Does Shale Gas Development Impact Infant Health through Drinking Water?*

Elaine L. Hill† Lala Ma‡

December 2017

Abstract

Widespread hydraulic fracturing of shale formations has yielded a range of economic and environmental benefits. There are, however, various costs associated with shale gas development (SGD) that remain uncertain. This paper looks to SGD operations in Pennsylvania to assess the magnitude of drinking water impacts and whether there exists health risks associated with SGD through this medium. Using the universe of birth records in Pennsylvania from 2003-2015 and all ground water-based Community Water System (CWS) drinking water contaminant measurements between 2011-2015, we investigate this question by building a novel data set that links gas well activity to infant health and public drinking water outcomes based on a water system's source location. This is the first study to examine the impacts of SGD on *public* drinking water quality and to identify the health impacts of SGD through the specific mechanism of water. Our difference-in-differences models find consistent evidence of an effect of shale gas development on both drinking water quality and infant health outcomes. The results are robust to the inclusion of various correlated threats that can threaten identification of impacts, fixed effects, and placebo tests. Together, our paper informs an important question of whether SGD affects reproductive health through the mechanism of drinking water. A better understanding of these potential external costs has important implications for regulatory policy and is crucial for weighing the costs of such operations against their economic and environmental benefits.

*We thank Richard DiSalvo and Andrew Boslett for excellent research support. We are grateful to seminar participants at AERE Summer Conference, IHEA conference, MEA conference, NBER EEE Summer Institute, VEAM-fest, Georgia State University, Penn State University, Northern Illinois University, University of Binghamton, University of Chicago, University of Kentucky, and University of Rochester for useful comments. We gratefully acknowledge funding from the University of Rochester Environmental Health Sciences Center (EHSC), an NIH/NIEHS-funded program (P30 ES001247), and the NIH Director's Early Independence Award funded by the NIH Common Fund (5 DP5 OD021338-02; PI Hill). The views expressed herein are those of the authors and do not necessarily reflect the views of the National Institutes of Health.

†Hill: Department of Public Health Sciences, University of Rochester School of Medicine, Rochester, NY 14642, elaine_hill@urmc.rochester.edu.

‡Ma: Gatton College of Business and Economics, Department of Economics, University of Kentucky, Business & Economics Building, Lexington, KY 40506, lala.ma@uky.edu.

Over the last decade, technological innovations in high-volume horizontal hydraulic fracturing, commonly known as 'fracking', has allowed for the cost-effective recovery of energy resources from tight rock formations. Most notably, widespread hydraulic fracturing of shale formations containing abundant amounts of natural gas has taken place in many regions across the US. These operations have yielded a range of benefits from reductions in energy costs to improvements in greenhouse gas emissions, which are timely in the face of rising energy demand (Allcott and Keniston, 2014; Bartik et al., 2016; Feyrer et al., 2199; Hausman and Kellogg, 2015; Mason et al., 2015). On the other hand, various costs associated with unconventional shale gas development (SGD) exist and are borne by populations that are exposed to these operations. Of particular concern is whether the recent shale gas boom has led to contamination of water resources and sustained potential health impacts to exposed populations (Finkel and Hays, 2015). Effluent from waste water treatment plants that process waste from shale gas operations has shown levels of contaminants that exceed the maximum allowable limit in drinking water (Vengosh et al., 2015). Various scientific studies have also documented elevated levels of contaminants in private groundwater drinking wells near SGD (Osborn et al., 2011; Warner et al., 2012; Jackson et al., 2013; Drollette et al., 2015; Hildenbrand et al., 2015). Concern over potential health risks has resulted in an approximately $19 million increase in bottled water purchases in 2010 in response to SGD in Pennsylvania (Wrenn et al., 2016). Homeowners in this state have also been found to require a discount of anywhere between 9.9 and 16.5 percent of housing prices to compensate for perceived ground water contamination risk as a result of living near SGD (Muehlenbachs et al., 2015). The magnitude of these risks, while potentially salient, remain uncertain, and opens the potential for direct health impacts for those who do not take costly measures of protection. Thus, research is needed on the extent of these risks more generally and whether the associated risks are large enough to produce measurable impacts on health.

This paper looks to SGD operations in Pennsylvania to assess the magnitude of drinking water impacts and whether there exists health risks associated with SGD by means of threats to drinking water quality. To tackle this question, we compile a novel data set that links gas well activity to both health outcomes from the universe of birth records in Pennsylvania from 2003-2015 and all ground water-based Community Water System (CWS) drinking water contaminant measurements between 2011-2015. This is accomplished by using the exact geographic locations of a mother's residence, gas wells, and public drinking water source intake locations to infer exposure at a high spatial resolution, as well as temporal information on births, well bore activity, and water sampling at high frequency to narrow the timing of exposure. Using a difference-in-differences approach, our baseline specifications estimate the impacts of increasing SGD activity within 1 km of public water system sources on drinking water contaminants and birth outcomes relative to impacts as measured by gas well threats at farther distances.

Our contributions are twofold. First, aside from our companion paper Hill and Ma (2017), we are the first study to investigate the impacts of SGD on *public* drinking water quality. All field work to date are regional studies that examine private drinking water quality impacts. This 'first stage' question is extremely relevant from a policy perspective. If such an externality exists with respect to SGD operations, it justifies the role for regulation of firms from an efficiency standpoint, since any measurable impact would be indicative that the industry is not internalizing these unintended, environmental consequences.[1] Furthermore, given the focus on *public* drinking water quality, our paper yields an additional implication regarding the regulation of drinking water standards. Although the US Environmental Protection Agency (EPA), under the Safe Drinking Water Act, regulates over 90 contaminants by establishing testing frequencies and Maximum Contaminant Level (MCL) thresholds, many of the contaminants that are suspected to be tied to SGD are unregulated. Conditional on there being measurable drinking water pollution and health impacts, there may be a need to assess whether the benefits of increasing the set of regulated contaminants or the stringency of drinking water contaminants associated with the industry (e.g. through frequency or MCL thresholds) will outweigh its regulatory costs.

Next, while there have been work in both epidemiology and economics that examine the health impacts of SGD, all have used measures of exposure based on residence proximity to gas wells, leaving the mechanism of impacts unknown. Our second contribution is to identify the health impacts of SGD through the specific mechanism of water. We are able to distinguish in utero exposure to SGD via water source proximity to gas wells as opposed to exposure based on residence proximity to gas wells. Although the correlation between the two types of exposure is high (~ 0.6), it is not perfect. We use this variation to recover the health impacts from SGD through water by defining SGD exposure based on water source proximity and then limiting consideration to

[1] Currently, there are no federal regulations to limit externalities from unconventional natural gas extraction. While regulation of industry practices is more common at state and municipal levels, those that exist are mostly related to permitting of operations, where de facto regulation to deal with the external consequences comes through individual lease negotiations (Timmins and Vissing, 2015). If the negotiation process functions properly, then the Coase Theorem would guarantee the efficiency of bargaining outcomes. However, Timmins and Vissing (2015) find evidence that outcomes of lease contracts depend on characteristics of homeowners (race/ethnicity and language ability) even after controlling for willingness-to-pay attributes, which suggests that information asymmetries may cause efficient Coasian bargaining to break down.

infants born to mothers who are unlikely to be as threatened by exposure based on residence proximity. We then take several measures to investigate alternative drivers tied to SGD that could threaten the identification of health impacts. These include controlling for air emissions recorded by Toxic Release Inventory, distance to the nearest state road, and removing areas without active or historical coal mining. Taken together, our paper informs an important question of whether SGD affects reproductive health through the mechanism of drinking water.

We find consistent evidence that unconventional drilling impacts both drinking water quality and birth outcomes. First, our results indicate that drilling an additional well bore within 1 km of ground water source locations increases shale gas-related contaminants by 1.3 to 2.2 percent, on average. Impacts are driven by well bores that are up-gradient from ground water intake locations and by wells that have ever produced natural gas. The results are robust to multiple specifications and falsification tests. This suggests that shale gas development is systematically impacting public drinking water quality. These results are striking considering that our data are based on water sampling measurements taken after municipal treatment. Furthermore, our estimated water quality impacts are likely to be understated as they do not include sampling of chemicals that are *not* regulated under the Safe Drinking Water Act. These water quality results lead us to measure the infant health impacts of SGD through water contamination.

We find that a standard deviation increase in *in utero* exposure to well bores near water sources increases prematurity and low birth weight respectively by 0.42 and 0.37 percentage points (approximately equal to a 4 to 5 percent increase from the mean) for mothers who do not have a well adjacent to their home (i.e. no wells within 1 km). We also employ a mother fixed effect model to compare siblings and find somewhat attenuated impacts but still suggestive of an impact on gestation and birth weight for well bores drilled within 0.5 km of a CWS source. We include a number of robustness checks to assess these results. Of utmost concern is that SGD-related air quality changes could be driving our results (e.g. due to increased traffic and congestion that are not necessarily adjacent to gas wells). Our estimates are robust to inclusion of air quality controls that are proxied by either ambient air quality using the EPA's Risk-Screening Environmental Indicators (RSEI) Model or distance to the nearest state road, as well as when limiting estimation to a subgroup of infants that are far from coal mining areas. We also assess the extent to which our results are driven by mobility and fertility decisions in response to SGD and find no evidence of this with respect to SGD at water sources as measured by compositional changes in maternal characteristics or by the likelihood of moving. This suggests that moms are unlikely to suspect that public drinking water is affected by SGD, which is plausible given the hedonic literature that found only negative impacts for private ground water serviced homes (Muehlenbachs et al., 2015). Moreover, our finding that mothers do not respond to drilling at water source locations, but, interestingly, do respond if they live within 0.5 km of gas wells, bolsters our identification strategy using water source locations to define exposure.

SGD is taking place near a non-trivial portion of the population in Pennsylvania: 5 percent of births in our sample were exposed to at least 1 well bore within 10 km during gestation; a smaller share of 1 percent was serviced by community water systems with intakes that were in close proximity of a drilled well bore. Assessment of both the benefits and costs associated with SGD is much needed for input into policy-making in order to estimate the net welfare impact, a challenging task given the range of related benefits and costs. The most recent study finds that there is, on average, positive willingness-to-pay (WTP) for SGD-induced amenity changes after consideration of both economic benefits and amenity costs (Bartik et al., 2016). As WTP estimates reveal individual preferences that are formed based on what is currently known about SGD impacts, a better understanding of the external costs of these operations is crucial for the formation of such welfare estimates. Our paper contributes to an increasing body of research that estimates the causal impacts of SGD on the environment and health in order to weigh the extent of these potential costs against their economic and environmental benefits.

Our paper is organized as follows. Section 1 overviews shale gas development and related literature on ground water contamination and infant health. We next describe our data sources and provide summary statistics in section 2. Section 3 outlines our model and provides graphical evidence to support our empirical strategy. We present results in section 4. Finally, section 5 concludes.

1 Background

1.1 Shale Gas Development Overview

Shale gas is natural gas produced from organic shale formations (U.S. EIA, 2017). In Pennsylvania, shale gas development involves both vertical and horizontal drilling into, primarily, the Marcellus Shale, but more recently, the Utica Shale. After drilling, the process uses a technique to stimulate wells called hydraulic fracturing (HF), where water is pumped into a well bore at high pressure in order to fracture the shale beneath the ground. Shale plays are heterogeneous and so the distance drilled and quantity of water required differ across varied geological formations. On average, in Pennsylvania, it involves injecting 3-4 million gallons of water mixed with sand and fracturing chemicals into the well and using pressure to fracture the shale about 7,000 ft below the surface (ALL Consulting, 2009). The drilling process takes on average 5 weeks for a horizontal well in the Marcellus Shale and 15 days to complete the hydraulic fracturing process (NY DEC 2011). Before fracturing begins, a steel casing ranging from 1,000 - 4,000 feet is inserted and cement is poured into the area between the casing and the well bore (FracFocus.org, 2017). These measures are taken to both aid collection of natural gas and to protect water supplies by keeping fracturing fluids from migrating outward.

Hydraulic fracturing fluids are comprised of 99 percent water and proppant and 1 percent other chemicals used for a variety of purposes such as as viscosity adjusters, friction reducers, biocides, surfactants, scaling inhibitors, and acids to help dissolve minerals (U.S. Department of Energy, 2009; King, 2012). While many of the additives have low toxicity, some are known to be toxic and/or carcinogenic. In addition, approximately one-third of the chemicals studied by Stringfellow et al. (2014) do not have toxicity data. Of the HF fluid that is injected during the fracturing process, 5-50 percent is estimated to return to the surface as flow back or produced water (Byrnes 2011; King 2012),[2] which are then collected in impoundments and taken to be treated at a waste water facility (ALL Consulting, 2009).

1.2 Shale Gas Development and Ground Water Contamination

Water quality impacts is of top concern for this innovative process to recover natural gas. While there are numerous channels through which shale gas operations can impact water resources,[3] in the current literature, the two least controversial pathways of ground water contamination are faulty well casings or from abandoned nearby wells (Jackson et al., 2013; Lyverse and Unthank, 1988; Osborn et al., 2011).[4] The PADEP issued 90 violations in 2010 and 119 violations in 2011 for faulty casing and cementing. More controversial sources of ground water contamination are pathways between the shale formation and the aquifer, or if the drilling process occurs too close to a drinking water aquifer (Warner et al., 2012; DiGiulio et al., 2011). Migration of brine is

[2] Flow back is the fracturing fluid that quickly returns to the surface, and produced water is the fracturing fluid that takes longer to return to the surface (NY DEC 2011).

[3] For a comprehensive review, see Kuwayama et al. (2013).

[4] The PA DEP estimated that it only had records for 141,000 of 325,000 oil and gas wells drilled historically in the state, leaving the status and location unknown for approximately 184,000 abandoned wells (PA DEP, 2000). The likelihood of abandoned wells being conduits of groundwater contamination in Pennsylvania remains unknown at this time.

theoretically possible, but the likelihood remains debated in the literature (Myers, 2012; Saiers and Barth, 2012).

To date, there are only a few studies addressing ground water contamination concerns. One EPA study found that wells in Pavilion, Wyoming near drilling sites had elevated levels of methane, hydrocarbons associated with the shale play, and solvents used in the drilling process (DiGiulio et al., 2011).[5] Another study, using a sample of 60 private water wells in northeastern PA, found that drinking-water wells within a 1 km radius of a well head had methane concentrations 17 times higher than wells outside of the 1 km radius, with no measurable contamination of brine or fracturing fluids (Osborn et al., 2011).[6] The authors sampled an additional 81 water wells to enhance their previous findings and found methane in water wells 82 percent of the time, with concentrations 6 times higher for homes less than 1 km from a shale gas well (Jackson et al., 2013).[7]

Concerns over water quality impacts have led the Environmental Protection Agency on a six-year scientific assessment of the HF impacts on drinking water resources. In the final draft, released in December of 2016, the study concludes that HF activities *can* impact water resources. While this and the aforementioned studies find associations with water contamination, the results raise several additional questions that render the state of knowledge on water quality and SGD to be far from conclusive. For example, as many studies focus on private water wells, are the SGD impacts on water quality limited if applied to public water that arguably undergo more extensive treatment? Moreover, are the regional studies in a small portion of a state applicable more widely? Along with Hill and Ma (2017), we contribute additional evidence to this body of literature based on a state-wide investigation of SGD impacts on public drinking water quality.

