (0.003)	-7e-04 (0.0026)
After Test Period	6e-04	0.0026	5e-04	6e-04
	(0.0026) (0.0026)	(0.002) (0.0018)	(0.0025) (0.0025)	(0.0018) (0.0018)
Existing Wells Drilled				
Before School Year	-0.0053 (0.0045)	5e-04 (4e-04)	-0.0065 (0.0057)	-5e-04 (5e-04)
During School Year	-3e-04 (0.0027)	0.0011 (0.0017)	-5e-04 (0.0027)	4e-04 (0.0016)
Demo Ctrls	x	x	x	x
Weather Ctrls	x	x	x	x
Fixed Effects	x	x	x	x
N.Obs	158129	158129	14586	14586
N.Schools	1711	1711	158	158
R2	0.66	0.66	0.67	0.67

One-way (School) clustered standard-errors
*Signif. Codes: ***: 0.001, **: 0.01, *: 0.05, .: 0.1*

For well clusters 0-3 km from schools. Existing wells drilled before the school year are those drilled from 10 years before the start of the school year to the start of the school year; those drilled during the school year are those drilled from the start of the school year to 90 days before the start of the test period. Fixed effects for (1) year by grade by test and (2) school. Demographic controls for student teacher ratio, portion of students from out of district, portion of students not white, student mobility, and portion of students eligible for free and reduced lunch. Weather controls for total precipitation, count of days with max temps below 10° *C, count of days with max temps above* 25° *C, mean daily wind speed, and count of harmful humidity days during the test period.*

Table 4: Primary specification: all radii

Dependent Variable:	Portion Students Met Expectations											
	0-2 km	0-3 km	0-4 km	0-5 km	0-6 km	0-7 km	0-2 km	0-3 km	0-4 km	0-5 km	0-6 km	0-7 km
	Dummy						*Count*					
New Wells Drilled												
Before Test Period	-0.0107***	-0.0075***	-0.0046**	-0.0023	2e-04	0.0011	-0.0085***	-0.0045***	-0.0029**	-0.0019*	-0.0011	-4e-04
	(0.0029)	(0.0018)	(0.0017)	(0.0017)	(0.0017)	(0.0017)	(0.0023)	(0.0013)	(0.001)	(8e-04)	(7e-04)	(7e-04)
During Test Period	-0.0038	-0.0014	0.0011	-0.0029	-0.0032	-0.0017	-0.0028	1e-04	5e-04	-0.0027	-0.0022	-0.0012
	(0.0037)	(0.0032)	(0.0026)	(0.0025)	(0.0023)	(0.0022)	(0.0035)	(0.0025)	(0.0022)	(0.0021)	(0.0014)	(0.0013)
After Test Period	0.0046	6e-04	-8e-04	-0.0024	-0.0019	4e-04	0.0064*	0.0026	1e-05	-5e-04	-2e-04	0
	(0.0035)	(0.0026)	(0.0023)	(0.0022)	(0.0019)	(0.0019)	(0.0028)	(0.002)	(0.0015)	(0.0012)	(9e-04)	(7e-04)
Existing Wells Drilled												
Before School Year	-0.0024	-0.0053	-0.0021	-0.0014	-0.0034	-0.0021	0.0011	5e-04	3e-04	1e-04	1e-04	1e-04
	(0.0046)	(0.0045)	(0.0041)	(0.0036)	(0.0032)	(0.0029)	(8e-04)	(4e-04)	(2e-04)	(2e-04)	(1e-04)	(1e-04)
During School Year	-5e-05	-3e-04	5e-04	0.0013	0.0017	9e-04	9e-04	0.0011	0.0017	0.0012	0.001	4e-04
	(0.0031)	(0.0027)	(0.0027)	(0.0025)	(0.0022)	(0.0017)	(0.0027)	(0.0017)	(0.0014)	(0.001)	(8e-04)	(6e-04)
Demo Ctrls	x	x	x	x	x	x	x	x	x	x	x	x
Weather Ctrls	x	x	x	x	x	x	x	x	x	x	x	x
Fixed Effects	x	x	x	x	x	x	x	x	x	x	x	x
N.Obs	158129	158129	158129	158129	158129	158129	158129	158129	158129	158129	158129	158129
N.Schools	1711	1711	1711	1711	1711	1711	1711	1711	1711	1711	1711	1711
R2	0.66	0.66	0.66	0.66	0.66	0.66	0.66	0.66	0.66	0.66	0.66	0.66

One-way (School) clustered standard-errors
*Signif. Codes: ***: 0.001, **: 0.01, *: 0.05, .: 0.1*

For well clusters 0-3 km from schools. Existing wells drilled before the school year are those drilled from 10 years before the start of the school year to the start of the school year; those drilled during the school year are those drilled from the start of the school year to 90 days before the start of the test period. Fixed effects for (1) year by grade by test and (2) school. Demographic controls for student teacher ratio, portion of students from out of district, portion of students not white, student mobility, and portion of students eligible for free and reduced lunch. Weather controls for total precipitation, count of days with max temps below 10° *C, count of days with max temps above* 25° *C, mean daily wind speed, and count of harmful humidity days during the test period.*

Table 5: Heterogeneous impacts: by grade

Dependent Variable:	Portion Students Met Expectations			
	Elem. School	Middle School + High School	Elem. School	Middle School + High School
	Dummy		*Count*	
New Wells Drilled				
Before Test Period	-0.000511 (0.004)	-0.006802* (0.0028)	-0.001602 (0.0026)	-0.004173* (0.0018)
During Test Period	-0.009224 (0.0061)	0.003676 (0.0035)	-0.008119 (0.0056)	0.002494 (0.0029)
After Test Period	-0.008605. (0.0048)	0.005773 (0.0043)	-0.006955. (0.0037)	0.0052 (0.0032)
Existing Wells Drilled				
Before School Year	0.0051 (0.0056)	-0.007916 (0.0061)	0.000724* (3e-04)	-0.000791. (4e-04)
During School Year	0.007267 (0.0046)	-0.004444 (0.0034)	0.004052 (0.0028)	-0.001735 (0.002)
Weather Ctrls	x		x	
Fixed Effects	x		x	
N.Obs	158129		158129	
N.Schools	1711		1711	
R2	0.66		0.66	.66

One-way (School) clustered standard-errors
*Signif. Codes: ***: 0.001, **: 0.01, *: 0.05, .: 0.1*

For well clusters 0-3 km from schools, and a dummy for Elementary School interacted with all independent variables. Existing wells drilled before the school year are those drilled from 10 years before the start of the school year to the start of the school year; those drilled during the school year are those drilled from the start of the school year to 90 days before the start of the test period. Fixed effects for (1) year by grade by test and (2) school. Demographic controls for student teacher ratio, portion of students from out of district, portion of students not white, student mobility, and portion of students eligible for free and reduced lunch. Weather controls for total precipitation, count of days with max temps below 10° *C, count of days with max temps above* 25° *C, mean daily wind speed, and count of harmful humidity days during the test period.*

