Bonetti, P., Leuz, C., & Michelon, G. (2021). Large-sample evidence on the impact of unconventional oil and gas development on surface waters. *Science*, 373(6557), 896-902.
<https://doi.org/10.1126/science.aaz2185>

Peer reviewed version

Link to published version (if available):
[10.1126/science.aaz2185](https://doi.org/10.1126/science.aaz2185)

[Link to publication record in Explore Bristol Research](#)
PDF-document

This is the final published version of the article (version of record). It first appeared online via Science at <https://doi.org/10.1126/science.aaz2185> .Please refer to any applicable terms of use of the publisher.

University of Bristol - Explore Bristol Research

General rights

This document is made available in accordance with publisher policies. Please cite only the published version using the reference above. Full terms of use are available:
<http://www.bristol.ac.uk/red/research-policy/pure/user-guides/ebr-terms/>

Title: Large-Sample Evidence on the Impact of Unconventional Oil and Gas Development on Surface Waters

Accepted version

Science Aug 2021: Vol. 373, Issue 6557, pp. 896-902, DOI: [10.1126/science.aaz2185](https://doi.org/10.1126/science.aaz2185)

One Sentence Summary: First evidence that unconventional oil and gas development is related to increased salt concentrations in surface waters across several U.S. shales and watersheds.

Authors: Pietro Bonetti¹, Christian Leuz^{2*}, Giovanna Michelin³

Affiliations:

¹ University of Navarra - IESE, 21 Av. Pearson, Barcelona, 08007, Spain.

² University of Chicago & NBER, 5807 S. Woodlawn Ave, Chicago, IL 60614, USA.

³ University of Bristol, 15-19 Tyndall's Park Road, Bristol, BS8 1PQ, UK

*Correspondence to: cleuz@chicagobooth.edu

Abstract: The impact of unconventional oil and gas development on water quality is a major environmental concern. We built a large, geo-coded database that combines surface water measurements with horizontally drilled wells stimulated by hydraulic fracturing (HF) for several shales to examine whether temporal and spatial well variation is associated with anomalous salt concentrations in U.S. watersheds. We analyzed four ions that could indicate water impact from unconventional development. We found very small concentration increases associated with new HF wells for barium, chloride and strontium, but not bromide. All ions showed larger, but still small-in-magnitude increases 91-180 days after well spudding. Our estimates were most pronounced for wells with larger amounts of produced water, wells located over high-salinity formations, and wells closer and likely upstream from water monitors.

The rise of shale gas and tight oil development has triggered a major public debate about such unconventional development, in which horizontal drilling is combined with hydraulic fracturing (HF). Hydraulic fracturing is the high-pressure injection of water mixed with chemical additives and propping agents like sand to create fractures in low-permeability formations, allowing oil or gas to flow. Additives in the HF fluids vary with the geological characteristics of the formation and by operator, but the fluid mix usually contains friction reducers, surfactants, scale inhibitors, biocides, gelling agents, gel breakers and inorganic acid (1, 2). HF wells produce large amounts of wastewater, which initially consists of flowback of HF fluids and over time increasingly of produced water from the deep formations. The latter brine is naturally occurring water, into which organic and inorganic constituents from the formation have dissolved, resulting in high salt concentrations (1, 3-5).

While unconventional oil and gas (O&G) development has been important for energy production (6), we do not fully understand the associated environmental and social risks (7-11). These risks include hydrocarbon emissions, water usage, and pollution, along with potential human and ecological health consequences (7, 10, 12-16). Among these, the impact of unconventional O&G and HF on water quality remains a key concern (2, 8, 9, 17-20). In the U.S., one reason for this concern is that HF is exempt from the Underground Injection Control provisions of the Safe Drinking Water Act (8). Around the world, unconventional drilling has either just been introduced or is being considered by many countries with uncertain impacts on water quality (16).

The Environmental Protection Agency (EPA) reviewed and synthesized scientific evidence concerning the impact of HF on U.S. water resources. The final report concluded that hydraulic fracturing activities can impact drinking water resources under some circumstances (17), but the report did not identify widespread evidence of contamination. Groundwater studies primarily

examine contamination from stray gas or deep formation brines, which could occur because of cementing or casing failures or due to brine migration to shallow aquifers through faults or other pre-existing pathways (2, 21, 22). Instances of stray gas contamination have been found in Pennsylvania connected to shale gas development of the Marcellus Shale (23-29), but not in Arkansas for the Fayetteville Shale (30). Geochemical evidence of gas contamination has also been documented for the Barnett Shale in Texas (24) and the Denver-Julesburg basin in Colorado (31). Studies of brine migration from deep formations, mostly in Northeastern Pennsylvania, have provided mixed evidence (23, 32, 33). No evidence of brine contamination of groundwater has been documented for the Fayetteville Shale (30). For Pennsylvania, increases in shale gas-related contaminants have been documented at ground-water intake locations of community water systems that are in close proximity to shale gas wells (18).