1.3 Shale Gas Development and Infant Health

A growing body of literature has attempted to address the potential reproductive health effects of shale gas development. All of these studies are retrospective analyses of birth certificate records and focus on proximity to maternal residences as the definition of "exposure." In the epidemiological literature, studies use inverse distance weighted counts of wells within 10 miles compared to areas with zero wells (McKenzie et al., 2014; Stacy et al., 2015) or indexes of SGD exposure (Casey et al.,

[5] Due to mounting criticism regarding the report and the interpretation of its findings, USGS has released quality control well data with no interpretation and Pavilion, Wyoming is part of the large EPA study currently underway (Wright et al., 2012; EPA, 2012).

[6] The authors indicate that the presence of the well itself may be the conduit for methane migration, not necessarily the process of hydraulic fracturing.

[7] The authors also studied ethane and propane, two hydrocarbons that are only associated with gas extraction activities, and found that ethane was 23 times higher for homes less than 1 km from a gas well.

2015). Three studies from the economic literature have used a difference-in-differences approach (Hill, 2013b,a; Currie et al., 2017). Overall, these studies find an increased risk of low birth weight (Hill, 2013b; Stacy et al., 2015; Currie et al., 2017) and premature birth in Pennsylvania (Casey et al., 2015). Hill (2013a) also found an increased risk of low birth weight and premature birth in Colorado for residences within 1 km, while McKenzie et al. (2014) found an increased risk of birth defects for babies born near natural gas wells. None of these studies have assessed the likely mechanisms of exposure and this is where we contribute to this body of literature by directly assessing the effects of drinking water contamination from the industry on birth outcomes.

2 Data and Summary Statistics

We draw upon four main sources of data to produce a unique data set linking shale gas operations to infant birth outcomes through its impact on drinking water. We begin by describing the sources for our two main outcomes of interest, infant birth outcomes and water quality. We follow with the data that allows us to construct our main explanatory variable of interest, unconventional gas wells drilled. Our last data source describes public drinking water intake locations, which we use to link the explanatory variable to our outcomes of interest. We augment our main data with various other sources that we also list briefly at the end. After summarizing our data sources, we describe the data construction process and provide summary statistics for our estimation samples.

2.1 Data Sources

Birth Outcomes Confidential birth certificate records for the universe of births in PA beginning from 2003 through 2015 come from the Pennsylvania Department of Health (PADOH), where the maternal address associated with each birth is geocoded to longitude and latitude. Birth outcomes in the data include birth weight and gestational period (calculated from conception and birth dates). Demographic information of mothers such as race, age, marital status and education are available, as well as maternal health behaviors (e.g. smoking) and pregnancy risks.

Public Water System Water Quality Public drinking water quality and service areas are from the Pennsylvania Department of Environmental Protection (PADEP). Sampling and testing of drinking water are regulated by the PA Safe Drinking Water program for all public drinking water systems in Pennsylvania. The results of these monitoring efforts are recorded in the PADEP

Drinking Water Reporting System (DWRS). We obtained the water sampling data for all Community Water Systems (CWS)[8] from 2005 through 2015. Because information was electronically submitted by drinking water systems only beginning in 2011, we use the water measurements beginning in 2011 as our main estimation sample.[9] The data give detailed sampling information such as the sample date and time-of-day, the unique drinking water system identifier for which the sample was taken (PWSID), the contaminant sampled and its measured amount (in parts per million), and a variety of testing information (such as the testing method and laboratory). We supplement the DWRS drinking water data with additional information on whether the contaminants are indicative of SGD activity based on a list of chemicals that are associated with SGD (US House of Representatives, 2011; U.S. Environmental Protection Agency, 2016). We list these contaminants, the numbers of samples measured in Pennsylvania from 2011-2015 and whether they are SGD related because they are fracturing fluid or produced water chemicals in Table A1 of the Appendix. We also group contaminants into 12 categories (including a catchall group) based on categorizations used in the Remediation Technologies Screening Matrix and Reference Guide[10] and the EPA contaminant rules for public water systems.[11] Lastly, digitized public drinking water system maps from PADEP provide service area boundaries, which we spatially merge with maternal addresses to determine the public drinking water system on which a mother relies. Figure 1 shows these boundaries including the gas wells used in our analyses.

Unconventional Gas Wells Next, the Carnegie Museum of Natural History Pennsylvania Unconventional Natural Gas Wells Geodatabase (UNCGDB) provides unconventional natural gas well data through the third quarter of 2015 (Whitacre and Slyder, 2015). The UNCGDB unifies the major natural gas data sets made available by the PADEP and shows the life of each well from permit to production. The PADEP provides eight primary reports on natural gas well activity

[8]The EPA classifies a CWS as a subset of public water systems that supplies water to the same population year-round. See https://www.epa.gov/dwreginfo/information-about-public-water-systems.

[9]Table A2 in the Appendix shows that water measurements before 2011 for water systems that are eventually exposed to SGD are generally lower than systems that are not. As such, inclusion of pre-2011 water measurements would increase the magnitude of our results. However, because it is unclear whether the difference in water quality before 2011 is truly because water quality is better or if it is a result of switching recording systems, we take the conservative approach by using the water measurements beginning in 2011.

[10]The reference guide was compiled through an inter-agency effort to review available remediation technologies that can be used in hazardous waste cleanup projects. See https://frtr.gov/matrix2/top_page.html

[11]Broad categories include inorganic compounds (IOCs), halogenated or non-halogenated volatile organic compounds (VOCs), halogenated or non-halogenated semi-volatile organic compounds, synthetic organic compounds (SOCs), fuel contaminants, and radionuclides. Based on previous literature, we also separate out trihalomethanes (THM), pesticides, coliform, nitrates and nitrites, acids, and disinfectant byproducts (DBPs).

to the public.[12] All reports can be unified through a unique well permit number, known as an American Petroleum Institute (API) number. We primarily make use of the Oil & Gas Locations report, which identifies the longitude and latitude coordinates of unconventional natural gas wells, and the Permits Issued and Spud Data reports, which respectively provide the date on which a well bore was permitted and the date on which drilling began (or will begin).

Public Water System Source Intakes Our last set of data gives a snap-shot of near ∼8000 ground water intake locations in Pennsylvania as of 2015, which we build from the PADEP website. Crucially, these data allow us to link shale gas operations at well bores to both the quality of water provided by CWSs as well as infants that are born to mothers that rely on public water provided by those systems.

Other Sources In addition to the data discussed above, we draw upon several other sources to augment our main data set and check for robustness. We briefly list these sources here and discuss, in a later section, how we utilize each within the context of our empirical strategy. These other sources include average daily temperature and precipitation from Schlenker and Roberts (2009), water sampling data from U.S. Geological Survey (USGS) ground water monitors,[13] digitized elevation maps from the USGS National Elevation Dataset,[14] digitized maps for PA coal seams and state roads, both of which are provided by Pennsylvania Spatial Data Access (PASDA),[15] and US EPA Toxic Release Inventory (TRI) data.[16]

2.2 Data Construction and Summary Statistics

We build two main data sets to investigate our question of interest: 1) an infant health data set in which a unit of observation is a birth, and 2) a water quality data set, where a unit of observation is a contaminant sampling measurement from the DWRS.

Beginning with all community water systems that draw from ground water sources, we first map all births to these water systems based on maternal address and the spatial boundaries of

[12]These are Permits Issued, SPUD Data, Production Reports, Waste Reports, Compliance Reports, Public Utility Commission (PUC) Act 13 Unconventional Wells Spud Report, PA DEP Oil & Gas Locations - Conventional Unconventional (hosted by PA Spatial Data Access), and Well Formations Report.

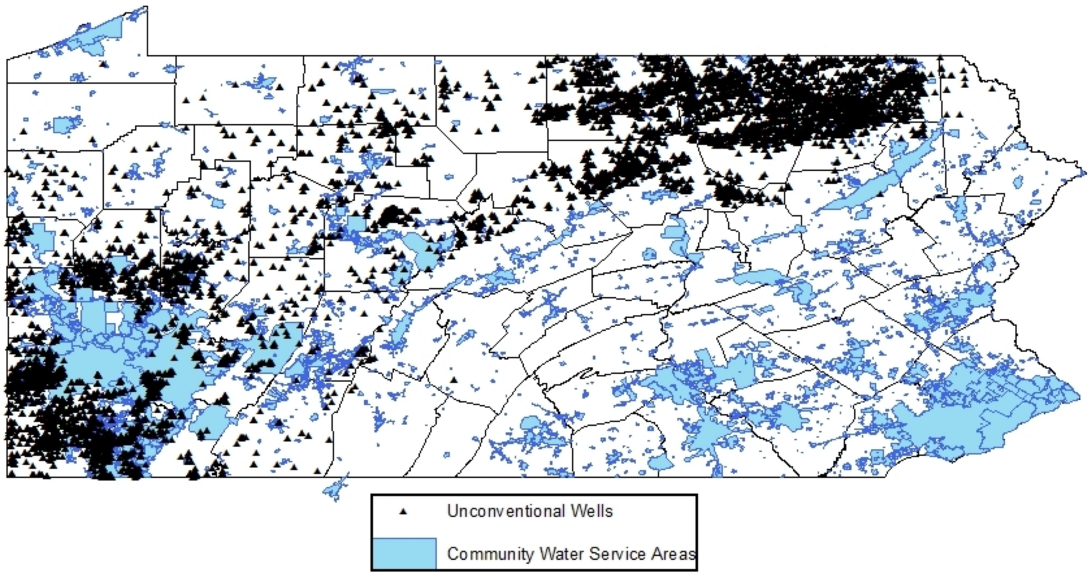
[13]Monitor locations and water quality measurements are downloaded from https://www.waterqualitydata.us/portal/.

[14]Raster data are downloaded from https://lta.cr.usgs.gov/NED

[15]Shapefiles are downloaded from http://www.pasda.psu.edu/.

[16]We assessed the Basic Data Files from https://www.epa.gov/toxics-release-inventory-tri-program/tri-basic-data-files-calendar-years-1987-2016

Figure 1: Gas Wells and Community Water Systems



the service area using Geographic Information Systems (GIS) software. Next, we spatially link the sources of these CWS to all well bores within a large buffer of 10 km.[17,18] Figure 1 overlays Pennsylvania natural gas wells and community water systems. As evidenced by the spread of natural gas wells in the figure, the Marcellus shale play stretches from the southwest corner of the state to the northeast. As such, regions exposed to SGD will be predominantly rural, and comparisons of either births or water quality in these areas with that in cities (i.e. Philadelphia) would be inappropriate. The importance of limiting the comparison to births in CWS with intakes within 10 km of a gas well is apparent in Table 1, which compares characteristics of the sample of mothers within 10 km of drilling to those from the population. Out of a total of 1.50 million births in PA through the third quarter of 2015,[19] 115,690 are born in CWS's with ground water intakes that are within 10 km of gas wells, and 340,827 are born to mothers who live within 10 km of any gas well. Maternal characteristics for these groups differ from the population in a number of important ways that may affect birth outcomes: mothers exposed to SGD, whether it be through intake

[17] A system can have multiple intakes. In this case, we consider the system within the vicinity as long as any one of its intake locations are within the 10 km buffer.

[18] We limit our investigation to ground water systems because we do not have surface water protection areas, which would delineate the exposure area to surface water systems. We abstract from these systems for the purposes of a cleaner exposure definition since surface water exposure areas can be large and can vary in exposure range (e.g. the exposure area for a pond versus a river), and leave investigation of surface drinking water impacts for future work.

[19] There are approximately 1.84 million births in PA between 2003-2015. After removing births to mothers with missing characteristics or birth outcomes, we are left with 1.53 million births, of which 1.50 million are before the last quarter of 2015.

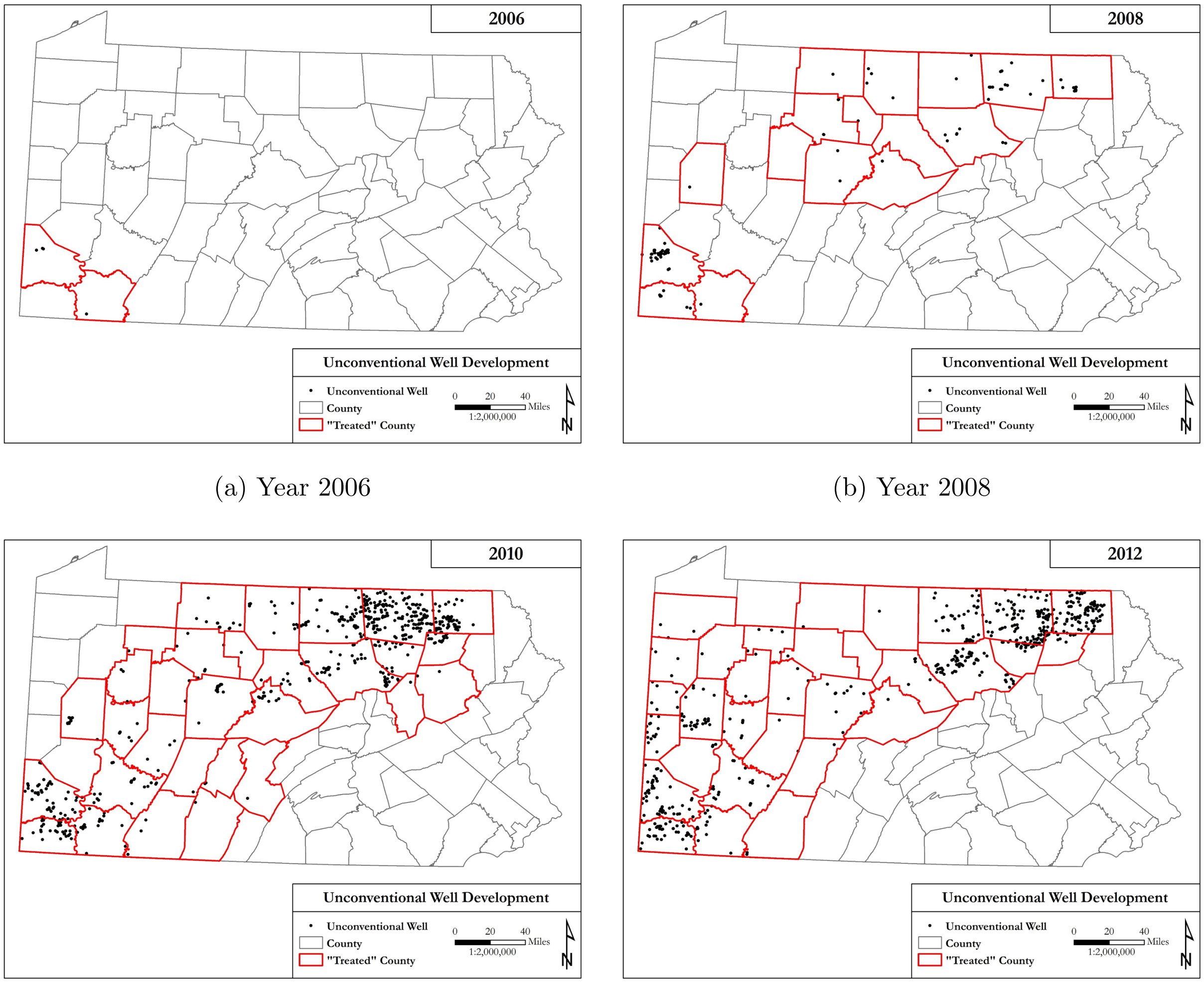
or residence proximity, are much less likely to be black or Hispanic, and are more likely to have been born in PA, to have smoked during the pregnancy, and to have paid for hospital services with Medicaid and participated in the Special Supplemental Nutrition Program for Women, Infants, and Children (WIC).[20] Crucially, this suggests additional, unobserved differences between births who are exposed to SGD and the population more generally. As such, we retain all births in community water systems (CWS) whose ground water intakes lie within 10 km of a gas well.