Table 6: Heterogeneous impacts: by subject

Dependent Variable:	Portion Students Met Expectations			
	English	Math	English	Math
	Dummy		*Count*	
New Wells Drilled				
Before Test Period	-0.0062** (0.0019)	-0.003123 (0.0027)	-0.004253*** (0.0013)	-0.002523 (0.0017)
During Test Period	0.000868 (0.0033)	-0.00577 (0.0057)	0.001211 (0.0029)	-0.007072 (0.0053)
After Test Period	-0.000461 (0.0024)	0.003735 (0.0035)	0.000451 (0.0017)	0.001791 (0.0026)
Existing Wells Drilled				
Before School Year	-0.004645 (0.0053)	-0.001193 (0.0029)	-0.00038 (4e-04)	-2.6e-05 (2e-04)
During School Year	0.001703 (0.0025)	-0.005382 (0.004)	0.001482 (0.0016)	-0.002711 (0.0024)
Demo Ctrls	x		x	
Weather Ctrls	x		x	
Fixed Effects	x		x	
N.Obs	158129		158129	
N.Schools	1711		1711	
R2	0.66		0.66	.66

One-way (School) clustered standard-errors
*Signif. Codes: ***: 0.001, **: 0.01, *: 0.05, .: 0.1*

For well clusters 0-3 km from schools, and a dummy for Math subject tests interacted with all independent variables. Existing wells drilled before the school year are those drilled from 10 years before the start of the school year to the start of the school year; those drilled during the school year are those drilled from the start of the school year to 90 days before the start of the test period. Fixed effects for (1) year by grade by test and (2) school. Demographic controls for student teacher ratio, portion of students from out of district, portion of students not white, student mobility, and portion of students eligible for free and reduced lunch. Weather controls for total precipitation, count of days with max temps below 10° *C, count of days with max temps above* 25° *C, mean daily wind speed, and count of harmful humidity days during the test period.*

Table 7: Heterogeneous impacts: top performers

Dependent Variable:	Portion Students in Top Performance Level							
	0-2 km	0-3 km	0-4 km	0-5 km	0-2 km	0-3 km	0-4 km	0-5 km
	Dummy				*Count*			
New Wells Drilled								
Before Test Period	-0.0013 (0.0028)	-0.002 (0.0018)	-0.0016 (0.0016)	-0.0017 (0.0015)	4e-04 (0.0024)	-7e-04 (0.0011)	-0.001 (9e-04)	-9e-04 (7e-04)
During Test Period	4e-04 (0.004)	0.0012 (0.0032)	7e-04 (0.0025)	-3e-04 (0.0024)	0.0017 (0.0042)	0.0018 (0.0027)	1e-05 (0.0022)	-4e-05 (0.0019)
After Test Period	0.0044 (0.0032)	-9e-04 (0.0021)	-0.0016 (0.0016)	-0.0011 (0.0016)	0.0042. (0.0025)	1e-04 (0.0016)	-0.0018 (0.0011)	-0.0013 (8e-04)
Existing Wells Drilled								
Before School Year	-0.001 (0.0036)	-0.004 (0.0031)	0.0019 (0.0033)	0.0071* (0.0032)	9e-04 (8e-04)	6e-04 (4e-04)	2e-04 (2e-04)	1e-04 (1e-04)
During School Year	-7e-04 (0.0026)	-0.0017 (0.0021)	-0.0026 (0.0019)	-0.003 (0.0018)	-2e-04 (0.002)	4e-05 (0.0013)	-2e-04 (9e-04)	-2e-04 (7e-04)
Demo Ctrls	x	x	x	x	x	x	x	x
Weather Ctrls	x	x	x	x	x	x	x	x
Fixed Effects	x	x	x	x	x	x	x	x
N.Obs	158129	158129	158129	158129	158129	158129	158129	158129
N.Schools	1711	1711	1711	1711	1711	1711	1711	1711
R2	0.74	0.74	0.74	0.74	0.74	0.74	0.74	0.74

One-way (School) clustered standard-errors
*Signif. Codes: ***: 0.001, **: 0.01, *: 0.05, .: 0.1*

For well clusters 0-3 km from schools. Existing wells drilled before the school year are those drilled from 10 years before the start of the school year to the start of the school year; those drilled during the school year are those drilled from the start of the school year to 90 days before the start of the test period. Fixed effects for (1) year by grade by test and (2) school. Demographic controls for student teacher ratio, portion of students from out of district, portion of students not white, student mobility, and portion of students eligible for free and reduced lunch. Weather controls for total precipitation, count of days with max temps below 10° *C, count of days with max temps above* 25° *C, mean daily wind speed, and count of harmful humidity days during the test period.*

Table 8: Robustness in clustering decisions

Dependent Variable:	Portion Students Met Expectations					
	Not Clustered	0.2 km, 15 days	2 km, 40 days	Not Clustered	0.2 km, 15 days	2 km, 40 days
		Dummy			*Count*	
New Wells Drilled						
Before Test Period	-0.0079***	-0.0079***	-0.0075***	-4e-04	-0.0038***	-0.0045***
	(0.002)	(0.002)	(0.0018)	(3e-04)	(0.001)	(0.0013)
During Test Period	-8e-04	-0.0012	-0.0014	-5e-04	3e-04	1e-04
	(0.0029)	(0.003)	(0.0032)	(5e-04)	(0.0024)	(0.0025)
After Test Period	0.0033	0.0028	6e-04	7e-04*	0.003*	0.0026
	(0.0026)	(0.0024)	(0.0026)	(3e-04)	(0.0014)	(0.002)
Existing Wells Drilled						
Before School Year	-0.005	-0.005	-0.0053	1e-04	3e-04	5e-04
	(0.0046)	(0.0046)	(0.0045)	(1e-04)	(2e-04)	(4e-04)
During School Year	-0.003	-0.0012	-3e-04	-1e-04	4e-04	0.0011
	(0.0028)	(0.0027)	(0.0027)	(3e-04)	(9e-04)	(0.0017)
Demo Ctrls	x	x	x	x	x	x
Weather Ctrls	x	x	x	x	x	x
Fixed Effects	x	x	x	x	x	x
N.Obs	158129	158129	158129	158129	158129	158129
N.Schools	1711	1711	1711	1711	1711	1711
R2	0.66	0.66	0.66	0.66	0.66	0.66

One-way (School) clustered standard-errors
*Signif. Codes: ***: 0.001, **: 0.01, *: 0.05, .: 0.1*

For well clusters 0-3 km from schools. Existing wells drilled before the school year are those drilled from 10 years before the start of the school year to the start of the school year; those drilled during the school year are those drilled from the start of the school year to 90 days before the start of the test period. Fixed effects for (1) year by grade by test and (2) school. Demographic controls for student teacher ratio, portion of students from out of district, portion of students not white, student mobility, and portion of students eligible for free and reduced lunch. Weather controls for total precipitation, count of days with max temps below 10° *C, count of days with max temps above* 25° *C, mean daily wind speed, and count of harmful humidity days during the test period.*