For surface water, the evidence is more limited. Instances of contamination have been ascribed primarily to discharges of inadequately treated wastewater, HF fluid leaks as well as spills and other mishandling of flowback and produced waters (2, 4, 8, 17, 20, 34, 35). Specifically, increased chloride and bromide concentrations downstream of effluents from wastewater treatment plants have been found in Western Pennsylvania up to 2011, when the release of wastewaters from unconventional wells into streams through municipal wastewater treatment plants was not prohibited (2, 4, 36). Further, a high frequency of brine spills in North Dakota has resulted in elevated levels of salts and other contaminants in surface waters up to 4 years after the spills occurred (37). A large-sample statistical examination of the effects of shale gas development activities on surface water in Pennsylvania found higher chloride concentrations downstream from wastewater treatment facilities and an association between gas well density in a watershed and increased total suspended solid (but not chloride) concentrations during the period 2000-2011 (38).

The authors of this study suggest that both insufficient wastewater treatment and building infrastructure for unconventional O&G extraction could explain the results. Evidence also exists for barium concentrations in Pennsylvania being higher in areas with unconventional wells than in areas without them, but the authors of this study point out that this evidence cannot be uniquely ascribed to unconventional wells as it could also reflect the presence of basin brines or a sulfate decrease in acid rains (39). In sum, prior studies document localized instances of surface water contamination related to unconventional O&G development, with spills and leaks as the most common pathway (20).

We investigated the potential impact of unconventional O&G development on surface water quality using a large-sample statistical approach. We combined a geo-coded database of 46,479 HF wells from 24 shales with 60,783 surface water measurements over eleven years (2006–2016) across 408 watersheds (HUC10s) with HF activity (Fig. 1, Fig. S2, Table S1-S3). Our analysis focuses on concentrations of bromide (Br^-), chloride (Cl^-), barium (Ba) and strontium (Sr) in watersheds exposed to unconventional O&G development. We chose these four ions for three reasons. First, they are usually found in high concentrations in flowback and produced water from HF wells and hence could indicate surface water impact, if and when it exists (1, 2, 4, 5, 8, 20, 27, 33, 36, 39, 40). Moreover, unlike some organic components of HF fluids, the four ions do not experience biodegradation and their presence has been measured several years after HF spill events (20, 37). Thus, an analysis of ions is a likely mode of detection (2, 4). Second, many of the water quality concerns associated with HF wastewaters are related to the chemistry of the deep formation brines as well as the salinity in flowback and produced water (8). Third, the four ions are measured in many watersheds with reasonable frequency, which is not the case for other potential signatures of HF wastewater. As high salt concentrations can also occur in surface waters for many natural

and anthropogenic reasons such as brine migration or road de-icing (2, 4, 33, 39), we need sufficient data to estimate reliable, local and time-varying baselines for the background ion concentrations.

We construct these baselines with regression analysis and then exploit temporal and spatial variation in the spudding of HF wells within and across U.S. watersheds to identify anomalous changes in ion concentrations associated with newly spudded HF wells in the same watersheds. Our regression model includes temperature and precipitation control variables as well as an extensive set of fixed effects that construct local and time-varying baselines for background ion concentrations (41, 42). Specifically, our model allows for arbitrary monthly variation in the average background ion concentrations across sub-basins (HUC8) and within a sub-basin over time. This flexible regional baseline controls for sub-basin differences in water quality, geochemistry, salinity, climate, water body types, or economic activity, and also for over-time changes in sub-basin concentrations due to seasons, weather patterns and related road de-icing, salinization trends, or economic development. Our model also has a local baseline, using each water monitoring station as its own control, which accounts for arbitrary differences in average local ion concentrations. Our model combines these two baselines and the weather control variables to estimate the association between anomalous concentration changes and new HF wells in the same watersheds (42).

Our model explains over 80% (in many cases over 90%) of the variation in ion concentrations across watersheds and through time (Table S4), suggesting the model estimates precise baselines for background ion concentrations. We estimated the association between new HF wells and ions concentrations at the watershed level using a variable that counts the number of HF wells in a watershed at a point in time (*#wellsHUC10*). We estimated our regression model for all U.S.

watersheds with HF wells and, separately, for Pennsylvania because Pennsylvania accounts for almost 41% of the sample. We found a robust association between new HF wells in a watershed and elevated ion concentrations in its surface waters (Fig. 2 and Table S4). For watersheds in Pennsylvania (PA), the coefficients on *#wellsHUC10* are positive for all ions and significant for three of them (Table S4, Column 2, HUC8 model, Br^- : 0.00019, $P = 0.865$; Cl^- : 0.00071, $P = 0.031$; Ba: 0.00038, $P = 0.086$; Sr: 0.00041, $P < 0.001$). The lack of significance for Br^- could reflect measurement issues (42). For watersheds throughout the U.S. (ALL), the coefficients on *#wellsHUC10* are generally comparable to those for Pennsylvania in terms of magnitude and significance, except for Ba, which is rarely measured outside of Pennsylvania and for which results are weaker (Fig. 2).