With each birth spatially linked to every well bore within 10 km of its intake, we then attribute shale gas activity of well bores to infants by aggregating the total number of well bores within 10 km of the infant's CWS intake that were drilled *within the gestation period* of that infant. We additionally aggregate these 'threats' at various buffers within 10 km (i.e. between 0.5 and 2 km) to distinguish the impact of threats at different intake proximities. We do this as there are typically multiple well bores (drilled at different times) located near any given groundwater intake, and as such, no clear 'before' or 'after' exposure period. While this complicates our definition of a treatment period, it provides good variation in exposure to shale gas operations that one can exploit. Figure 2, which delineates the new well bores drilled and the affected counties by year, is indicative of this as drilling varies both on the extensive and intensive margins.

We use a similar procedure to construct our water quality data for water measurements beginning from 2011 through the third quarter of 2015. Upon limiting the sample of contaminant measurements to water systems within 10 km, we take each DWRS water measurement and aggregate the total number of well bores within 10 km of the CWS intake (and various proximities within) that have been drilled by the time that water measurement was taken. We remove samples that are greater than the 99th percentile of the sampling result distribution to prevent outliers from driving our results. Directing our attention to the set of contaminants that have been associated with SGD: of the 171,615 water measurement observations from systems within 10 km of CWS intakes, approximately 40% (or 69,239) are of contaminants that have been tied to SGD. For this SGD-related sample, Table 3 finds that there are, respectively, 0.18, 0.45, and 27 well bores drilled, on average, within 1, 1.5 and 10 kilometers of intake locations.

[20]This sample limitation is similarly important for water quality: Table A2 in the Appendix gives CWS characteristics for systems with intakes within 10 km of drilling compared to systems without. Water systems with intakes near well bores serve much smaller populations. Furthermore, a comparison of the average results in each contaminant category before 2009 shows that the systems outside of 10 km of well bores generally have higher levels of contaminants than those found in the water systems within 10 km.

Figure 2: New Well Bores Drilled by Year



(a) Year 2006

(b) Year 2008

(c) Year 2010

(d) Year 2012

3 Empirical Strategy

Our baseline specification follows a difference-in-differences (DD) approach that compares changes in birth outcomes in response to drilling for infants born in systems with drinking water intakes near well bores to similar changes for infants in systems with intakes that are farther away but still within 10 km. Specifically, we regress an indicator (Y_{ijt}) of either low birth weight (weight $<$ 2500 grams) or prematurity (gestation length $<$ 37 weeks) for a birth i in CWS j on the number of well bores within x km and 10 km of the CWS intake that are drilled during the infant's gestation period:

$$Y_{ijt} = \beta_1 \sum_{k}^{K} d_{jk}^{<x} w_{kt} + \beta_2 \sum_{k}^{K} d_{jk}^{<10} w_{kt} + X_{it} + \nu_{jt} + q_t + cws_j + \epsilon_{ijt} \quad (1)$$

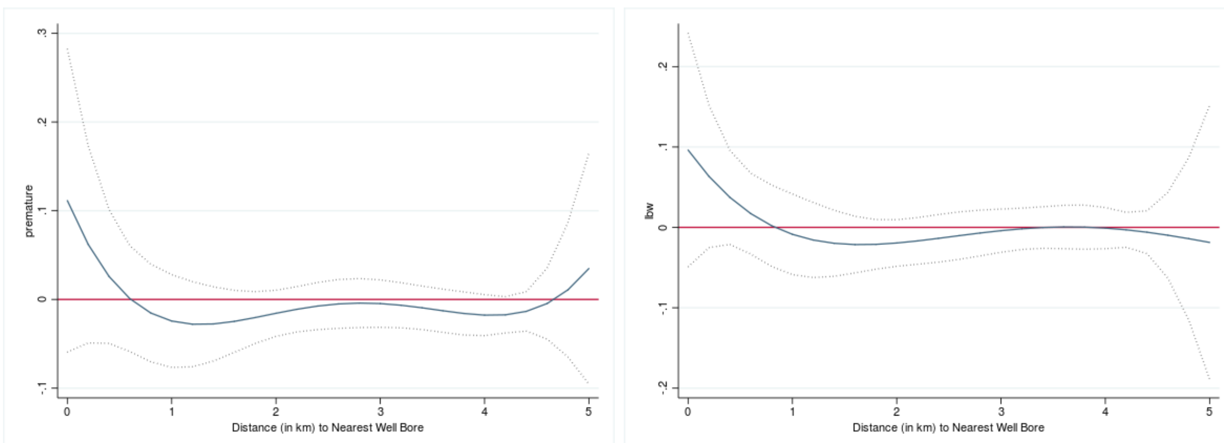
The indicator w_{kt} is equal to 1 if well bore k was drilled during the gestation period of the infant, and $d_{ik}^{<x}$ is an indicator for intake proximity at a buffer of x km. The interaction of these terms and our variable of interest, $\sum_k^K d_{jk}^{<x} w_{kt}$, sums over all K well bores within 10 km of an infant's intake, giving the total number of well bores within x km that are drilled during gestation. The corresponding variable at 10 km, $\sum_k^K d_{jk}^{<10} w_{kt}$, returns the exposure to wells within 10 km of source intakes, capturing air exposure from trucking activity, etc. We divide these variables by the standard deviation of well exposure during gestation at 1 km (0.1) so as to better gauge the magnitude of the impacts. The main coefficient of interest β_1 returns the change in birth outcome given a standard deviation increase in exposure to well bores within x km of source intakes. Importantly, this estimate controls for changes in birth outcomes that would have happened in absence of drilling near water sources, which is captured by more distant drilling activity that is still within 10 km, $\sum_k^K d_{jk}^{<10} w_{kt}$. Controls for maternal characteristics, X_{it}, include the mother's age, race, education, Medicaid and WIC enrollment at birth, and a host of pregnancy risks (e.g. pre-gestational diabetes and smoking).[21] We also include the following controls: temperature and precipitation near maternal residence, which can directly impact birth outcomes (Deschênes et al., 2009) as well as moderate exposure to water contaminants; a direct measure of changes in water quality that is *not* related to SGD, which is the number of coliform and disinfectant by-product exceedances of federally established legal limits during gestation; air quality controls as proxied by TRI emissions. In addition, we include county-by-year fixed effects (ν_{jt}), month-by-year fixed effects (q_t), and a fixed effect for each CWS, cws_j. These help to control for differential trends in birth outcomes across regions, seasonal differences in birth outcomes, and unobserved differences across water systems that might impact health. As there can be unobserved neighborhood characteristics within public water systems that could impact health outcomes (e.g. access to public facilities such as parks), we remove other time-invariant neighborhood differences at the ZIP code level by using

[21] Specifically, controls for maternal characteristics include dummy variables for mother's age group (19 to 24, 25 to 35, and 35 or older), race/ethnicity (Hispanic or black), educational attainment (high school only, some college, associates degree, and college or more), marital status, WIC enrollment at birth, and Medicaid payment. Controls for pregnancy risks include number of cigarettes smoked in the last trimester, parity, as well as dummy variables for smoking habits (separate indicators for ever smoked in first, second, and third trimester), pre-gestational and gestational diabetes, having poor outcome for previous birth, vaginal bleeding, previously had a pre-term birth, and infertility risk. In addition to maternal characteristics, we control for gender of the infant.

ZIP code fixed effects in a second specification. Still, comparisons within ZIP codes cannot control for differences in family background, an important contributor to health. We attempt to limit unobserved differences in family backgrounds by also estimating health impacts with comparisons within siblings, i.e. through the use of mother fixed effects.[22]

Figure 3 presents graphical evidence of drilling impacts on birth outcomes using a simplified definition of spatial exposure that only looks at the nearest well bore to the infant's CWS intake. Specifically, the figures plot, by distance to the well bore, the difference between birth outcomes of infants whose nearest well bore was drilled during gestation and that of infants who were not exposed during gestation. The figures show a large increase in the incidences of premature births

Figure 3: Birth Outcome Impact by Distance between Intake and Nearest Well Bore



(a) Prematurity

(b) Low Birth Weight

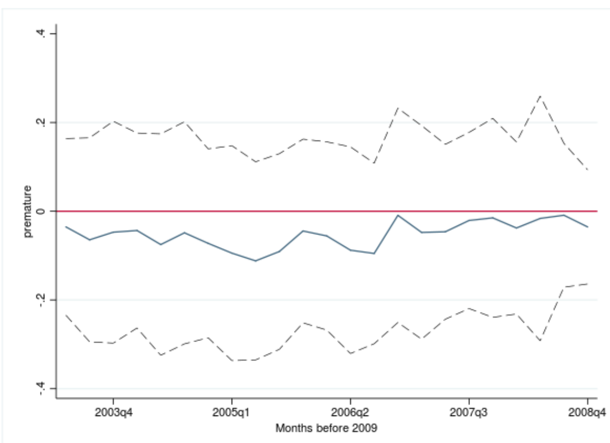
and low birth weight for those exposed during gestation relative to those who are not exposed, where the impacts are around 0.1 percentage points (pp) and decrease continuously until approximately 1 km. While these impacts are imprecisely estimated, we note that there are multiple well bores near a CWS intake and that the simple measure of exposure employed in these figures may be inappropriate. We thus rely on our regression specification to capture infant exposure to multiple wells, which can be drilled both at different distances from the intake and on different dates.

For the relationship in the figures to be a causal one rests on the assumption that birth impacts captured by drilling activities that are 'far' from intakes over the same period represent changes in infant health that would have occurred in the absence of drilling near the intake as water quality

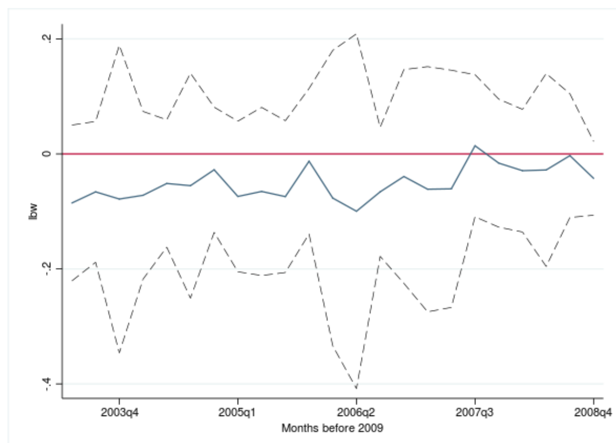
[22] We note that inclusion of mother fixed effects does not avoid other forms of time-varying endogeneity (e.g. delaying fertility); moving out of state so that we do not observe a second birth or miscarriage that could be due to exposure (Grossman et al., 2017).

would not have been impacted. Since graphical evidence potentially point to water quality impacts within 1 km, we can separate our infants into a treatment and control group based on a 1 km exposure buffer to check for pre-existing trends in birth outcomes before SGD began. Figures 4a and 4b respectively plot the quarterly averages of prematurity and low birth weight comparing infants with source intakes within 1 km of well bores to those on systems that are farther away from well bores. For the most part, point estimates suggest that birth outcomes in areas that would be exposed to SGD within 1 km are better than those that are farther away, although the relative differences are not statistically different from 0. Relative birth outcomes over time are generally stable, except for a slight reversion to 0 in birth outcome between 2007 and 2008, which coincides with the Great Recession. This could be cause for concern if infant outcomes for our treatment and

Figure 4: Pre-Existing Trends in Birth Outcomes



(a) Prematurity



(b) Low Birth Weight

control groups may begin to trend differentially over time at the onset of SGD due to factors that are unrelated to water contamination. Previous investigations have shown that macroeconomic conditions affect fertility (Adsera, 2004, 2011; Currie and Schwandt, 2014), and a recent paper by Kearney and Wilson (2017) finds fertility impacts for SGD specifically; moreover, the responses to economy-wide changes could differ by age, socioeconomic status, and employment status (Schultz, 1985; Becker, 1981; Gustafsson and Kalwij, 2006). With the shale gas boom coinciding with the recovery of the recession, one might expect responses to changing macroeconomic environments in the form of selection into childbearing to be different between areas near and far from drilling. To deal with this, we include a set of controls for the number of permitted well bores during an infant's gestation both near (within x km) and far (within 10 km) from the CWS source. As we show in the water quality results later in the paper, permitting does not impact water quality, and thus any response of infant outcomes to permitting activity should be unrelated to water quality changes. Under the assumption that the relative change in infant outcomes for those near permitting of well bores is a valid counterfactual for birth outcomes of infants affected by SGD had they not been exposed to drilling, then inclusion of permits captures the correlated trends in birth outcomes over this period of time.