Table 9: Robustness in clustering decisions: stationary clusters

Dependent Variable:	Portion Students Met Expectations			
	Primary Clustering	Stationary Clusters	Primary Clustering	Stationary Clusters
	Dummy		*Count*	
New Wells Drilled				
Before Test Period	-0.0075*** (0.0018)	-0.00871*** (0.0026)	-0.0045*** (0.0013)	-0.00706** (0.0022)
During Test Period	-0.0014 (0.0032)	-0.00078 (0.0039)	1e-04 (0.0025)	-2e-05 (0.0034)
After Test Period	6e-04 (0.0026)	-0.00077 (0.0029)	0.0026 (0.002)	0.00058 (0.0022)
Existing Wells Drilled				
Before School Year	-0.0053 (0.0045)	0.00221 (0.0049)	5e-04 (4e-04)	0.00053 (7e-04)
During School Year	-3e-04 (0.0027)	-0.00225 (0.0031)	0.0011 (0.0017)	-0.00034 (0.0023)
Demo Ctrls	x	x	x	x
Weather Ctrls	x	x	x	x
Fixed Effects	x	x	x	x
N.Obs	158129	158129	158129	158129
N.Schools	1711	1711	1711	1711
R2	0.66	0.66	0.66	0.66

One-way (School) clustered standard-errors
*Signif. Codes: ***: 0.001, **: 0.01, *: 0.05, .: 0.1*

For well clusters 0-3 km from schools that are stationary across different well clustering results. Existing wells drilled before the school year are those drilled from 10 years before the start of the school year to the start of the school year; those drilled during the school year are those drilled from the start of the school year to 90 days before the start of the test period. Fixed effects for (1) year by grade by test and (2) school. Demographic controls for student teacher ratio, portion of students from out of district, portion of students not white, student mobility, and portion of students eligible for free and reduced lunch. Weather controls for total precipitation, count of days with max temps below 10° C, count of days with max temps above 25° C, mean daily wind speed, and count of harmful humidity days during the test period.

Table 10: Robustness in impact windows

Dependent Variable:	Portion Students Met Expectations	
	Dummy	*Count*
New Wells Drilled		
136-180 Days Before Test Period	0.002 (0.003)	0.0016 (0.0024)
91-135 Days Before Test Period	0.0022 (0.0032)	0.002 (0.0027)
46-90 Days Before Test Period	-0.012*** (0.0029)	-0.0113*** (0.0025)
1-45 Days Before Test Period	-0.0012 (0.0029)	1e-04 (0.002)
During Test Period	-0.0019 (0.0034)	-2e-04 (0.0027)
1-45 Days After Test Period	-0.0017 (0.0027)	-8e-04 (0.0023)
46-90 Days After Test Period	0.0039 (0.0035)	0.0051 (0.0033)
91-135 Days After Test Period	0.0029 (0.003)	0.004 (0.0026)
136-180 Days After Test Period	0.0024 (0.0031)	0.0028 (0.0021)
Existing Wells Drilled		
Before School Year	-0.0056 (0.0045)	6e-04 (4e-04)
During School Year	-4e-04 (0.0035)	-1e-04 (0.0026)
Demo Ctrls	x	x
Weather Ctrls	x	x
Fixed Effects	x	x
N.Obs	158129	158129
N.Schools	1711	1711
R2	0.66	0.66

One-way (School) clustered standard-errors

*Signif. Codes: ***: 0.001, **: 0.01, *: 0.05, .: 0.1*

For well clusters 0-3 km from schools. Existing wells drilled before the school year are those drilled from 10 years before the start of the school year to the start of the school year; those drilled during the school year are those drilled from the start of the school year to 180 days before the start of the test period. Fixed effects for (1) year by grade by test and (2) school. Demographic controls for student teacher ratio, portion of students from out of district, portion of students not white, student mobility, and portion of students eligible for free and reduced lunch. Weather controls for total precipitation, count of days with max temps below 10° *C, count of days with max temps above* 25° *C, mean daily wind speed, and count of harmful humidity days during the test period.*

References

Aesthetic and Noise Control Regulations, 2014. Aesthetic and noise control regulations. Colorado Oil and Gas Conservation Commission. URL: https://cogcc.state.co.us/documents/reg/Rules/LATEST/800Series.pdf.

Agerton, M., Hartley, P.R., III, K.B.M., Temzelides, T., 2017. Employment impacts of upstream oil and gas investment in the united states. Energy Economics 62, 171–180.

Allcott, H., Keniston, D., 2018. Dutch disease or agglomeration? thelocal economic effects of natural resource booms in modern america. The Review of Economic Studies 85, 695–731.

Allshouse, W., McKenzie, L., Barton, K., Brindley, S., Adgate, J., 2019. Community noise and air pollution exposure during the development of a multi-well oil and gas pad. environmental science technology. Environmental Science and Technology 53, 7126–7135.

Archsmith, J., Heye, A., Saberian, S., 2018. Air quality and error quantity: Pollution and performance in a high-skilled, quality-focused occupation. Journal of the Association of Environmental and Resource Economists 5.

Bartik, A.W., Currie, J., Greenstone, M., Knittel, C.R., 2016. The local economic and welfare consequences of hydraulic fracturing. The Albert Shanker Institute January 2017.

Behrans and Associates, 2020. Drilling rig sound control systems. URL: https://www.drillingnoisecontrol.com/brakeshroud.

Boslett, A., Guilfoos, T., Lang, C., 2015. Valuation of expectations: A hedonic study of shale gas development and new york's moratorium. Journal of Environmental Economics and Management 77, 14–30.

Brummet, Q., 2014. The effect of school closings on student achievement. Journal of Public Economics 114, 108–124.

CDE, 2019b. Colorado department of education assessment unit. URL: https://www.cde.state.co.us/cdereval.

CDE, 2020a. Colorado department of education assessment unit. URL: https://www.cde.state.co.us/assessment.

CDE Achievement Levels, 2006. Csap achievement levels. Colorado Student Assessment Program. URL: http://165.127.85.56/sites/default/files/documents/assessment/documents/pld/achievement_level_overview.pdf.

CDE Assessment FAQ, 2021, 2021. Colorado department of education assessment fact sheet and frequently asked questions. URL: https://www.cde.state.co.us/assessment/factsheetsandfaqs#cmasscores.

CDE Interpretive Guide, 2019. Interpretive guide to assessment results; a guide for parents and educators. Colorado Department of Education. URL: https://www.cde.state.co.us/assessment/cmas_coalt_interpretiveguide_2019.

CDPHE, 2020. Cdphe cdoe school locations and district office locations. CDPHE Open Data. URL: https://data-cdphe.opendata.arcgis.com/datasets/cdphe-cdoe-school-locations-and-district-office-locations.

Chetty, R., Friedman, J.N., Rockoff, J.E., 2014. Measuring the impacts of teachers ii: Teacher value-added and student outcomes in adulthood. American Economic Review 104, 2633–2679.

Clark, C., Martin, R., van Kempen, E., Alfred, T., Head, J., Davies, H.W., Haines, M.M., Barrio, I.L., Matheson, M., Stansfeld, S.A., 2005. Exposure-effect relations between aircraft and road traffic noise exposure at school and reading comprehension. American Journal of Epidemiology 163, 27–37.

Climatology Lab, 2020. gridmet. URL: http://www.climatologylab.org/gridmet.html.

Cockx, L., Francken, N., 2016. Natural resources: A curse on education spending? Energy Policy 92, 394–408.

COGCC, 2020, 2020. Well surface location data (updated daily). URL: https://cogcc.state.co.us/data2.html#/downloads.