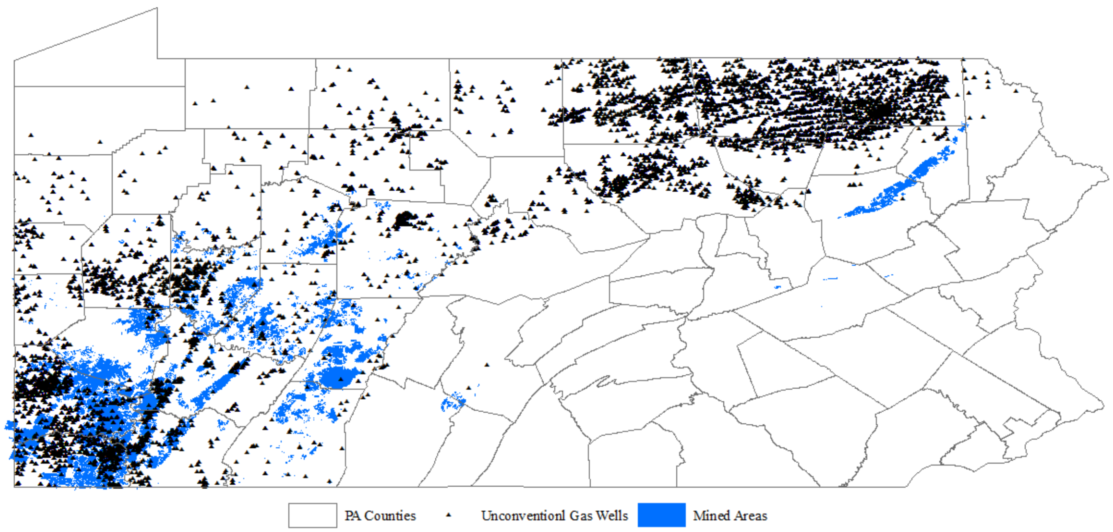
Finally, we augment our baseline specification to ensure that the impacts we are recovering are through the mechanism of water contamination. Of most concern is that our estimated infant health impacts could be driven by changes in air quality (Chay and Greenstone, 2003; Currie and Walker, 2011). The negative impacts on health from other media of contamination that would most affect mothers living in close physical proximity to well bore activity would cause us to overstate the impacts of water quality changes. Of course, there are potential benefits from living in close proximity of drilling activity if a household receives royalties or lease payments for allowing drilling on its property. We control for these 'adjacent' impacts, which combine both the costs and benefits from living near well bores, by additionally mapping the distance between maternal residences and well bores, and estimating all models with a subsample of mothers who live at some minimum distance from the nearest well bore, i.e. mothers who are threatened by drinking water intake proximity to gas wells but not by physical well bore proximity. We take this minimum distance to be 1 km, following previous work that have found proximity impacts on infant outcomes within this buffer (Hill, 2013a; Currie et al., 2017).

We next address air quality concerns more directly by creating several controls to capture potential air quality impacts on birth outcomes. First, we calculate U.S. EPA Toxic Release Inventory (TRI) on-site releases in the vicinity of the maternal address.[23] Since the amount of TRI releases does not take into consideration the toxicity of contaminants, we additionally use an alternative measure of ambient air quality at the block-group-by-year level that is calculated from TRI data using EPA's Risk-Screening Environmental Indicators (RSEI) Model.

As Pennsylvania has had a long history of coal mining dating back to the 1920's, a concern that arises is that our estimates are picking up the impact of these activities, which oftentimes coincides with areas that are currently engaged in SGD. Figure 5 overlays unconventional gas wells with all coal seams in Pennsylvania that have ever been mined as of 2017. The mined areas shown include

[23]In practice, we include 1) the number of TRI sites within 1, 3, and 5 kilometers of the maternal address, and 2) an inverse distance^2-weighted average of total on-site releases within 10 kilometers of the maternal address.

Figure 5: Coal Seams as of 2017



those that were historically mined but no longer active, currently active mines (as of 2017), as well as areas for which the last-known mining date is unknown (assumed to be inactive).[24] Immediately, one can see that there is a fair amount of overlap between SGD and coal seams in the southwest region of the state, which is cause for concern. To assess the extent to which our estimates are driven by mining, we identify the public water systems that have any drinking water intakes within 1, 5 and 10 km of historical coal seams or any coal seams (active or historical), and estimate our model on a sub-sample that removes infants belonging to any of those groups with the idea being that by removing observations that are exposed to coal seams, we can limit the impact of coal mining on our estimates.

Lastly, SGD-related transport is hypothesized to increase air pollutants; we control for the distance between maternal address and the closest PA state-owned and maintained public road to assess the possibility that our results are caused by traffic-induced air quality changes.

An important limitation of the above specifications is the inability to deal with avoidance behavior. If individuals are aware of water contamination risks, then they could take various measures to protect themselves from switching to bottled water (Graff Zivin et al., 2011; Wrenn et al., 2016), altering fertility decisions (Kearney and Wilson, 2017), or moving (Banzhaf and Walsh, 2008). If mothers engage in such avoidance behavior, then our estimates would be biased

[24]This data only provides the date on which mines are no longer active, but not the start dates. We make an assumption that all mined areas that are active or no longer active during our sample period have been active as of the start of our sample period (2003).

against finding an impact. While it is difficult to control for these confounders, we assess avoidance behavior by estimating birth outcome impacts using subgroups based on socioeconomic status (e.g. education and Medicaid use). If individuals are indeed taking measures to mitigate exposure, then we would likely see the largest negative impacts on health to be concentrated among the low SES groups, those with arguably less ability to take costly avoidance measures. In addition, we test whether maternal characteristics change in response to increasing well bore threats, as well as whether drilling increases the probability that a mother moves to assess the extent to which fertility decisions and migration drive our results.

3.1 Water Quality Impacts

Whether SGD has impacted birth outcomes by affecting drinking water quality requires understanding whether drinking water is actually impacted. Currently, there is no consensus regarding this 'first stage' question from the scientific community, as there is only a handful of studies examining SGD's impact on private drinking water quality and no work, aside from Hill and Ma (2017), that have investigated its impact on public drinking water quality. As such, establishing this relationship is an important, necessary step to asking the question of whether SGD impacts health through water; if no direct water quality impacts exist, then the scope for SGD's impacts to be mediated through water would be indeed limited.

The model to estimate water quality impacts builds upon previous work in Hill and Ma (2017) and follows that for infant health closely. Our specification is again a DD approach that compares water quality changes in response to drilling for systems with intakes near well bores to water quality changes over the same time period for systems with intakes that are farther away but still within 10 km. Specifically, we model the logarithm of water quality measurements i, r_{ijt}, for a community water system j to depend on all K well bores within 10 km that have been drilled by the sampling date. We then allow well bores (indexed by k) within a smaller buffer of x km of an intake ($d_{ik}^{<x} = 1$) to have an additional impact compared to those that are within 10 km but outside of x km ($d_{ik}^{<10} = 1$). The regression controls for sample-specific attributes (X_{it}) such as hour-of-day of when a sample was collected, the laboratory at which sampled results were measured, the contaminant group to which a pollutant belongs, and temperature and precipitation. We also include county-by-year fixed effects (ν_{jt}), quarter fixed effects (q_t), and a fixed effect for each CWS,

cws_j. The following gives our baseline specification:

$$r_{it} = \beta_1 \sum_k^K d_{jk}^{<x} w_{kt} + \beta_2 \sum_k^K d_{jk}^{<10} w_{kt} + X_{it} + \nu_{jt} + q_t + cws_j + \epsilon_{it} \quad (2)$$

where $w_{kt} = 1$ if well bore k has been spudded by time t. The parameter of interest, β_1, returns the impact of drilling an additional well bore on SGD-related contaminants, in addition to changes in water quality trends over the same period as captured by water quality changes of CWS with more distant gas well threats, $\sum_k^K d_{jk}^{<10} w_{kt}$.

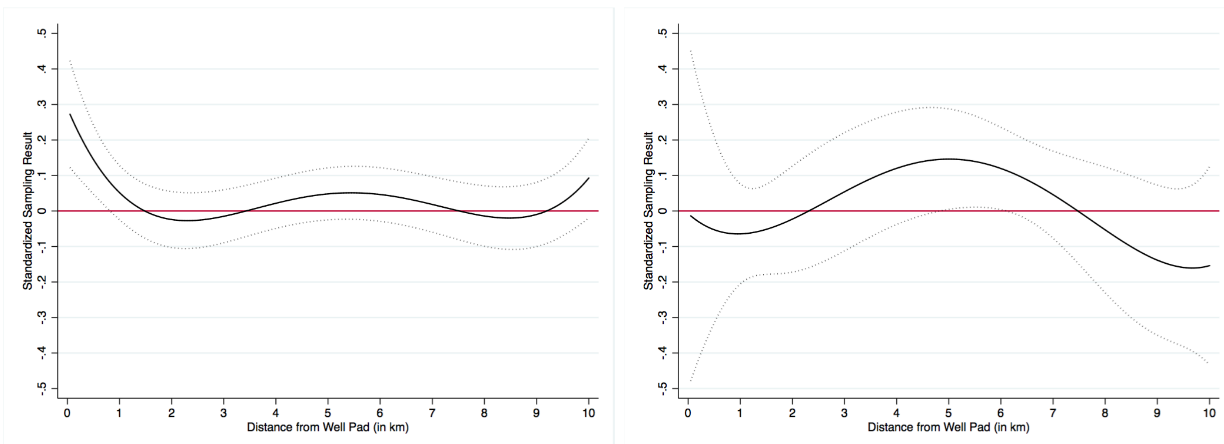
We can explore heterogeneity of effects by distinguishing the impacts from well bores that are drilled uphill versus downhill from intakes, and those that ever produce any oil or gas as opposed to never-produce. In each case, the total number of threats within a certain proximity can be decomposed into those from each type of threat for a given way of distinguishing threats,

$$\sum_k^K d_{jk}^{<1} = \sum_k^K \left(TypeA_{jk}^{<1} + TypeB_{jk}^{<1} \right) \quad (3)$$

where 'TypeA' and 'TypeB' would refer to, for example, the number of up- and down- gradient threats within 1 km when separately estimating impacts by elevation. Gas well threats are defined to be 'uphill' from a ground water source if the surface elevation of the well bore is higher than the surface elevation at the source intake. If elevation affects ground water flow, one would expect uphill threats to have stronger impacts on drinking water quality than those down hill of intake wells. Unproductive wells are typically left inactive because the cost is often prohibitive to permanently plug and abandon the wells (Muehlenbachs, 2015). A priori, we do not know whether producing wells are more likely to contaminate nearby drinking water sources than wells that are just drilled and never produce. Separately testing these dimensions not only serves as robustness checks, but provides insight into potential mechanisms of contamination. We additionally perform a number of placebo tests and estimate the impact of gas well threats on non-SGD related chemicals as well as the impact on SGD-related chemicals of gas well threats that occur after water measurements are taken. As with the birth outcomes model, we estimate our water model after removing water measurements from water systems that have ever been exposed to coal mining. Furthermore, we build an additional data set that matches water sampling data from U.S. Geological Survey (USGS) ground water monitors to gas wells. Construction of the data follows the same procedure as that used for public water system water quality, except the water sampling data is matched to gas wells via the location (i.e. longitude and latitude) of the USGS water monitor. These checks would further bolster our case that our estimated impacts are, in fact, causal.

We present graphical evidence of water quality impacts in Figure 6a, which plots the difference in water sampling results before and after the nearest well pad was spudded as a quartic function of distance to the nearest well pad, controlling for county-by-year, quarter, and CWS fixed effects.[25] Sampling results are standardized by their respective contaminant categories so that each has mean 0 and a standard deviation of 1. Figure 6a shows that SGD-related contaminant levels after drilling are up to 0.25 standard deviations larger compared to samples taken prior to drilling. The magnitude of the differences decreases continually until 2 km, but is no longer significantly different at about 1 km. We further find no evidence that the water quality of our 'near' systems (i.e. those

Figure 6: Water Quality Impacts by Distance between Intake and Nearest Well Pad



(a) SGD Chemicals (b) Coliform

with intakes within 1 km) had been worsening relative to those that are 'far' prior to 2009 (shown in the Appendix), indicating that these impacts are not a result of pre-existing trends.[26] We contrast the figure for SGD chemicals with the distance gradient for coliform (Figure 6b), a contaminant that is unlikely to be associated with SGD. Coliform sampling exhibits no clear relationship with

[25] As there are typically multiple well bores located near any given ground water intake, complicating our definition of exposure, we examine the impacts of the nearest well pad (groups of well bores within 1 acre of each other) for this graphical exercise. We account for activity at all well bores nearby in the regression analysis.

[26] If there was evidence of this, it suggest that we are attributing water quality impacts to SGD when, in fact, the impacts would have likely occurred even in the absence of the shale gas boom. Figure 1 in the Appendix plots quarterly averages of SGD-related chemicals before 2009 for systems that are within 1 km relative to those that are not. The relative difference is estimated to be negative at all quarters, signaling that, before any drilling began, CWS that would eventually be exposed to SGD within 1 km had better water quality relative to our control group. Moreover, the quarterly averages do not exhibit any discernible trend over time, lending some support for the DD identification assumption of common trends that allows one to infer that these exposed systems would have continued on a path similar to that of the control group had they not been exposed to SGD activity.

well pad distance. Taken together, these figures suggest water quality is potentially impacted by drilling, and that the largest impacts are within 1 km.[27]

4 Results

4.1 Water Quality Impacts

Table 4 reports our main water quality results. Each column within a panel is a separate regression. In all regressions, the comparison group is community water systems located within 10 km of a well bore, whereas the definition of the exposure buffer changes across regressions as indicated by the column headings. These results restrict the sample to the set of contaminants that are considered related to SGD. The dependent variable is the log of the sampling result so that the coefficients can be interpreted as percent changes. In all regressions, we control for system fixed effects, county-by-year fixed effects and quarter fixed effects in addition to sample-specific characteristics as discussed previously.

For the overall impacts on the sample of contaminants related to shale gas (Panel A), we find that drilling an additional well bore within 0.5 km of intake locations increases average sampling of contaminants by 1.96 percent (5% level of statistical significance). As the buffer is relaxed to 1 km, impacts decrease to 1.25 percent. However, well bore threats at distances farther than 1.5 km are an order of magnitude smaller and are not statistically significant, indicative of no effect. This is intuitive as we would expect that systems with intakes further away from well bores are less likely to be affected by surface spills or activity that might impact ground water. Most of the scientific papers to date investigating ground water impacts use a 1 km buffer (Yan et al., 2016) and our graphical analyses indicate that most of our anticipated impact would be within 1 km.

The negative effects of well bore threats on drinking water systems may vary depending on the characteristics of the gas well. Equation 3 augments the baseline specification to explore heterogeneity. Panel B of Table 4 displays these results. First, because elevation affects groundwater flow, we differentiate the water quality impacts of uphill well bores from those downhill. Unsurprisingly distinguishing by up- as opposed to down-hill threats, we find evidence that it is the uphill threats that are disproportionately affecting drinking water quality. The estimated impacts of uphill bores

[27]While we have no a priori belief as to the maximum distance a well bore's effluent can contaminate a water source, the distance of 1 km is in the range of what has been found in previous literature. Vengosh et al. (2014) have found gas wells drilled within 1 km to as far as 5 km of intakes to pose a risk to drinking water quality; ground water contamination risks, as measured by housing price impacts from Muehlenbachs et al. (2015), have also been perceived up to 2 km from well pads for houses relying on non-public water systems.

are 2.16 and 1.33 percent for threats within 0.5 and 1 km, respectively (5% level of statistical significance). We next examine whether the adverse effect of well bore threats on drinking water quality is driven by producing well bores. In panel C of Table 4, we find that the effect of an additional well bore threat is driven primarily by producing wells, where the estimated impacts of producing wells within 0.5 and 1 km are respectively 2.16 and 1.33 percent.

We perform a number of placebo regressions in Table 5. In panel A, we estimate our baseline using only chemicals that are unlikely to be related to SGD. If our estimated water quality impacts are due to changes in the environment that are correlated with shale gas, then one would likely see increases in non-SGD related chemicals as well. We do not find much evidence of this, and in fact see that overall water quality within 1 km is somewhat improving over time in these systems. Panel B estimates the impact of well bores that are drilled after the water measurements were taken (i.e. threats incurred in the future) and find that there is no effect of future threats on drinking water quality, which would be the case if our estimates are causal. In panel C, we examine the impact of well bores that were permitted but never drilled and again we do not find any effects. These placebo regressions confirm our findings by showing that our results are not just spurious correlations.

Table 6 estimates SGD's impact on water quality by removing water measurements from water systems whose source intake location is within 1, 5 or 10 km of a historically active or any coal seams. The estimated impact after removing samples from CWS with intakes within 1 km of historical coal seams is 1.81 and 1.19 percent (5% level of significance) for well bores drilled within 0.5 and 1 km, respectively. These are slightly smaller, but qualitatively and statistically similar to our baseline estimates that include samples near coal mining activities (1.96 and 1.25 percent for bores within 0.5 and 1 km). When we additionally remove samples near areas with active coal mining, the estimated impacts of 0.5 and 1 km well bores are smaller at 0.67 and 0.66 percent. This would suggest that there may indeed be some interactions between SGD operations and existing coal seams that could impact water quality, especially in areas experiencing active coal mining. Still, none of the estimates are statistically different from each other. Moreover, when removing all samples near any coal seams, there is not a clear dose-response relationship when moving farther from areas with coal as one would anticipate if coal is the primary cause of water quality degradation.

We examine alternative characterizations of water quality measures in Table 7. Panel A switches the dependent variable (in order from left to right) to be detection (i.e. an indicator of a positive readings[28]), an indicator for exceeding the Maximum Contaminant Level goal (MCLg), an indicator for exceeding the Maximum Contaminant Level (MCL), the sampling result, and the result that is standardized to have mean 0 and 1 within its respective contaminant group. We find a 3.11 pp increase in the probability of detection, which is approximately a 13% increase from a mean detection rate of 24%. We similarly find a 12% increase in exceedance of an MCL goal, which is unsurprising as many of the goals for these contaminants are 0. We do not find a measurable impact on MCL exceedances, a more stringent and legally binding threshold for contaminants. This suggests that many of the contaminant increases that we detect are not large enough to trigger MCL violations from the state water authority. Finally, switching our dependent variable to be either the continuous sampling result or a standardized version does not qualitatively alter our results. Panel B estimates detection by individual chemical groups. We find increases in the group of inorganic compounds (5% level of significance) and synthetic organic compounds (10% level of significance), but no impact on other chemical groups that are statistically significant at conventional levels.

Finally, we explore the robustness of our results by examining water quality changes near gas wells as measured by ground water monitor data from USGS. The same specification is used as before, where controls for weather and contaminant group indicators are included, as well as fixed effects for quarter and county-year. Table 8 gives the impacts of an increase in number of well bores drilled on SGD related chemicals (Panel A) and Non-SGD related chemicals (Panel B). SGD chemicals increase by approximately 3% (statistically significant at the 1% level) in response to drilling within 1 km of water monitors relative to changes in chemicals farther away (i.e. more than 1 km and within 10 km) from gas wells. This impact is much smaller and not statistically significant for non-SGD related chemicals. We find results that are consistent with that for public water systems when separating impacts by elevation and well production.[29] That the impacts from water quality monitors are somewhat larger suggests that public water systems are partially successful at mitigating the impacts of water contaminants. Furthermore, some of the chemicals measured in the water monitoring data are not present in the public drinking water data (e.g. bromides and chlorides) because they are not regulated under SDWA. This suggests that the drinking water quality impacts based on measurements of regulated contaminants is likely to understate the total impact on drinking water quality.

[28]Chemical measurements that are below the limit of detection, i.e. 'non-detects', are coded as 0's in our data. This is generally more optimistic than the truth, in which case our estimates would be biased downward.

[29]We provide these results in Table A3 of the Appendix. We again find that the effects are concentrated at bores up-gradient from water monitors and those that produce natural gas.

4.2 Infant Health Impacts

The finding from the previous section of SGD impacts on water quality through exposure of CWS intake location opens the potential for health impacts. We present the average impacts on the incidence of prematurity and low birth weight for a one standard deviation increase in the number of drilled well bores in Tables 9. Standard errors are clustered at the CWS level for these and all remaining specifications. Panel A estimates our models without restrictions on the distance between a mother's residence and well bores, whereas panel B removes all infants born to mothers who live within 1 km of at least one well bore. We omit results using ZIP code fixed effects as they are similar to that estimated using CWS fixed effects.

Without restrictions on residence proximity (panel A), our CWS fixed effects models find that a standard deviation increase in well bores drilled within 1 km increases prematurity by 0.265 pp (5% significance level). The 1 km impact on incidence of low birth weight is similar, but muted, with an impact of 0.199 pp. When removing infants that are directly 'adjacent' to SGD activity (panel B), the magnitude of impacts is slightly larger: At 1 km, average impacts on prematurity is 0.418 pp (statistically significant at the 1% level). This impact increases for gas well threats at 0.5 km to 1.08 pp (significant at the 10% level). With an average rate of prematurity at 9%, the impact for a one standard deviation increase in well bores drilled within 1 km (or 0.5 km) is about 4.5% (or 11.6%), on average. Results for low birth weight are similar, with impacts of 0.365 and 0.555 pp (or about 4.7% and 7.2% given a 7.7 percent incidence of low birth weight) for gas well threats within 1 and 0.5 km of CWS intakes. We note our estimated impacts of well bore threats within 0.5 km for low birth weight are not statistically significant at conventional levels even though the magnitude of effects are larger than 1 km threats. That our estimated birth outcomes impacts increase when removing mothers who are adjacent to drilling, while somewhat unexpected, provides insights into the proximity impacts of shale gas. If residence proximity to shale gas embodied a purely negative biological effect from air pollution, then one would expect the resulting impact on birth outcomes to be less negative after removing infants that are exposed through both water and air. However, as proximity is also correlated with a potential to increase income through royalties, the direction of the change in estimates from leaving out SGD-adjacent mothers is ambiguous. For our sample of infants, the results are suggestive of an income effect that corresponds with residence proximity.

We next explore heterogeneity in impacts by re-estimating our CWS fixed effects models using sub-groups of infants based on the mother's socioeconomic status (SES) in Table 10. As impacts were previously estimated to be within 1 km, we present estimates for gas well threats within 0.5 and 1 km only. Focusing on the 1 km threats first, we find that the magnitude of prematurity impacts are larger for mothers who participate in WIC (0.466 pp) and for those that paid for hospital services with Medicaid (0.576 pp) compared to the 0.418 pp increase in prematurity estimated for the overall population that live at least 1 km away from the nearest well bore. As WIC and Medicaid enrollment are potential proxies for income, this is suggestive of these groups taking less avoidance behavior and thus suffer larger impacts to reproductive health (Neidell, 2009; Currie et al., 2013). Estimates for the group with an education of high school or less (conditional on being older than 18) is smaller than that for the average population (0.148 pp, not statistically significant). What is somewhat surprising is that we find the subgroup of college-educated women (who are at least 22 years of age) sees impacts on prematurity of 0.797 pp, comparable to the size of impacts for the WIC subgroup. We find similar results for impacts on low birth weight. If educational attainment is also an indicator of income, then this would suggest avoidance behavior is less likely to be an issue. That mothers are not so responsive to drilling in our setting is potentially reasonable if SGD near CWS sources are less observable than that near the residence: the two types of exposure are highly, but not perfectly, correlated (correlation of 0.6), and mothers may assume that piped public water protects them from contamination from the industry, as indicated by perceived risks being primarily associated with private ground water wells (Muehlenbachs et al., 2015). In robustness checks that assesses sorting, we find additional evidence suggesting that mothers are unaware of SGD threats at their water source location.

We explore the overall impacts on prematurity and low birth weight using within-sibling comparisons. Table 11 gives impacts on both birth outcomes for estimates with and without restrictions on residence proximity to well bores. Focusing on the estimates in panel B, where we remove mothers who live within 1 km of well bores, impacts on prematurity and low birth weight for well bore threats within 0.5 km are respectively 0.667 (not statistically significant) and 0.786 (5% significance level). For well bore threats at 1 km, we find a similar impact for prematurity (0.342 pp), but a counter-intuitive negative impact for low birth weight (-0.204 pp). We note that none of the estimates for threats at 1 km are statistically significant. We keep in mind that while inclusion of mother fixed helps control for unobserved, time-invariant differences in family background, it reduces our sample size by over a half and is based on a selected sample of mothers who have multiple births. Still, the results using mother fixed effects are qualitatively similar to those estimated from our CWS fixed effects models. In our view, these results support the general finding that both birth weight and gestation are negatively affected through SGD activity's impact on drinking water quality.

Before concluding, we provide some robustness checks through alternative specifications of our threat variable of interest and inclusion of additional controls. We also assess the extent to which compositional changes due to maternal sorting in response to gas well development could threaten our identification of impacts. First, we quantify the impact on birth outcomes if *any* well bore is drilled with 'd' km of a water intake (Table 12). The impact of there being any well bores drilled within 1 km of ground water intakes leads to an increase in premature births of about 4.0 pp and an increase in incidence of low birth weight of 2.8 pp.

To directly address the potential that air quality changes are driving our results, we re-estimate our model by removing areas near coal mines and including additional proxies for ambient air quality (not captured by our TRI controls). Table 13 estimates the impacts on birth outcomes that removes areas that are susceptible to coal mining (either historical or currently active). The magnitude of impacts remain stable for most specifications, regardless of whether we limit coal seams to be a minimum of 1, 5, or 10 km away from the source intake. Impacts are mostly statistically significant except when removing infants exposed to coal seams within 10 km, which reduces our sample to about 30 percent of its original size. Table 14 introduces (from left to right) controls for distance to the nearest state road (in meters) and ambient air quality controls at the block-group-by-year level using the EPA RSEI model. Estimated impacts are virtually the same. Additionally, the estimates are stable after inclusion of controls for residence distance to the nearest gas well (shown in Table A5 of the Appendix).[30] Taken together, these findings lead us to believe that unconventional drilling has an independent impact on birth outcomes through contamination of public drinking water.[31]

Lastly, we test whether our estimated impacts merely reflect compositional changes in the types of mothers who select into fertility near gas well development. First, Table 15 regresses an indicator for some maternal characteristic (e.g. mothers who participate in WIC) on the number of well bore threats, both near and far, where the same set of controls as those in Table 9 are used. Panel A of this table defines gas well threats using CWS intake proximity, with residence proximity

[30] We re-estimate the overall drilling impacts from the specification in Table 9, which includes moms in close proximity of well bores, but include distance bins for the minimum distance to the nearest gas well at each kilometer.

[31] Our estimates are additionally robust to inclusion of paternal age and education (Table A4 of the Appendix). To maintain a consistent a sample, we include dummy variables where the paternal characteristics are missing. As there may be non-singleton births, we assess our results limiting to singleton births only and find similar results (Table A6 of the Appendix).

limited to being at least greater than 1 km, and Panel B defines threats using the distance to maternal address. Regardless of how threats are defined, the estimates find generally economically insignificant impacts on maternal characteristics, reflecting a change of less than 2% from the mean for a standard deviation increase in the number of gas wells drilled for all characteristics. Table 16 then estimates the impact of additional gas wells drilled either within the vicinity of CWS intakes (Panel A) or near the maternal address (Panel B) on whether mothers switch water systems or ZIP codes using the sample of mothers who have multiple births. As with before, the sample in Panel A is limited to infants born to mothers whose residence is at least 1 km away from the nearest well bore and the sample in Panel B makes no restrictions on gas well distance to residence. The same set of controls used in the specifications of Table 9 are used, and CWS fixed effects are included for all specifications. We find no statistically significant impacts of gas well activity near source water intakes on likelihood to move. Interestingly, we do find evidence that well bores drilled in close proximity (0.5 km) of the maternal residence increases the chance of moving. That mothers may sort in response to drilling activity nearby but not in response to drilling at water intake locations bolsters our identification strategy using source intake locations to define exposure. We infer from these tests that our estimated impacts of SGD on birth outcomes are not driven by changes in the type of mothers who selects into fertility near gas well development.

5 Conclusion

This study seeks to understand and quantify the impacts of shale gas development on drinking water quality and infant health for families living in community water systems with intakes near SGD. We assembled a unique data set with the latitude and longitude of new mothers' residences, community water system's water source locations and the locations of shale gas wells in Pennsylvania. We find robust and consistent evidence of an effect of shale gas development on water quality: an additional well bore drilled within 1 km of a CWS intake increases contamination by 1.3 percent. These estimated impacts are primarily driven by well bores up-hill of the water intake and are associated with wells in production. Our findings are robust to placebo tests as well as sample restrictions that remove areas with any historical or active coal mining. Additionally, we find consistent water quality impacts as measured by USGS water quality monitors in the vicinity of gas wells, which suggests that our estimated water impacts from public water systems data represent a lower bound as water systems only sample a subset of known SGD chemicals.

We find evidence that water quality changes due to SGD indeed produces measurable impacts on birth outcomes: the incidences of prematurity and low birth weight respectively increase by 4.4 and 4.7 percent from the mean in response to a standard deviation increase in gas wells drilled during gestation. The overall impacts of increased drilling at a closer buffer of 0.5 km of water sources are upwards of 11.6 (prematurity) and 7.2 (low birth weight) percent. We determine that our results are unlikely to be driven by correlated air quality changes associated with congestion/traffic or coal, nor are they driven by maternal mobility and fertility decisions in response to SGD.