Cook, N., Heyes, A., 2020. Brain freeze: outdoor cold and indoor cognitive performance. Journal of Environmental Economics and Management 101.

Deming, D., 2009. Early childhood intervention and life-cycle skill development: Evidence from head start. American Economic Journal: Applied Economics 1, 111–134.

Dockrell, J., Shield, B., 2006. Acoustical barriers in classrooms: the impact of noise on performance in the classroom. British Education Research Journal 32, 509–525.

Dunn, S., 2016. Fracking 101: Breaking down the most important part of today's oil, gas drilling. Greeley Tribune. URL: https://www.greeleytribune.com/2016/10/14/fracking-101-breaking-down-the-most-important-part-of-todays-oil-gas-drilling/.

Dunn, S., 2017. Older vertical wells are slowly disappearing from weld county, colorado landscapes. Greeley Tribune. URL: https://www.greeleytribune.com/2017/04/29/older-vertical-wells-are-slowly-disappearing-from-weld-county-colorado-landscapes/.

Duque, V., Gilraine, M., 2020. Coal use and student performance. Annenberg Institute at Brown University Ed Working Paper: 20-251.

Ebenstein, A., Lavy, V., Roth, S., 2016. The long-run economic consequences of high-stakes examinations: Evidence from transitory variation in pollution. American Economic Journal: Applied Economics 8, 36–65.

Fedan, J., 2020. Biological effects of inhaled hydraulic fracturing sand dust. Toxicology and Applied Pharmacology 409.

Feyrer, J., Mansur, E., Sacerdote, B., 2017. Geographic dispersion of economic shocks: Evidence from the fracking revolution. American Economic Review 107, 1313–34.

Gittings, R.K., Roach, T., 2019. Who really benefits from a resource boom? evidence from the marcellus and utica shale plays. Energy Economics 87.

Goodman, P.S., Galatioto, F., Thorpe, N., Namdeo, A.K., Davies, R.J., Bird, R.N., 2016. Investigating the traffic-related environmental impacts of hydraulic-fracturing (fracking) operations. Environment International 89-90, 248–260.

Gopalakrishnan, S., Klaiber, H., 2014. Is the shale energy boom a bust for nearby residents? evidence from housing values in pennsylvania. American Journal of Agricultural Economics 96, 43–66.

Guettabi, M., James, A., 2020. Who benefits from an oil boom? evidence from a unique alaskan data set. Resource and Energy Economics 62.

Gylfason, T., 2001. Natural resources, education,and economic development. European Economic Review 45, 847–859.

Hays, J., McCawley, M., Shonkoff, S.B., 2016. Public health implications of environmental noise associated with unconventional oil and gas development. Science of the Total Environment .

Heissel, J., Persico, J., Simon, D., 2019. Does pollution drive achievement? the effect of traffic pollution on academic performance. National Bureau of Economic Research Working Paper No. 25489 .

Hill, E., 2018. Shale gas development and infant health: Evidence from pennsylvania. Journal of Health Economics , 134–150.

Jamison, E., Jamison, D., Hanushek, E., 2019. The effects of education quality on income growth and mortality decline. Economics of Education Review 26.

Jones, J., Ready, E., Hazel, A., 2020. Social network analysis for anthropologists; materials for a short course. Stanford Human Evolutionary Ecology and Health. URL: https://eehh-stanford.github.io/SNA-workshop/intro-igraph.html#adjacency-matrices.

Kohler, J., 2020. Colorado takecolorado poised to takeanother inanother major step inregulating regulating oil, gas industry —this setbacksthis time on well setbacks. The Denver Post. URL: https://www.denverpost.com/2020/09/28/colorado-regulators-favoring-big-well-setbacks/.

Komisarow, S., Pakhtigian, E.L., 2020. Are power plant closures a breath of fresh air? air pollution, absences, and achievement. Annenberg Institute at Brown University Ed Working Paper: 20-290.

Langley, D., 2011. Rig mobility on the fast track. Drilling Contractor. URL: https://www.drillingcontractor.org/rig-mobility-on-the-fast-track-10666.

Laramie Energy, 2020. Operations — williams fork. Laramie Energy. URL: https://www.laramie-energy.com/operations/.

Maniloff, P., Mastromonaco, R., 2017. The local employment impacts of fracking: A national study. Resource and Energy Economics 49.

Marchand, J., Weber, J., 2015. The local effects of the texas shale boom on schools, students, and teachers. University of Alberta, Department of Economics Working Paper 15.

Marcotte, D., 2017. Something in the air? air quality and children's educational outcomes. Economics of Education Review 56.

Mason, R., 2014. Walking tall: Pad drilling prompts rig design evolution. Hart Energy. URL: https://www.hartenergy.com/exclusives/walking-tall-pad-drilling-prompts-rig-design-evolution-19989.

Mejía, J.F., Choy, S.L., Mengersen, K., Morawska, L., 2011. Methodology for assessing exposure and impacts of air pollutants in schoolchildren: Data collection, analysis and health effectse – a literature review. Atmospheric Environment 45, 813–823.

Metcalfe, R., Burgess, S., Proud, S., 2011. Student effort and educational attainment: Using the england football team to identify the education production function. Centre for Market and Public Organisation 11.

Muehlenbachs, L., Spiller, E., Timmins, C., 2015. The housing market impacts of shale gas development. American Economic Review 105, 3633–3659.

NYS DEC, 2009. Impact assessment of natural gas production in the new york city water supply watershed. rapid impact assessment report. hazen and sawyer report on behalf of the new york city department of environmental protection. URL: http://www.nyc.gov/html/dep/pdf/natural_gas_drilling/rapid_impact_assessment_091609.pdf.

Parades, D., Komarek, T., Loveridge, S., 2015. Income and employment effects of shale gas extraction windfalls: Evidence from the marcellus region. Energy Economics 47, 112–120.

Park, J., 2016. Hot temperature and high stakes exams: Evidence from new york city public schools. Unpublished Mansuscript; Harvard University Economics Department .

Pearce, D.W., Atkinson, G.D., 1993. Capital theory and the measurement of sustainable development: an indicator of "weak" sustainability. Ecological Economics 8, 103–108.

Pearson, M., 2020. The seven steps of oil and natural gas extraction. URL: https://www.cred.org/seven-steps-of-oil-and-natural-gas-extraction/.

Persico, C., Venator, J., 2018. The effects of local industrial pollution on students and schools. Zurich: IZA Conference on Labor Market Effects of Environmental Policies .

Pickett, A., 2015. Pad drilling; leading operators improve efficiency and effectiveness of multiwell pad operations. The American Oil & Gas Reporter. URL: https://www.aogr.com/magazine/cover-story/leading-operators-improve-efficiency-and-effectiveness-of-multiwell-pad-ope.

Raimi, D., Newell, R.G., 2016. Us state and local oil and gas revenues. Resources for the Future Discussion Paper , 16–50.

Ransom, M., Pope, C.A., 1992. Elementary school absences and pmlo pollution in utah valley. Environmental Research 58, 204–219.

Ronfeldt, M., Loeb, S., Wyckoff, J., 2013. How teacher turnover harms student achievement. American Educational Research Journal 50, 4–36.