The shale gas revolution has undoubtedly yielded benefits. An estimate by Hausman and Kellogg (2015) puts the annual welfare increase between 2007 to 2013 at $48 billion. It is important, however, to keep in mind that many of the unintended environmental consequences of these technological advancements are not accounted for in this welfare figure, due, in part, to the difficulty in measuring these external costs. Our paper takes a first step to better understand one of these consequences by estimating the impacts on infant health through contamination of drinking water. Combined with the economics literature regarding the short and long term impacts of neonatal health, our findings suggest that the social costs of our water and resulting birth impacts are non-trivial. Researchers have shown that neonatal health has a significant effect on both mortality within one year and mortality up to age 17 (Oreopoulos et al., 2008). Further, these outcomes are strong predictors of a host of longer term outcomes, such as human capital accumulation, welfare take-up, earnings, and labor force participation (Black et al., 2007; Oreopoulos et al., 2008; Johnson and Schoeni, 2011; Figlio et al., 2014). These findings provide an impetus for the regulation of drinking water policy and/or the shale gas industry; policies that mitigate SGD's water pollution impacts could enhance efficiency as long as the benefits of mitigation outweigh their regulatory costs.

Tables

Table 1: Birth Outcome & Maternal Characteristics
of Population and Estimation Sample

	Population		Intake ≤ 10 km			Residence ≤ 10 km		
Variables	Mean	St. Dev.	Mean	St. Dev.	t-stat	Mean	St. Dev.	t-stat
Premature	0.10	0.29	0.09	0.29	-0.98	0.09	0.29	-1.42
Low Birth Weight	0.08	0.27	0.08	0.27	-0.96	0.07	0.26	-1.44
Age	28.36	5.94	27.26	5.76	-2.69	27.84	5.81	-1.25
Teen	0.04	0.20	0.05	0.22	0.79	0.04	0.20	-0.14
Born in PA	0.69	0.46	0.78	0.41	5.24	0.79	0.41	9.37
Hispanic	0.02	0.15	0.01	0.10	-3.02	0.01	0.08	-4.77
Black	0.15	0.36	0.04	0.21	-2.41	0.05	0.22	-2.21
HS or Less	0.38	0.49	0.43	0.50	2.14	0.39	0.49	0.51
College Educated	0.44	0.50	0.38	0.48	-1.76	0.43	0.49	-0.27
Married	0.63	0.48	0.58	0.49	-0.52	0.63	0.48	0.15
Smoker	0.22	0.41	0.34	0.47	6.45	0.29	0.45	4.54
WIC	0.34	0.48	0.47	0.50	2.68	0.40	0.49	1.03
Paid with Medicaid	0.28	0.45	0.39	0.49	2.01	0.34	0.47	1.32
Observations	1,502,874		115,690			340,827		

Note. Table compares average births and mother characteristics for the population in Pennsylvania to that for the sample of births whose CWS intake location is within 10 km of at least one gas well, as well as the average characteristics for the sample of births to mothers who live within 10 km of a gas well. T-statistics that test the difference in characteristics between each sample and the rest of the population are provided.

Table 2: Birth Outcome & Maternal Characteristics by Buffers within 10 km

Intake Distance	d ≤ 10 km		d ≤ 0.5 km		d ≤ 1 km		d ≤ 1.5 km		d ≤ 2 km	
Variables	Mean	St. Dev.	Mean	t-stat	Mean	t-stat	Mean	t-stat	Mean	t-stat
Premature	0.09	0.29	0.09	-0.47	0.09	-0.74	0.09	-1.06	0.08	-2.00
Low Birth Weight	0.08	0.27	0.07	-0.75	0.07	-0.64	0.07	-1.00	0.07	-2.71
Age	27.26	5.76	26.61	-1.79	26.56	-2.27	26.47	-2.82	27.34	0.14
Teen	0.05	0.22	0.05	-0.59	0.05	1.40	0.05	1.33	0.05	-0.65
Born in PA	0.78	0.41	0.78	0.04	0.81	1.49	0.80	1.15	0.77	-0.29
Hispanic	0.01	0.10	0.01	-1.04	0.00	-1.58	0.00	-1.67	0.00	-1.69
Black	0.04	0.21	0.01	-3.82	0.01	-3.99	0.01	-4.01	0.01	-3.91
HS or Less	0.43	0.50	0.46	0.79	0.48	1.51	0.48	1.87	0.41	-0.39
College Educated	0.38	0.48	0.36	-0.55	0.33	-1.51	0.32	-2.02	0.41	0.50
Married	0.58	0.49	0.59	0.22	0.59	0.07	0.58	-0.09	0.63	1.19
Smoker	0.34	0.47	0.34	0.25	0.36	1.40	0.36	1.51	0.31	-0.72
WIC	0.47	0.50	0.51	0.89	0.51	0.99	0.52	1.39	0.44	-0.58
Paid with Medicaid	0.39	0.49	0.43	1.25	0.43	1.42	0.43	1.55	0.36	-0.61
Observations	115,690		1,374		7,725		10,897		23,131	

Note. Table compares average births and mother characteristics of infants with CWS intakes at various distances to gas wells within 10 km. T-statistics that test the difference in characteristics between each sample within 10 km and the 10 km sample are provided.

Table 3: Water Quality Summary Statistics

	SGD-Related		Not SGD-related	
Variables	Mean	St. Dev.	Mean	St. Dev.
# of Well Bore Spud Threats				
0.5 km	0.029	0.207	0.028	0.194
1 km	0.175	0.905	0.199	1.027
1.5 km	0.451	1.558	0.470	1.625
2 km	0.853	2.528	0.849	2.490
10 km	27.455	50.969	27.284	51.102
Sampling Result (ppm)	0.035	0.170	0.028	0.182
MCL Exceedance	0.004	0.059	0.006	0.079
MCL Goal Exceedance	0.078	0.268	0.011	0.104
Inorganic Compounds	0.353	0.478	0.003	0.058
Volatile Organic Compounds				
Halogenated	0.207	0.405	0.199	0.399
non-Halogenated	0.023	0.149	0.000	0.000
Halogenated, semi-	0.129	0.336	0.160	0.367
non-Halogenated, semi-	0.038	0.192	0.000	0.000
Synthetic Organic Compounds	0.192	0.394	0.638	0.481
Disinfectant Byproducts	0.161	0.368	0.092	0.290
Fuels	0.153	0.360	0.098	0.297
Radionuclides	0.011	0.105	0.007	0.084
Coliform	0.000	0.000	0.033	0.178
Nitrates / Nitrites	0.077	0.267	0.000	0.000
Observations	69,239		102,376	

Table 4: Water Quality Impacts of SGD Chemicals

A. Overall Impacts

Variables	d ≤ 0.5 km	d ≤ 1 km	d ≤ 1.5 km	d ≤ 2 km
Bores in ‘d’ km	0.0196**	0.0125**	0.000226	0.000538
	(0.00921)	(0.00487)	(0.00161)	(0.000750)
Bores in 10 km	-5.13e-05	-0.000111*	-3.14e-05	-5.42e-05
	(5.50e-05)	(6.69e-05)	(6.22e-05)	(6.39e-05)
Observations	69,237	69,237	69,237	69,237

B. By Elevation

Variables	d ≤ 0.5 km	d ≤ 1 km	d ≤ 1.5 km	d ≤ 2 km
Uphill Bores in ‘d’	0.0216**	0.0133**	0.000267	0.000306
	(0.0103)	(0.00541)	(0.00172)	(0.000874)
Downhill Bores in ‘d’	-0.00184	0.00644	-0.000291	0.00147
	(0.00573)	(0.00399)	(0.00360)	(0.00116)
Observations	69,237	69,237	69,237	69,237

C. By Well Production

Variables	d ≤ 0.5 km	d ≤ 1 km	d ≤ 1.5 km	d ≤ 2 km
Producing Bores in ‘d’	0.0216**	0.0133*	-0.000337	0.000541
	(0.0103)	(0.00749)	(0.00238)	(0.00107)
Never Producing Bores in ‘d’	-0.00184	0.0110**	0.00132	0.000531
	(0.00573)	(0.00518)	(0.00175)	(0.000823)
Observations	69,237	69,237	69,237	69,237

Note. Each column is a separate regression. Sample includes SGD-related chemicals only. Robust standard errors in parentheses are clustered at the CWS level. The dependent variable is log(water sampling result). *** p<0.01, ** p<0.05, * p<0.1.

Table 5: Water Quality Impacts of SGD - Placebo Checks

A. Non-SGD Chemicals

Non-SGD Chemicals	d ≤ 0.5 km	d ≤ 1 km	d ≤ 1.5 km	d ≤ 2 km
Bores in ‘d’ km	0.00117	-0.00336**	-0.000846	3.46e-05
	(0.00398)	(0.00167)	(0.000982)	(0.000449)
Bores in 10 km	9.66e-05**	0.000122**	0.000115**	9.62e-05*
	(4.80e-05)	(5.11e-05)	(5.40e-05)	(5.21e-05)
Observations	102,370	102,370	102,370	102,370

B. Future Threats

SGD Chemicals	d ≤ 0.5 km	d ≤ 1 km	d ≤ 1.5 km	d ≤ 2 km
Bores next 180	0.00571	0.00246	0.00183	0.000691
days in ‘d’ km	(0.0114)	(0.00339)	(0.00154)	(0.000724)
Observations	69,237	69,237	69,237	69,237

C. Permitted but Never Spudded Wells

SGD Chemicals	d ≤ 0.5 km	d ≤ 1 km	d ≤ 1.5 km	d ≤ 2 km
Permitted in ‘d’ km	0.0170	-0.000572	-0.00254*	-0.00108
	(0.0184)	(0.00262)	(0.00144)	(0.00103)
Permitted in 10 km	-2.04e-05	1.17e-05	6.10e-05	6.09e-05
	(8.76e-05)	(8.65e-05)	(7.25e-05)	(9.29e-05)
Observations	75,290	75,290	75,290	75,290

Note. Each column within each panel is a separate regression. The dependent variable is log(water sampling result). Robust standard errors in parentheses are clustered at the CWS level. *** p<0.01, ** p<0.05, * p<0.1.

Table 6: Water Impacts Limiting to Areas without Historical or Active Coal Seams

Sample		No Historical Coal in 'd' km			No Coal within 'd' km		
Limit:	Baseline	d≤1 km	d≤5 km	d≤10 km	d≤1 km	d≤5 km	d≤10 km
Bores in 0.5 km	0.0196**	0.0181**	0.00894***	0.00909***	0.00669*	0.0110***	0.0110***
	(0.00921)	(0.00753)	(0.00330)	(0.00303)	(0.00362)	(0.00374)	(0.00353)
Observations	69,237	64,487	52,096	41,605	60,481	45,609	36,428
Sample		No Historical Coal in 'd' km			No Coal within 'd' km		
Limit:	Baseline	d≤1 km	d≤5 km	d≤10 km	d≤1 km	d≤5 km	d≤10 km
Bores in 1 km	0.0125**	0.0119***	0.00699**	0.00790***	0.00657**	0.00778***	0.00836***
	(0.00487)	(0.00435)	(0.00274)	(0.00286)	(0.00272)	(0.00280)	(0.00282)
Observations	69,237	64,487	52,096	41,605	60,481	45,609	36,428

Note.

Table 7: Alternative Characterizations of Water Quality Impacts

A. Alternative Outcomes

Dep. Var.:	Detected	Exceed MCLg	Exceed MCL	Result (ppm)	Std. Result
Bores in 1 km	0.0311***	0.00994**	0.000937	0.0203**	0.106**
	(0.0112)	(0.00492)	(0.000790)	(0.00801)	(0.0428)
Bores in 10 km	0.000213	0.000175	-1.25e-05	-0.000175*	-0.000308
	(0.000220)	(0.000126)	(3.39e-05)	(0.000104)	(0.000586)
Mean	0.242	0.0778	0.00352	0.0349	-6.54e-05
Observations	69,237	67,538	67,538	69,237	65,568

B. Detection by Chemical Subgroups

Chemical Subgroup:	IOC	SOC	VOC	DBP	Nitrate/Nitrite
Bores in 1 km	0.0655**	0.00617*	0.00156	-0.0373	-0.00944**
	(0.0305)	(0.00349)	(0.000982)	(0.0459)	(0.00384)
Bores in 10 km	-0.000142	0.000159*	-1.45e-05	0.00138**	5.20e-05
	(0.000456)	(8.58e-05)	(2.86e-05)	(0.000626)	(7.94e-05)
Mean	0.324	0.00581	0.00790	0.673	0.00587
Observations	23,867	10,669	6,835	11,165	10,557

Note. Panel A examines alternative water quality outcomes as the dependent variable. 'Detected' indicates a non-zero water sampling results, 'MCL' represent Maximum Contaminant Level, and 'MCLg' represents MCL goals. Panel B regresses an indicator for a positive detection on shale gas threats. Robust standard errors in parentheses are clustered at the CWS level. The dependent variable is log(water sampling result). *** p<0.01, ** p<0.05, * p<0.1.

Table 8: USGS Ground Water Quality Impacts of SGD

A. SGD Chemicals Only

Variables	d ≤ 0.5 km	d ≤ 1 km	d ≤ 1.5 km	d ≤ 2 km
Bores in 'd' km	0.0331*	0.0313***	0.00992	0.00198
	(0.0175)	(0.00898)	(0.00641)	(0.00396)
Bores in 10 km	0.000594*	0.000437	0.000446	0.000588
	(0.000354)	(0.000358)	(0.000379)	(0.000382)
Observations	4,256	4,256	4,256	4,256

B. Non-SGD Chemicals Only

Variables	d ≤ 0.5 km	d ≤ 1 km	d ≤ 1.5 km	d ≤ 2 km
Bores in 'd' km	-0.0258	0.0103	0.00123	0.00304
	(0.0413)	(0.0206)	(0.0162)	(0.0103)
Bores in 10 km	0.000220	4.42e-05	0.000113	1.94e-05
	(0.000988)	(0.00100)	(0.00105)	(0.00107)
Observations	623	623	623	623

Note. Each column is a separate regression. The dependent variable is log(water sampling result) as measured by USGS ground water monitors. Specifications include temperature and precipitation controls, contaminant group, quarter, and county-by-year fixed effects. *** p<0.01, ** p<0.05, * p<0.1.