Roth, S., 2016. The contemporaneous effect of indoor air pollution on cognitive performance: Evidence from the uk. Zurich: IZA Conference on Labor Market Effects of Environmental Policies .

Shield, B., Dockrell, J., 2008. The effects of environmental and classroom noise on the academic attainments of primary school children. The Journal of the Acoustical Society of America 123.

Shogren, K., Shaw, L., Raley, S., Wehmeyer, M., 2018. Exploring the effect of disability, race-ethnicity, and socioeconomic status on scores on the self-determination inventory: Student report. Exceptional Children 85, 10–27.

Smith, A.J.B., 2020. Geographic dispersion of economic shocks: Evidence from the fracking revolution: Comment. American Economic Review 110.

Statewide Setbacks, 2013. Statewide setbacks – final rules. Colorado Oil and Gas Conservation Commission. URL: https://cogcc.state.co.us/documents/reg/Rules/2012/setback/Final_SetbackRules.pdf.

Thuot, K., 2014. On the launch pad: The rise of pad drilling. Enverus Drilling Info. URL: https://www.enverus.com/blog/launch-pad-rise-pad-drilling/.

US Census, 2010-2014;2014-2018a. Us census bureau american community survey. Colorado State Demography Office Census API. URL: https://demography.dola.colorado.gov/CensusAPI/.

US Census, 2019b. Us census bureau tracts. TIGRIS. URL: https://www2.census.gov/geo/pdfs/reference/GARM/Ch10GARM.pdf.

USDA, 2019, 2019. Irrigation and water management; results from the 2018 irrigation and water management survey. 2017 Census of Agriculture Highlights. United States Department of Agriculture, National Agriculture Statistics Service. URL: https://www.nass.usda.gov/Publications/Highlights/2019/2017Census_Irrigation_and_WaterManagement.pdf.

Waskom, R., Neibauer, M., 2014. Water conservation in and around the home. Colorado State University Extension .

Weiner, C., 2014. Oil and gas development in colorado. Colorado State University Extension .

Wright, D., 1999. Student mobility: A negligible and confounded influence on student achievement. The Journal of Educational Research 92, 347–353.

Zachary, L., Ratledge, N., 2017. Public education impacts of unconventional oil and gas development. Resources for the Future .

Zhang, X., Chen, X., Zhang, X., 2017. The impact of exposure to air pollution on cognitive performance. Proceedings of the National Academy of Sciences .

Zivin, J.G., Hsiang, S.M., Neidell, M., 2018. Temperature and human capital in the short and long run. Journal of the Association of Environmental and Resource Economics 5, 77–105.

Appendix A: Additional Tables and Figures

Table A1: Primary specification — fractional logit

Dependent Variable:	Portion Students Met Expectations	
	Dummy	*Count*
New Wells Drilled		
Before Test Period	-0.0846*** (0.02)	-0.0621*** (0.0141)
During Test Period	-0.0018 (0.0365)	-0.0052 (0.0303)
After Test Period	0.0114 (0.0291)	0.014 (0.0199)
Existing Wells Drilled		
Before School Year	-0.0459 (0.0604)	-0.0029 (0.0044)
During School Year	-0.0144 (0.0301)	0.0008 (0.0199)
Controls		
Portion Students Non-White	-1.6479*** (0.178)	-1.6524*** (0.1778)
Portion Students Free/Reduced Lunch Eligible	-0.4478*** (0.0853)	-0.4482*** (0.0853)
Portion Students Out of District	0.4383** (0.1618)	0.4379** (0.1618)
Teacher Turnover Rate	-0.3633** (0.1303)	-0.3632** (0.1303)
Student Mobility Rate	-0.0067*** (0.0009)	-0.0067*** (0.0009)
Student Teacher Ratio	-0.0007 (0.0004)	-0.0007 (0.0004)
Mean Daily Wind Speed (km/h)	0.0097 (0.0202)	0.0094 (0.0202)
Total Precipitation (mm)	0.0002 (0.0004)	0.0002 (0.0004)
Max Temp Below 10C (days)	-0.0051* (0.0023)	-0.0051* (0.0023)
Max Temp Above 25C (days)	-0.0099** (0.0038)	-0.0101** (0.0038)
Harmful Humidity (days)	0.0022 (0.0021)	0.0022 (0.0021)
Constant	3.0887*** (0.1953)	3.0933*** (0.1952)
Fixed Effects	x	x
N.Obs	158129	158129
N.Schools	1711	1711

One-way (School) clustered standard-errors
*Signif. Codes: ***: 0.001, **: 0.01, *: 0.05, .: 0.1*

For well clusters 0-3 km from schools. Existing wells drilled before the school year are those drilled from 10 years before the start of the school year to the start of the school year; those drilled during the school year are those drilled from the start of the school year to 90 days before the start of the test period. Fixed effects for (1) year by grade by test and (2) school.

Table A2: Primary specification — complete model results

Dependent Variable:	Portion Students Met Expectations	
	Dummy	*Count*
New Wells Drilled		
Before Test Period	-0.0075*** (0.00177)	-0.0045*** (0.00126)
During Test Period	-0.0014 (0.00322)	1e-04 (0.00254)
After Test Period	6e-04 (0.00259)	0.0026 (0.00203)
Existing Wells Drilled		
Before School Year	-0.0053 (0.00455)	5e-04 (0.00038)
During School Year	-3e-04 (0.00268)	0.0011 (0.00167)
Controls		
Portion Students Non-White	-0.1331*** (0.01637)	-0.1326*** (0.01641)
Portion Students Free/Reduced Lunch Eligible	-0.0515*** (0.00931)	-0.0516*** (0.00932)
Portion Students Out of District	0.048*** (0.01173)	0.0483*** (0.01174)
Teacher Turnover Rate	-0.039*** (0.01088)	-0.0385*** (0.01082)
Student Mobility Rate	-4e-04*** (8e-05)	-4e-04*** (8e-05)
Student Teacher Ratio	-1e-04 (5e-05)	-1e-04 (5e-05)
Mean Daily Wind Speed (km/h)	0 (0.00159)	0 (0.00159)
Total Precipitation (mm)	1e-04. (3e-05)	1e-04. (3e-05)
Max Temp Below 10C (days)	-6e-04*** (0.00018)	-6e-04*** (0.00018)
Max Temp Above 25C (days)	-0.0015*** (0.00032)	-0.0016*** (0.00032)
Harmful Humidity (days)	3e-04 (0.00017)	3e-04. (0.00017)
Fixed Effects	x	x
N.Obs	158129	158129
N.Schools	1711	1711
R2	0.66	0.66

One-way (School) clustered standard-errors
*Signif. Codes: ***: 0.001, **: 0.01, *: 0.05, .: 0.1*

For well clusters 0-3 km from schools. Existing wells drilled before the school year are those drilled from 10 years before the start of the school year to the start of the school year; those drilled during the school year are those drilled from the start of the school year to 90 days before the start of the test period. Fixed effects for (1) year by grade by test and (2) school.

Table A3: Concentric radii

Dependent Variable:	Portion Students Met Expectations					
	Before	During	After	Before	During	After
		Dummy			*Count*	
Wells Drilled						
0-1 km	-0.0021 (0.00477)	0.0065 (0.00701)	0.0013 (0.00556)	-0.001 (0.00466)	0.0051 (0.00743)	0.0015 (0.00483)
1-2 km	-0.0118*** (0.00336)	-0.0052 (0.00401)	0.008* (0.00367)	-0.01*** (0.00267)	-0.0057 (0.00386)	0.0084* (0.00331)
2-3 km	-0.003 (0.00223)	2e-04 (0.00404)	-0.0017 (0.00256)	-0.0023 (0.00175)	7e-04 (0.00357)	2e-04 (0.00223)
3-4 km	-0.002 (0.00229)	0.0011 (0.00326)	-0.0038 (0.00251)	-0.0012 (0.00182)	0.0012 (0.00317)	-0.0032 (0.00202)
4-5 km	-6e-04 (0.00179)	-0.008* (0.0033)	-0.0023 (0.00211)	-3e-04 (0.00134)	-0.0078** (0.00294)	-0.0018 (0.00177)
5-6 km	3e-04 (0.00165)	-0.0014 (0.00286)	0.0018 (0.0018)	0.0011 (0.00131)	-0.001 (0.00239)	0.0013 (0.00132)
6-7 km	0.0015 (0.00182)	0.0011 (0.00275)	0.0021 (0.00199)	0.0012 (0.00108)	0.0019 (0.00241)	0.0013 (0.00138)
Existing Well Ctrls		x			x	
Demo Ctrls		x			x	
Weather Ctrls		x			x	
Fixed Effects		x			x	
N.Obs		158129			158129	
N.Schools		1711			1711	
R2		0.66			0.66	

One-way (School) clustered standard-errors
*Signif. Codes: ***: 0.001, **: 0.01, *: 0.05, .: 0.1*

Fixed effects for (1) year by grade by test and (2) school. Existing well controls for the wells drilled in the 10 years leading up to the start of the school year and between the start of the school year and 90 days before the start of the test period. Demographic controls for student teacher ratio, portion of students from out of district, portion of students not white, student mobility, and portion of students eligible for free and reduced lunch. Weather controls for total precipitation, count of days with max temps below 10° *C, count of days with max temps above* 25° *C, mean daily wind speed, and count of harmful humidity days during the test period.*

Table A4: Heterogeneous impacts: population density

Dependent Variable:	Portion Students Met Expectations	
	Dummy	*Count*
New Wells Drilled		
Before Test Period	-0.011144***	-0.005193**
	(0.00317)	(0.00189)
During Test Period	0.002695	0.001623
	(0.00361)	(0.00303)
After Test Period	0.007835*	0.005095
	(0.00374)	(0.00311)
Before Test Period * Density	-1e-06	-1.3e-05
	(2e-05)	(1e-05)
During Test Period * Density	2.6e-05	2.7e-05
	(3e-05)	(3e-05)
After Test Period * Density	-4.9e-05*	-3.5e-05.
	(2e-05)	(2e-05)
Existing Wells Drilled		
Before School Year	0.011751	-0.000232
	(0.00729)	(0.00058)
During School Year	0.000232	0.001848
	(0.00365)	(0.00267)
Before School Year * Density	-0.000163***	-6e-06*
	(5e-05)	(0)
During School Year * Density	-7e-06	-5e-06
	(2e-05)	(2e-05)
Demo Ctrls	x	x
Weather Ctrls	x	x
Fixed Effects	x	x
N.Obs	109569	109569
N.Schools	1670	1670
R2	0.65	0.65

One-way (School) clustered standard-errors
*Signif. Codes: ***: 0.001, **: 0.01, *: 0.05, .: 0.1*

We have fewer observations due to the density data only available from 2010-2018.
For well clusters 0-3 km from schools, and the density of the students grade 1-8 in the school's census tract that year. Existing wells drilled before the school year are those drilled from 10 years before the start of the school year to the start of the school year; those drilled during the school year are those drilled from the start of the school year to 90 days before the start of the test period. Fixed effects for (1) year by grade by test and (2) school. Demographic controls for student teacher ratio, portion of students from out of district, portion of students not white, student mobility, and portion of students eligible for free and reduced lunch. Weather controls for total precipitation, count of days with max temps below 10° *C, count of days with max temps above* 25° *C, mean daily wind speed, and count of harmful humidity days during the test period.*

Table A5: Heterogeneous impacts: score gap index

Dependent Variable:	Portion Students Met Expectations	
	Dummy	*Count*
New Wells Drilled		
Before Test Period	0.000131 (0.0032)	0.001262 (0.0021)
During Test Period	-0.007201 (0.0052)	-0.004492 (0.0053)
After Test Period	-0.002236 (0.0048)	-0.003793 (0.0038)
Before Test Period * Index	-0.018594* (0.0083)	-0.015761** (0.0054)
During Test Period * Index	0.015008 (0.0168)	0.008836 (0.0166)
After Test Period * Index	0.006155 (0.0127)	0.010923 (0.0098)
Existing Wells Drilled		
Before School Year	-0.007709 (0.0089)	-0.001026. (6e-04)
During School Year	0.001083 (0.0047)	-0.000824 (0.0027)
Before School Year * Index	0.007199 (0.0262)	0.001579 (0.0016)
During School Year * Index	-0.00329 (0.0142)	0.003392 (0.0098)
Demo Ctrls	x	x
Weather Ctrls	x	x
Fixed Effects	x	x
N.Obs	158129	158129
N.Schools	1711	1711
R2	0.66	0.66

One-way (School) clustered standard-errors
*Signif. Codes: ***: 0.001, **: 0.01, *: 0.05, .: 0.1*

For well clusters 0-3 km from schools and the score gap index of the mean of the annual school's portion of students not white and eligible for free and reduced lunch. Existing wells drilled before the school year are those drilled from 10 years before the start of the school year to the start of the school year; those drilled during the school year are those drilled from the start of the school year to 90 days before the start of the test period. Fixed effects for (1) year by grade by test and (2) school. Demographic controls for student teacher ratio, portion of students from out of district, portion of students not white, student mobility, and portion of students eligible for free and reduced lunch. Weather controls for total precipitation, count of days with max temps below 10° *C, count of days with max temps above* 25° *C, mean daily wind speed, and count of harmful humidity days during the test period.*

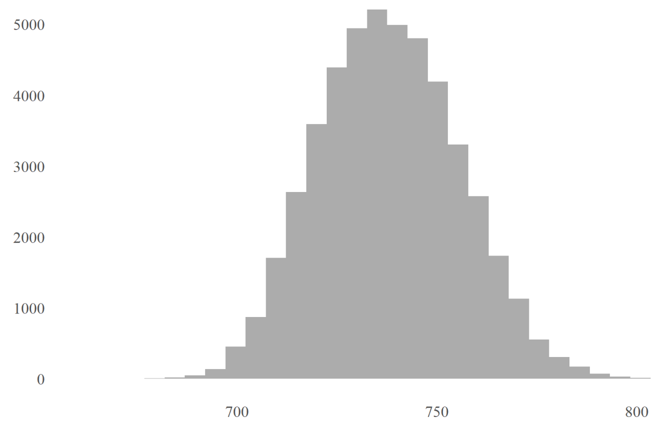


Figure A1: Score distributions.
Mean scale scores by content, grade, year,
and school from 2015-2019