Table 9: Impacts on Birth Outcomes

A. *No Restrictions on Residence Proximity to Well Bores*

Dep. Var.:	Prematurity			Low Birth Weight		
Exposure Buffer:	d≤0.5 km	d≤1 km	d≤1.5 km	d≤0.5 km	d≤1 km	d≤1.5 km
SD Bores in 'd' km	0.00974***	0.00265**	0.000226	0.00295	0.00199**	0.000671
	(0.00348)	(0.00107)	(0.000316)	(0.00323)	(0.000828)	(0.000549)
SD Bores in 10 km	9.59e-05**	8.84e-05**	9.22e-05**	5.57e-05	5.07e-05	4.75e-05
	(4.21e-05)	(4.17e-05)	(4.18e-05)	(3.61e-05)	(3.65e-05)	(3.66e-05)
Dep. Var. Mean	0.0931	0.0931	0.0931	0.0769	0.0769	0.0769
Observations	115,041	115,041	115,041	115,041	115,041	115,041

B. *Residence Proximity to Well Bores >1 km*

Dep. Var.:	Prematurity			Low Birth Weight		
Exposure Buffer:	d≤0.5 km	d≤1 km	d≤1.5 km	d≤0.5 km	d≤1 km	d≤1.5 km
SD Bores in 'd' km	0.0108*	0.00418***	5.93e-05	0.00555	0.00365***	0.000976*
	(0.00553)	(0.00151)	(0.000328)	(0.00494)	(0.00129)	(0.000584)
SD Bores in 10 km	7.06e-05	6.35e-05	6.72e-05	6.19e-05	5.63e-05	5.00e-05
	(4.68e-05)	(4.53e-05)	(4.62e-05)	(4.06e-05)	(4.09e-05)	(4.15e-05)
Dep. Var. Mean	0.0933	0.0933	0.0933	0.0770	0.0770	0.0770
Observations	112,525	112,525	112,525	112,525	112,525	112,525

Note. Table presents estimates of the impacts of well bores drilled within 'd' km to a CWS intake on the birth outcomes of prematurity and low birth weight. The sample in panel A includes all infants whose CWS intake is within 10 km of a well bore, whereas the sample in panel B restricts to infants born t mothers who live at least 1 km away from the nearest well bore. The variable of interest, number of well bores drilled, are in units of standard deviations. Each column represents a separate regression. All regressions control for maternal characteristics and pregnancy risks, variation in temperature and precipitation, the number of coliform and disinfectant by-product MCL exceedances, Toxic Release Inventory (TRI) facilities and their on-site releases, and year-by-month fixed effects. 'FE' represents included fixed effects that change across each regression, where standard errors are clustered at the public water system (or CWS) level.

Table 10: Impacts by Mother's Socioeconomic Status (SES)

Dep. Var.:	Threats 'd' $\leq$ 0.5 km					Threats 'd' $\leq$ 1 km				
Premature	WIC	Medicaid	HS or Less	College	Born in PA	WIC	Medicaid	HS or Less	College	Born in PA
SD Bores in 'd' km	0.0192**	0.0207***	0.00947*	0.0179*	0.00778	0.00466***	0.00576***	0.00148	0.00797***	0.00365*
	(0.00856)	(0.00721)	(0.00500)	(0.00959)	(0.00690)	(0.00163)	(0.00212)	(0.00139)	(0.00288)	(0.00192)
Mean	0.0971	0.103	0.101	0.0855	0.0938	0.0971	0.103	0.101	0.0855	0.0938
Observations	53,470	43,425	43,455	41,617	87,625	53,470	43,425	43,455	41,617	87,625

Dep. Var.:	Threats 'd' $\leq$ 0.5 km					Threats 'd' $\leq$ 1 km				
Low Birth Weight	WIC	Medicaid	HS or Less	College	Born in PA	WIC	Medicaid	HS or Less	College	Born in PA
SD Bores in 'd' km	0.0126	0.0132*	-0.000745	0.0141	0.00365	0.00384**	0.00353	0.00248	0.00518**	0.00262**
	(0.00781)	(0.00674)	(0.00377)	(0.0112)	(0.00572)	(0.00187)	(0.00236)	(0.00207)	(0.00253)	(0.00124)
Mean	0.0878	0.0968	0.0929	0.0617	0.0773	0.0878	0.0968	0.0929	0.0617	0.0773
Observations	53,470	43,425	43,455	41,617	87,625	53,470	43,425	43,455	41,617	87,625

Note. Table re-estimates the impacts of well bores drilled within 'd' km to a CWS intake on birth outcomes for subgroups of mothers based on maternal characteristics. 'WIC' and 'Medicaid' represent whether the mother participated in the WIC program or paid for hospital services with Medicaid. 'HS or Less' and 'College' represent the highest level of education achieved and are conditional upon being older than 18 or 22, respectively. Sample restricts the minimum distance between a mother's residence and well bores nearby to be at least 1 km. The set of controls used in the specifications of Table 9 are used.

Table 11: Impacts on Birth Outcome with Mother Fixed Effects

A. No Restrictions on Well Bore Proximity

Dep. Var.:	Prematurity			Low Birth Weight		
	d≤0.5 km	d≤1 km	d≤1.5 km	d≤0.5 km	d≤1 km	d≤1.5 km
SD Bores in 'd' km	0.00751*	0.00245	-0.00161	0.00368	-0.00316	0.000623
	(0.00441)	(0.00389)	(0.00132)	(0.00612)	(0.00305)	(0.00260)
SD Bores in 10 km	0.000121	0.000121	0.000144	7.70e-05	8.81e-05	7.68e-05
	(0.000146)	(0.000143)	(0.000145)	(8.84e-05)	(8.79e-05)	(8.49e-05)
Dep. Var. Mean	0.119	0.119	0.119	0.0978	0.0978	0.0978
Observations	45,501	45,501	45,501	45,501	45,501	45,501

B. Residence Proximity to Well Bore > 1 km

Dep. Var.:	Prematurity			Low Birth Weight		
	d≤0.5 km	d≤1 km	d≤1.5 km	d≤0.5 km	d≤1 km	d≤1.5 km
SD Bores in 'd' km	0.00667	0.00342	-0.00166	0.00786**	-0.00204	0.00124
	(0.00708)	(0.00517)	(0.00157)	(0.00305)	(0.00407)	(0.00296)
SD Bores in 10 km	9.60e-05	9.81e-05	0.000122	7.65e-05	9.32e-05	7.37e-05
	(0.000158)	(0.000155)	(0.000158)	(9.66e-05)	(9.61e-05)	(9.18e-05)
Dep. Var. Mean	0.119	0.119	0.119	0.0982	0.0982	0.0982
Observations	44,380	44,380	44,380	44,380	44,380	44,380

Table 12: Impact on Birth Outcome in the Presence of Any Well Bores

Dep. Var.:	FE: CWS		FE: Mom	
Premature	d≤0.5 km	d≤1 km	d≤0.5 km	d≤1 km
Any Bores in 'd' km	0.0677	0.0398***	0.0598	0.0409
	(0.0461)	(0.0149)	(0.0796)	(0.0606)
Any Bores in 10' km	-0.0193	-0.0227***	0.0116	0.00213
	(0.0202)	(0.00831)	(0.0401)	(0.0241)
Observations	112,525	112,525	44,380	44,380

Dep. Var.:	FE: CWS		FE: Mom	
Low Birth Weight	d≤0.5 km	d≤1 km	d≤0.5 km	d≤1 km
Any Bores in 'd' km	0.0195	0.0280*	0.0617*	-0.0350
	(0.0396)	(0.0153)	(0.0338)	(0.0559)
Any Bores in 10' km	-0.0181	-0.0176**	-0.0237	0.0954***
	(0.0182)	(0.00822)	(0.0276)	(0.0272)
Observations	112,525	112,525	44,380	44,380

Note. Sample is limited to infants born to mothers whose residence is at least 1 km away from the nearest well bore. The regressor of interest, 'Any Wells in 'd' km' is an indicator variable equal to 1 if any well bores were drilled during the gestation of the infant. The set of controls used in the specifications of Table 9 are used.

Table 13: Birth Impacts Limiting to Areas without Historical or Active Coal Seams

Sample Limit:		No Historical Coal in 'd' km			No Coal within 'd' km		
Dep. Var.: *Premature*	Baseline	d≤1 km	d≤5 km	d≤10 km	d≤1 km	d≤5 km	d≤10 km
SD Bores in 1 km	0.00418***	0.00446**	0.00387*	0.00300	0.00530**	0.00390*	0.00307
	(0.00151)	(0.00190)	(0.00213)	(0.00214)	(0.00219)	(0.00207)	(0.00205)
Mean	0.0933	0.0938	0.0886	0.0857	0.0940	0.0881	0.0848
Observations	112,525	102,051	62,540	37,611	98,287	56,377	32,611
Sample Limit:		No Historical Coal in 'd' km			No Coal within 'd' km		
Dep. Var.: *LBW*	Baseline	d≤1 km	d≤5 km	d≤10 km	d≤1 km	d≤5 km	d≤10 km
SD Bores in 1 km	0.00365***	0.00379***	0.00477***	0.00426***	0.00473***	0.00463***	0.00432***
	(0.00129)	(0.00142)	(0.00164)	(0.00164)	(0.00159)	(0.00164)	(0.00161)
Mean	0.0770	0.0775	0.0720	0.0702	0.0775	0.0716	0.0691
Observations	112,525	102,051	62,540	37,611	98,287	56,377	32,611

Note. Table re-estimates the impacts of well bores drilled within 1 km to a CWS intake on birth outcomes from Table 9 on a sample that removes infants whose CWS intakes are within 'd' km of historically active or any coal seams. For comparison, Table 9 estimates for an exposure buffer of 1 km is given in the first column. All specifications include CWS fixed effects.

Table 14: Birth Impacts Conditional on Additional Air Quality Measures

Control:	Distance to Nearest Road			RSEI Ambient Air Quality		
Dep. Var.: *Premature*	d≤0.5 km	d≤1 km	d≤1.5 km	d≤0.5 km	d≤1 km	d≤1.5 km
SD Bores in 'd' km	0.0107*	0.00418***	5.87e-05	0.0110*	0.00372**	4.27e-05
	(0.00553)	(0.00151)	(0.000328)	(0.00581)	(0.00154)	(0.000345)
Dist. to Road (m)	-6.39e-06	-6.44e-06	-6.43e-06			
	(5.24e-06)	(5.24e-06)	(5.24e-06)			
RSEI Concentration				-9.35e-09	-9.35e-09	-9.35e-09
				(6.26e-09)	(6.27e-09)	(6.26e-09)
Mean	0.0933	0.0933	0.0933	0.0955	0.0955	0.0955
Observations	112,525	112,525	112,525	84,357	84,357	84,357

Control:	Distance to Nearest Road			RSEI Ambient Air Quality		
Dep. Var.: *LBW*	d≤0.5 km	d≤1 km	d≤1.5 km	d≤0.5 km	d≤1 km	d≤1.5 km
SD Bores in 'd' km	0.00554	0.00365***	0.000976*	0.00546	0.00319**	0.000880
	(0.00494)	(0.00129)	(0.000584)	(0.00535)	(0.00135)	(0.000663)
Dist. to Road (m)	-2.67e-06	-2.68e-06	-2.69e-06			
	(4.18e-06)	(4.18e-06)	(4.18e-06)			
RSEI Concentration				-2.00e-08***	-2.00e-08***	-2.00e-08***
				(7.37e-09)	(7.39e-09)	(7.39e-09)
Mean	0.0770	0.0770	0.0770	0.0792	0.0792	0.0792
Observations	112,525	112,525	112,525	84,357	84,357	84,357

Note. Table re-estimates the impacts of well bores drilled within 'd' km to a CWS intake on birth outcomes from Table 9 that includes additional controls for air quality proxied by 1) the distance to the nearest state road, and 2) the average RSEI concentration in the year of birth. All specifications include CWS fixed effects.

Table 15: Impact on Maternal Characteristics

A. Well Bore Threats Near CWS Intake (Residence >1 km)

	‘d’≤0.5 km					‘d’≤1 km				
Dep. Var.:	WIC	Medicaid	College	Smoker	Born in PA	WIC	Medicaid	College	Smoker	Born in PA
SD Bores in ‘d’ km	-0.00966**	0.00731*	0.00240	0.00540	0.0132**	-0.00458**	0.00166	-0.00353**	0.00503***	-0.000299
	(0.00437)	(0.00411)	(0.00560)	(0.00764)	(0.00579)	(0.00192)	(0.00120)	(0.00141)	(0.00185)	(0.00219)
SD Bores in 10 km	-1.27e-05	3.60e-05	-6.92e-06	-1.14e-05	-1.11e-05	-7.73e-06	3.09e-05	-1.77e-05	-1.27e-05	-1.10e-05
	(5.33e-05)	(5.09e-05)	(5.35e-05)	(5.99e-05)	(5.58e-05)	(5.33e-05)	(5.07e-05)	(5.35e-05)	(6.24e-05)	(5.61e-05)
Dep. Var. Mean	0.476	0.386	0.379	0.335	0.779	0.476	0.386	0.379	0.335	0.779
Observations	112,525	112,525	112,525	112,525	112,525	112,525	112,525	112,525	112,525	112,525

B. Well Bore Threats Near Maternal Residence

	‘d’≤0.5 km					‘d’≤1 km				
Dep. Var.:	WIC	Medicaid	College	Smoker	Born in PA	WIC	Medicaid	College	Smoker	Born in PA
SD Bores in ‘d’ km	-0.00148	0.000170	0.00271	-0.00543	-0.00545	-0.00273	4.07e-06	-0.000284	0.00223	-0.00129
	(0.00571)	(0.00279)	(0.00538)	(0.00354)	(0.00353)	(0.00166)	(0.00160)	(0.00181)	(0.00179)	(0.00168)
SD Bores in 10 km	7.93e-06	2.36e-05	-4.26e-05	-7.48e-07	-1.38e-05	1.08e-05	2.68e-05	-3.92e-05	-2.48e-06	-1.50e-05
	(5.41e-05)	(4.23e-05)	(6.78e-05)	(5.76e-05)	(7.09e-05)	(5.40e-05)	(4.22e-05)	(6.68e-05)	(5.76e-05)	(7.10e-05)
Dep. Var. Mean	0.475	0.385	0.380	0.335	0.780	0.475	0.385	0.380	0.335	0.780
Observations	115,041	115,041	115,041	115,041	115,041	115,041	115,041	115,041	115,041	115,041

Note. Table presents regressions of an indicator for a maternal characteristic on the number of gas wells drilled within the vicinity of CWS intake (Panel A) and the maternal address (Panel B). Additionally, the sample in Panel A is limited to infants born to mothers whose residence is at least 1 km away from the nearest well bore; the sample in Panel B makes no restrictions on gas well distance to residence. The set of controls used in the specifications of Table 9 are used, where the dependent variable is omitted. For all specifications, ZIP code fixed effects are included.