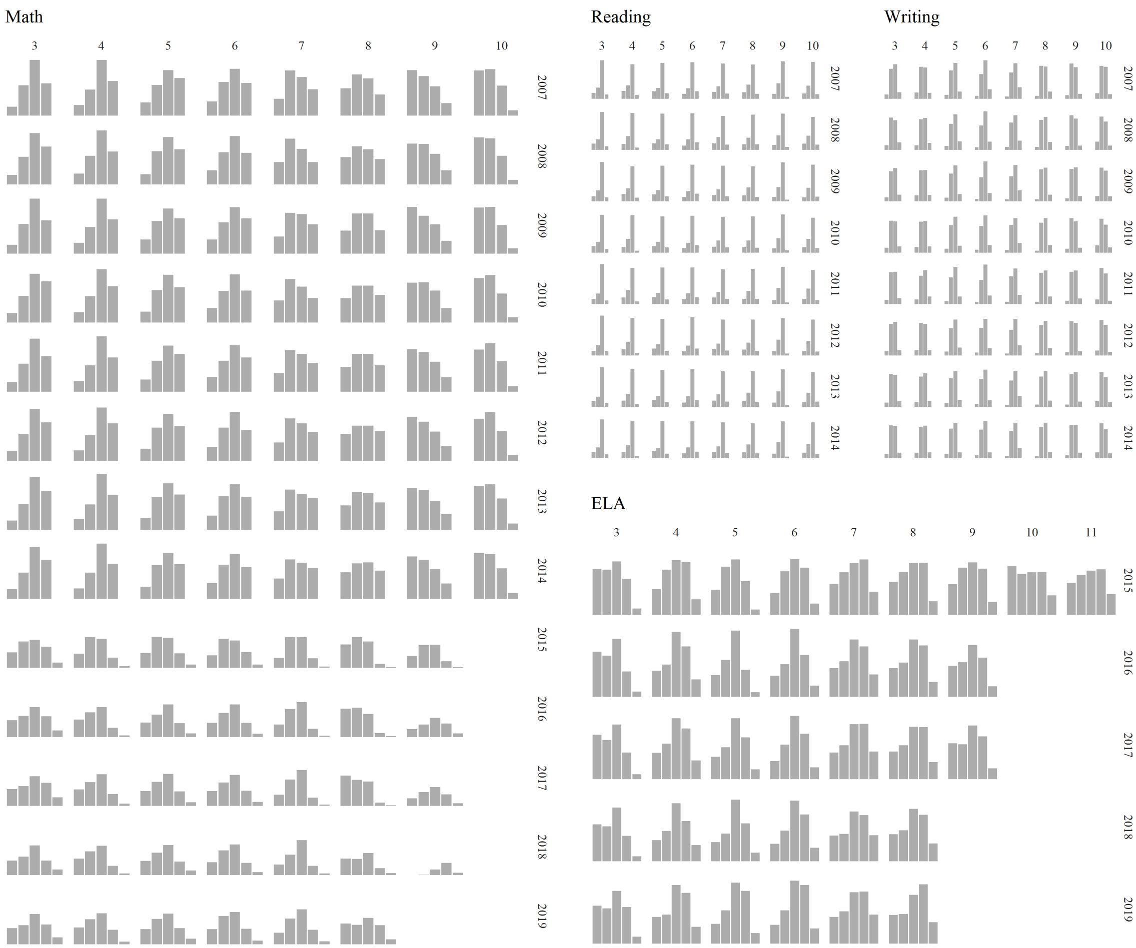
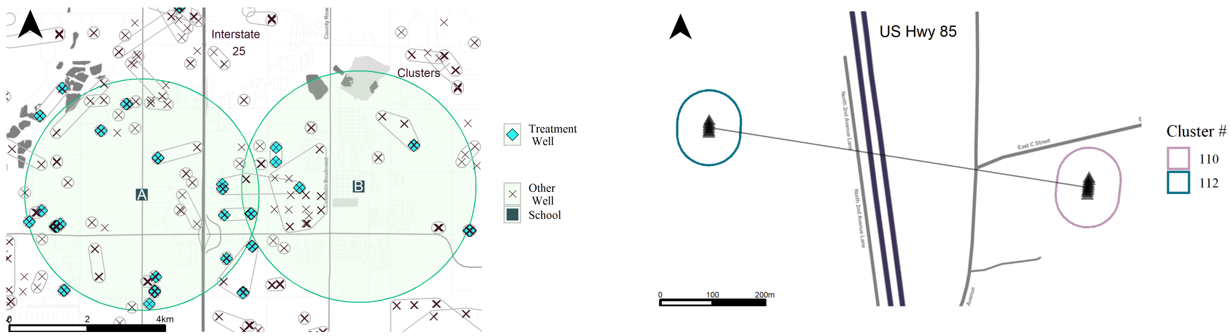


Figure A2: Score distributions.
Performance levels by content, grade, and year

Appendix B: Well Clustering

In order to accurately identify the UOGD activity that impacts students, we cluster wells so that our treatment variable assignments reflect the activities that occur on the well-pad level. The expansion of well pad drilling that started in 2006 (Mason, 2014; Langley, 2011) means that fracturing may not follow the drilling of each individual well. Instead, operators tend to prepare the well pad before they begin drilling any wells on the pad (NYS DEC). They then drill all wells on the pad in quick succession, with near continuous operation of the drill rig. When all wells have been drilled, fracking service providers prepare to fracture and then pump fracking fluid into all of the wells at a well pad in quick succession, with near continuous operation of the fracking rig (Pickett, 2015). In the case of adjacent well pads, fracking service providers build pipes from their rigs to the well pads to hydraulic fracture all wells in the same time period. Using a clustering method with well dates and locations as inputs, we accurately estimate relative timing of UOGD activities that occur on a well-pad level.

The method we apply uses graphs for each operator's wells constructed from three intersected filtered adjacency matrices: for days between spud dates, distances between well locations, and primary roads in between well locations. We split all horizontal wells in Colorado during the study period by well operator, and filter the adjacency matrices such that edge connections remain for an operators' wells that were drilled less than 40 days and 2 km apart, and are not split by primary roads. An edge connection is a true value for the ij^{th} element where row i and column j correspond to the pair of wells i and j that meet all criteria (since the matrices are symmetrical, there will also be a true value for the ji^{th} element). Wells are assigned to clusters closed under edge connection, which is to say that every well in a cluster is connected via an edge to at least one other well in that cluster. Our assumptions in doing so are that these critical values define the wells that are fractured sequentially. The temporal and spatial critical values allow for a cluster to contain multiple well pads if they are near enough to each other that they are likely fractured together. By splitting



(a) Wells and clusters in Weld County

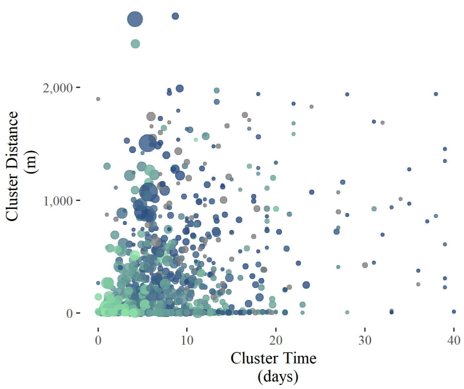
(b) Two clusters in Greeley, 2015

Figure B1: Spatial and temporal criteria in well clustering

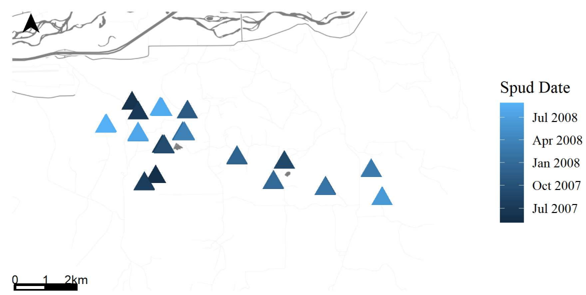
by roads, we incorporate the fact that fracking service providers are most unlikely to place fracking rigs on the opposite side of primary roads from the well pad.

Figure B1a shows how wells might be clustered. The wells are split on either side of Interstate 25, and no clusters span both sides of the interstate. On each side of the interstate, wells are assigned to distinct clusters due to time elapsed between well spud dates, the operators who drilled the wells, or distance between wells. Figure B1b shows wells that are temporally and spatially proximate but assigned to different clusters due to being split by roads.

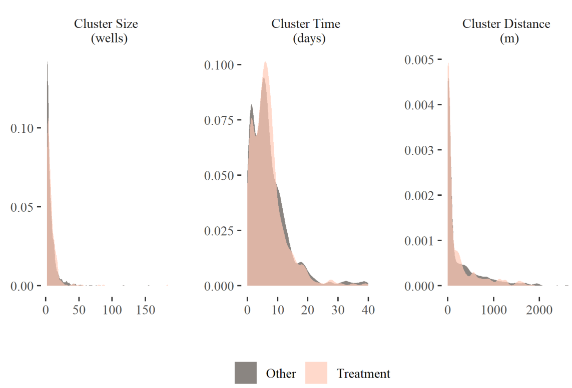
To analyze the results of our clustering method, we calculate cluster spread and size over time where cluster spread is the mean days or mean meters between all wells in a cluster when ordered temporally and cluster size is the number of wells in a cluster. Figure B2a depicts the properties of each well cluster. Since the wells are already ordered temporally, there are no wells with mean temporal spread greater than 40 days, however there are some outlier clusters in which sequential wells were spread more than 2 km. One of these outliers is found in Grand County, where an operator continuously drilled well pads in a non-linear pattern within a tight radius, perhaps moving around due to the mountainous terrain (Figure B2b). In Figure B2a we see that our clustering critical values bind some clusters on the top and right of the figure, but that largely our methods capture intuitive clustering, as found in the concentration of clusters in the bottom-left portion of the figure. This is depicted in Figure B2c, which shows that a plurality of wells have a mean temporal spread of less than 10 days and a mean distance spread of less than 500 m. Further, we see that over time cluster spread decreases while size increases (Figures B2a, B2d) due to the increasing affordability and adoption of mobile drill rigs, which make it cheaper and easier for operators to drill multiple wells in close proximity. It is also likely that distances between sequential wells decreased as technology advances allowed operators to drill longer distances horizontally at a similar cost and it subsequently became more profitable to drill more well heads in a single location.



(a) Spread, size, and time



(b) Outlier example, Grand County, 2018



(c) Study period distributions

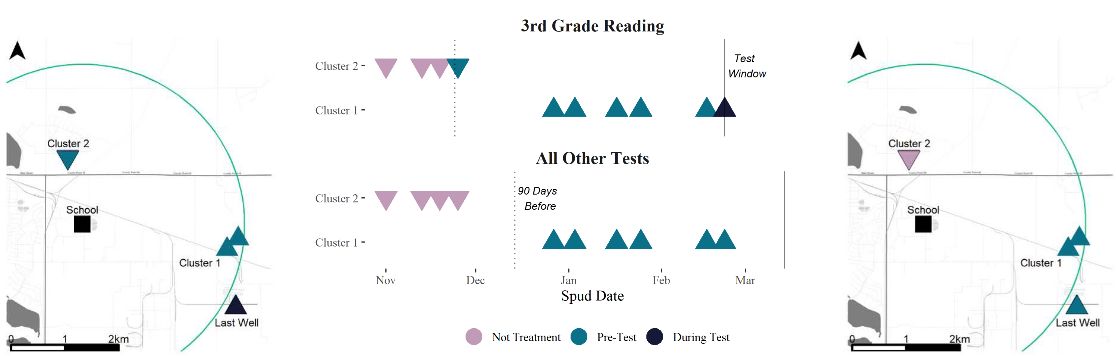
(d) Annual means

Figure B2: Cluster size and spread in all clusters and treatment clusters

We also find that the treatment well clusters are representative of all well clusters found in the data. Compared to those that do not treat tests, the well clusters drilled near schools around test windows have a similar distribution of size, mean temporal spread, and mean spatial spread (Figure B2c). Additionally, the annual means for each metric across all treated clusters is in line with the corresponding mean for clusters that do not treat tests (Figure B2d).

Clusters are assigned to test windows based on the spud date of the last well drilled in the cluster and the location of the nearest well to the school. Since well clusters are fracked after the last well is drilled, the spud date of the last well is the most accurate possible start date of post-drilling UOGD activity. To demonstrate how the logic is applied, we show the assignment of clusters to a school in Windsor, CO for the test periods in 2011 using a 3 km radius (Figures B3). For the years 2007-2014, there were two test windows: one for 3^{rd} grade reading, and another starting later for all other test subjects and grades. For both tests, the nearest well is found within 3 km and eligible for treatment assignment (Figures B3a, B3c).

However, the students are treated differently due to the overlap of the last wells' spud dates with the test windows (Figure B3b). Due to the timing of the spud date of its last well, Cluster 1 is classified as a treatment cluster during the test window for 3^{rd} grading reading and as a treatment cluster before the test window for the other tests. The students in 3^{rd} grade taking the reading test are also considered to be treated by post-drilling activity from Cluster 2 since the last well of that cluster is drilled less than 90 days before the start of the test period. In contrast, the other students in the same school are not treated by Cluster 2 since more than 90 days elapsed since the last well had been drilled — any fracking has likely ended by the time the testing begins.



(a) 3rd grade reading (b) Legend: temporal assignment (c) All other tests

Figure B3: An example of treatment assignments based on clustering to a school in Windsor, 2011.

Using graph theory to cluster wells allowed us to group wells with precision, and in doing so to approximate the missing data of fracking activity start and end dates. We believe that this method can be applied in other economic settings to impute missing data and improve the accuracy of treatment variables. Future applications of graphs to cluster treatments may incorporate greater flexibility into the critical values used to filter adjacency matrices. For instance in this setting, by incorporating granular industry knowledge to choose the critical values by year, season, region, or operator, clustering can more accurately reflect UOGD activity. Our clustering decisions reflect the nature of UOGD activity such that we can attribute the negative impacts of UOGD activity on student standardized test performance to fracturing activities (Table 3). Further, our robustness tests show that on the extensive margin, the use of the clustering method is more important than the precise critical values used to filter the adjacency matrices (Table 8).