Table 16: Impact on the Probability Mother Moves

	A. Well Bore Threats Near CWS Intake (Residence >1 km)					
Dep. Var.:	Switch CWS			Switch ZIP code		
	d≤0.5 km	d≤1 km	d≤1.5 km	d≤0.5 km	d≤1 km	d≤1.5 km
SD Bores in 'd' km	-0.00569	-0.00104	-3.42e-05	-0.00608	-0.00225	-0.00200
	(0.00606)	(0.00299)	(0.00112)	(0.00532)	(0.00373)	(0.00149)
SD Bores in 10 km	-2.49e-05	-1.32e-05	-2.28e-05	-3.64e-05	-1.90e-05	-1.07e-05
	(0.000100)	(9.88e-05)	(9.94e-05)	(0.000100)	(9.84e-05)	(9.74e-05)
Dep. Var. Mean	0.183	0.183	0.183	0.284	0.284	0.284
Observations	46,268	46,268	46,268	46,268	46,268	46,268

	B. Well Bore Threats Near Maternal Residence					
Dep. Var.:	Switch CWS			Switch ZIP code		
	d≤0.5 km	d≤1 km	d≤1.5 km	d≤0.5 km	d≤1 km	d≤1.5 km
SD Bores in 'd' km	0.0171***	0.00443	-0.000599	0.0145***	0.00676	-0.00230
	(0.00196)	(0.00408)	(0.00185)	(0.00311)	(0.00438)	(0.00183)
SD Bores in 10 km	-8.40e-05	-8.82e-05	-6.97e-05	-8.01e-05	-8.21e-05	-6.09e-05
	(0.000128)	(0.000128)	(0.000121)	(0.000144)	(0.000145)	(0.000140)
Dep. Var. Mean	0.183	0.183	0.183	0.284	0.284	0.284
Observations	46,268	46,268	46,268	46,268	46,268	46,268

Note. Table regresses an indicator for whether a mother switched water systems or ZIP codes on the number of gas wells drilled within the vicinity of CWS intake (Panel A) and the maternal address (Panel B) using the sample of mothers who have multiple births. Additionally, the sample in Panel A is limited to infants born to mothers whose residence is at least 1 km away from the nearest well bore; the sample in Panel B makes no restrictions on gas well distance to residence. The set of controls used in the specifications of Table 9 are used. For all specifications, CWS fixed effects are included.

References

Adsera, Alicia. 2004. "Changing fertility rates in developed countries. The impact of labor market institutions." *Journal of Population Economics*, 17(1): 17–43.

Adsera, Alicia. 2011. "Where are the babies? Labor market conditions and fertility in Europe." *European Journal of Population/Revue européenne de Démographie*, 27(1): 1–32.

ALL Consulting, LLC. 2009. "Modern Shale Gas Development in the United States: A Primer."

Allcott, Hunt, and Daniel Keniston. 2014. "Dutch disease or agglomeration? The local economic effects of natural resource booms in modern America."Technical report, National Bureau of Economic Research.

Banzhaf, Spencer H, and Randall P Walsh. 2008. "Do people vote with their feet? An empirical test of Tiebout's mechanism." *The American Economic Review*, 98(3): 843–863.

Bartik, Alexander Wickman, Janet Currie, Michael Greenstone, and Christopher R Knittel. 2016. "The Local Economic and Welfare Consequences of Hydraulic Fracturing."

Becker, GS. 1981. "A Treatise on the Family Harvard University Press Cambridge MA Google Scholar."

Black, Sandra E., Paul J. Devereux, and Kjell G. Salvanes. 2007. "From the cradle to the labor market? The effect of birth weight on adult outcomes." *The Quarterly Journal of Economics*, 122(1): 409–439.

Casey, Joan A, David A Savitz, Sara G Rasmussen, Elizabeth L Ogburn, Jonathan Pollak, Dione G Mercer, and Brian S Schwartz. 2015. "Unconventional Natural Gas Development and Birth Outcomes in Pennsylvania, USA.." *Epidemiology (Cambridge, Mass.)*.

Chay, K.Y., and M. Greenstone. 2003. "The Impact of Air Pollution on Infant Mortality: Evidence from Geographic Variation in Pollution Shocks Induced by a Recession*." *The Quarterly journal of economics*, 118(3): 1121–1167.

Currie, J., and R. Walker. 2011. "Traffic Congestion and Infant Health: Evidence from E-ZPass." *American Economic Journal: Applied Economics*, 3(1): 65–90.

Currie, Janet, Michael Greenstone, and Katherine Meckel. 2017. "Hydraulic Fracturing and Infant Health: New Evidence from Pennsylvania." *Science Advances*, 3(12): .

Currie, Janet, and Hannes Schwandt. 2014. "Short-and long-term effects of unemployment on fertility." *Proceedings of the National Academy of Sciences*, 111(41): 14734–14739.

Currie, Janet, Joshua S. Graff Zivin, Katherine Meckel, Matthew J. Neidell, and Wolfram Schlenker. 2013. "Something in the Water: Contaminated Drinking Water and Infant Health."

Deschênes, Olivier, Michael Greenstone, and Jonathan Guryan. 2009. "Climate change and birth weight." *The American Economic Review*, 99(2): 211–217.

DiGiulio, Dominic, Richard T Wilkin, Carlyle Miller, and Gregory Oberly. 2011. "Investigation of Ground Water Contamination near Pavillion." *Wyoming. Draft. US Environmental Protection Agency.*

Drollette, Brian D, Kathrin Hoelzer, Nathaniel R Warner, Thomas H Darrah, Osman Karatum, Megan P O'Connor, Robert K Nelson, Loretta A Fernandez, Christopher M Reddy, Avner Vengosh et al. 2015. "Elevated levels of diesel range organic compounds in groundwater near Marcellus gas operations are derived from surface activities." *Proceedings of the National Academy of Sciences*, p. 201511474.

EPA, US Environmental Protection Agency. 2012. "Study of the Potential Impacts of Hydraulic Fracturing on Drinking Water Resources: Progress Report." URL: http://www2.epa.gov/hfstudy/study-potential-impacts-hydraulic-fracturing-drinking-water-resources-progress-report-0, Accessed January, 2013.

Feyrer, James, Erin T Mansur, and Bruce Sacerdote., "Geographic Dispersion of Economic Shocks: Evidence from the Fracking Revolution." *American Economic Review*.

Figlio, David, Jonathan Guryan, Krzysztof Karbownik, and Jeffrey Roth. 2014. "The effects of poor neonatal health on children's cognitive development." *The American Economic Review*, 104(12): 3921–3955.

Finkel, Madelon L, and Jake Hays. 2015. "Environmental and health impacts of 'fracking': why epidemiological studies are necessary." *Journal of Epidemiology and Community Health*.

Graff Zivin, Joshua, Matthew Neidell, and Wolfram Schlenker. 2011. "Water Quality Violations and Avoidance Behavior: Evidence from Bottled Water Consumption." *American Economic Review: Papers and Proceedings*, 101(3): 448–453.

Grossman, Daniel S, David JG Slutsky et al. 2017. "The Effect of an Increase in Lead in the Water System on Fertility and Birth Outcomes: The Case of Flint, Michigan."Technical report.

Gustafsson, Siv, and Adriaan Kalwij. 2006. *Education and postponement of maternity: Economic analyses for industrialized countries.* 15: Springer Science & Business Media.

Hausman, Catherine, and Ryan Kellogg. 2015. "Welfare and Distributional Implications of Shale Gas." *Brookings Papers on Economic Activity*.

Hildenbrand, Zacariah L, Doug D Carlton Jr, Brian E Fontenot, Jesse M Meik, Jayme L Walton, Josh T Taylor, Jonathan B Thacker, Stephanie Korlie, C Phillip Shelor, Drew Henderson et al. 2015. "A comprehensive analysis of groundwater quality in the Barnett Shale region." *Environmental science & technology*, 49(13): 8254–8262.

Hill, Elaine L. 2013a. "The Impact of Oil and Gas Extraction on Infant Health in Colorado." *Charles H. Dyson School of Applied Economics and Management, Cornell University, Working Paper*.

Hill, Elaine L. 2013b. "Shale Gas Development and Infant Health: Evidence from Pennsylvania." *Charles H. Dyson School of Applied Economics and Management, Cornell University, Working Paper*.

Hill, Elaine, and Lala Ma. 2017. "Shale Gas Development and Water Quality." *American Economic Review, Papers and Proceedings*, 107(5): 522–525.

Jackson, Robert B, Avner Vengosh, Thomas H Darrah, Nathaniel R Warner, Adrian Down, Robert J Poreda, Stephen G Osborn, Kaiguang Zhao, and Jonathan D Karr. 2013. "Increased stray gas abundance in a subset of drinking water wells near Marcellus shale gas extraction." *Proceedings of the National Academy of Sciences*, 110(28): 11250–11255.

Johnson, Rucker C., and Robert F. Schoeni. 2011. "The influence of early-life events on human capital, health status, and labor market outcomes over the life course." *The BE journal of economic analysis and policy*, 11(3): .

Kearney, Melissa S, and Riley Wilson. 2017. "Male Earnings, Marriageable Men, and Nonmarital Fertility: Evidence from the Fracking Boom."Technical report, National Bureau of Economic Research.

Kuwayama, Yusuke, Sheila M Olmstead, and Alan Krupnick. 2013. "Water resources and unconventional fossil fuel development: linking physical impacts to social costs."Technical report, Resources for the Future.

Lyverse, M.A., and M.D. Unthank. 1988. "Assessment of Ground-Water Contamination in the Alluvial Aquifer near West Point, Kentucky."Technical report, US Geological Survey.

Mason, Charles F, Lucija Muehlenbachs, and Sheila M Olmstead. 2015. "The economics of shale gas development." *Resources for the Future Discussion Paper* 14–42.

McKenzie, Lisa M, Ruixin Guo, Roxana Z Witter, David A Savitz, Lee S Newman, and John L Adgate. 2014. "Birth Outcomes and Maternal Residential Proximity to Natural Gas Development in Rural Colorado." *Environmental Health Perspectives*, 122(4): , p. 412.

Muehlenbachs, Lucija. 2015. "A dynamic model of cleanup: Estimating sunk costs in oil and gas production." *International Economic Review*, 56(1): 155–185.

Muehlenbachs, Lucija, Elisheba Spiller, and Christopher Timmins. 2015. "The Housing Market Impacts of Shale Gas Development." *American Economic Review*, 105(12): 3633–59.

Myers, Tom. 2012. "Potential contaminant pathways from hydraulically fractured shale to aquifers." *Ground Water*, 50(6): 872–882.

Neidell, Matthew. 2009. "Information, Avoidance behavior, and health the effect of ozone on asthma hospitalizations." *Journal of Human Resources*, 44(2): 450–478.

Oreopoulos, Philip, Mark Stabile, Randy Walld, and Leslie Roos. 2008. "Short-, Medium-, and Long-Term Consequences of Poor Infant Health An Analysis Using Siblings and Twins." *Journal of Human Resources*, 43(1): 88–138.

Osborn, Stephen, Avner Vengosh, Nathaniel Warner, and Robert Jackson. 2011. "Methane Contamination of Drinking Water Accompanying Gas Well Drilling and Hydraulic Fracturing." *Proceedings of the National Academy of Sciences*, 108(20): 8172–8176.

PA DEP, Department of Environmental Protection. 2000. "Pennsylvania's Plan for Addressing Problem Abandoned Wells and Orphaned Wells." URL: http://www.elibrary.dep.state.pa.us/dsweb/Get/Version-48262/, Accessed January, 2012.

Saiers, James E., and Erica Barth. 2012. "Potential Contaminant Pathways from Hydraulically Fractured Shale Aquifers." *Ground Water*, 50(6): 826–828.

Schlenker, Wolfram, and Michael J Roberts. 2009. "Nonlinear Temperature Effects Indicate Severe Damages to US Crop Yields under Climate Change." *Proceedings of the National Academy of sciences*, 106(37): 15594–15598.

Schultz, T Paul. 1985. "Changing world prices, women's wages, and the fertility transition: Sweden, 1860-1910." *Journal of Political Economy*, 93(6): 1126–1154.

Stacy, Shaina L, LuAnn L Brink, Jacob C Larkin, Yoel Sadovsky, Bernard D Goldstein, Bruce R Pitt, and Evelyn O Talbott. 2015. "Perinatal outcomes and unconventional natural gas operations in Southwest Pennsylvania." *PloS one*, 10(6): , p. e0126425.

Stringfellow, William T, Jeremy K Domen, Mary Kay Camarillo, Whitney L Sandelin, and Sharon Borglin. 2014. "Physical, chemical, and biological characteristics of compounds used in hydraulic fracturing." *Journal of hazardous materials*, 275 37–54.

Timmins, Christopher, and Ashley Vissing. 2015. "Environmental Justice and Coasian Bargaining: The role of race and income in lease negotiations for shale gas."Technical report, Working Paper.

U.S. Environmental Protection Agency. 2016. "Hydraulic Fracturing for Oil and Gas: Impacts from the Hydraulic Fracturing Water Cycle on Drinking Water Resources in the United States."Technical report.

US House of Representatives. 2011. *Chemicals Used in Hydraulic Fracturing.* URL: http://www.conservation.ca.gov/dog/general_information/Documents/Hydraulic%20Fracturing%20Report%204%2018%2011.pdf, (Accessed: 2017-03-03).

Vengosh, Avner, Robert B Jackson, Nathaniel Warner, Thomas H Darrah, and Andrew Kondash. 2014. "A critical review of the risks to water resources from unconventional shale gas development and hydraulic fracturing in the United States." *Environmental science & technology.*

Vengosh, Avner, Nathaniel R Warner, Andrew Kondash, Jennifer S Harkness, Nancy Lauer, Romain Millot, Wolfram Kloppman, and Thomas H Darrah. 2015. "Isotopic Fingerprints for Delineating the Environmental Effects of Hydraulic Fracturing Fluids." *Procedia Earth and Planetary Science*, 13 244–247.

Warner, Nathaniel, Robert Jackson, Thomas Darrah, Stephen Osborn, Adrian Down, Kaiguang Zhao, Alissa White, and Avner Vengosh. 2012. "Geochemical Evidence for Possible Natural Migration of Marcellus Formation Brine to Shallow Aquifers in Pennsylvania." *Proceedings of the National Academy of Sciences*, 109(30): 11961–11966.

Whitacre, J. V, and J. B. Slyder. 2015. "Carnegie Museum of Natural History Pennsylvania Unconventional Natural Gas Wells Geodatabase (v. 2015 Q4)." URL: http://maps.carnegiemnh.org/index.php/projects/unconventional-wells/.

Wrenn, Douglas H, H Allen Klaiber, and Edward C Jaenicke. 2016. "Unconventional Shale Gas Development, Risk Perceptions, and Averting Behavior: Evidence from Bottled Water Purchases." *Journal of the Association of Environmental and Resource Economists*, 3(4): 779–817.

Wright, Peter R., Peter B. McMahon, David K. Mueller, and Melanie L. Clark. 2012. "Groundwater quality and quality control data for two monitoring wells near Pavillion, Wyoming, April and May 2012."Technical report, US. Geological Survey, Retrieved 2/9/2012.

Yan, Beizhan, Martin Stute, Reynold A Panettieri, James Ross, Brian Mailloux, Matthew J Neidell, Lissa Soares, Marilyn Howarth, Xinhua Liu, Pouné Saberi et al. 2016. "Association of groundwater constituents with topography and distance to unconventional gas wells in NE Pennsylvania." *Science of The Total Environment.*