To gauge the magnitude of the estimated effects, we multiply each coefficient by the respective sample mean ion concentration and the average number of wells per watershed to obtain the ion concentration increase in the average HUC10 implied by our estimation (HUC10 impact in Fig 2). With this approach, and focusing on coefficients from the HUC8 specification, we estimated an average increase of Cl^- by 1,322.44 $\mu\text{g/L}$ for PA and 2,232.55 $\mu\text{g/L}$ for ALL, Ba by 1.61 $\mu\text{g/L}$ for PA, and Sr by 5.19 $\mu\text{g/L}$ for PA and 8.88 $\mu\text{g/L}$ for ALL (Fig. 2). These magnitudes are very small but need to be interpreted in the context of our analysis. First, the *#wellsHUC10* coefficient is by construction a per-well estimate but this does not imply that each well is associated with a concentration increase. It is an average over all wells. Second, by using measurements from all monitors in a given watershed, the estimated well-ion association reflects the average exposure of monitors in the watershed. However, some monitors could be very far or upstream from wells, in which case they should not be exposed or affected, lowering the average. For these reasons, the

impact estimates in Fig. 2 are expected to be small; they should increase when the analysis focuses on the most relevant water measurements as reported below.

We found robust results in different sensitivity analyses (Table S5). (i) Estimating separate *#wellsHUC10* coefficients for watersheds in and outside of Pennsylvania shows that the findings for Pennsylvania and the other U.S. states are similar. (ii) Estimating separate effects for different time periods shows similar results over time. The coefficients in later periods tend to be smaller but higher in significance presumably because the frequency of water measurements increases over time. (iii) Our results are similar when the model is estimated over all watersheds within a sub-region (including those without HF wells), albeit in some cases slightly weaker, possibly because this specification uses less relevant baselines for the background ion concentrations. (iv) Adding further controls for snow (to account for related discharge of road salts) or for within-watershed seasonality does not alter the findings, suggesting that the monthly sub-basin baselines already control for local weather patterns. (v) Our results are also robust to alternative modelling choices, e.g., scaling the cumulative number of wells by watershed size as in (38); estimating the model with weighted least squares (WLS) to give more weight to observations, for which we have more readings to estimate the monthly sub-basin baselines; and using alternative transformations to address skewness in ion concentrations.

We also investigated whether the results could be driven by HF-related patterns in the frequency of water monitoring (e.g., more measurements shortly after spud dates or closer to wells). However, we found no evidence that water measurement is systematically related to new wells in a watershed (Table S6). We analyzed three other water quality proxies (dissolved oxygen, phosphorus and fecal coliforms) that are frequently measured but not as indicative of HF-related impacts. Concentration levels of these proxies could be related to other economic activities with

water impacts such as agriculture (43). We used this analysis to gauge how well our model controls for economic activity and other potential confounds. The estimated *#wellsHUC10* coefficients for the three analytes are not different from zero (Table S7), which is in contrast to the results for the ion concentrations we chose for the main analysis (Table S4).

Up to this point, our analysis estimated the long-run association between HF wells and ion concentrations because we did not restrict the sample and the estimation to a particular period after well spudding. However, concentration increases could be stronger early on and fade over time. Hence, we estimated the association in specific time windows around a new well spud date, allowing us to map out the estimates through time. For this temporal analysis, we modified eq. S1 by replacing *#wellsHUC10* with several time-specific well counts, defined for the following time windows measured in days: [-180, -91], [-90, 0], [1, 90], [91, 180], [181, 360], [> 360]. The coefficients were estimated relative to measurements collected 180 days or more before a new well spudding (Table S8). We found increases in ion concentrations 91–180 days after the spud date, consistently for all four ions, in Pennsylvania and all U.S. watersheds. In Fig. 3, we plotted the coefficients, estimated over all watersheds, for each window together with the 95% confidence interval. For the [91, 180] window, the well count coefficients are significant for all four ions (Fig. 3 and Table S8, Panel B, Br^- : 0.01095, $P = 0.036$; Cl^- : 0.00401, $P = 0.022$; Ba: 0.00347, $P = 0.017$; Sr: 0.00289, $P = 0.015$). These ion concentration increases in the [91, 180] window are at least one order of magnitude larger than the long-run estimates in Table S4 (shown as a red dot in Fig. 3 for comparison). The coefficients for the [91, 180] window (Fig. 3) correspond to an average (short-run) increase of 178.64 $\mu\text{g/L}$ for Br, 16,014.30 $\mu\text{g/L}$ for Cl^- , 15.46 $\mu\text{g/L}$ for Ba, and 71.34 $\mu\text{g/L}$ for Sr per watershed with HF wells. Water measurement is often sparse and, as shown earlier, does not increase around the well spud dates. Thus, most measurements naturally fall into the

benchmark period [<-180] days (for which no coefficient is estimated) or in the period beyond 360 days after spudding, which explains why the *#wellsHUC10 [>360]* coefficients always have the tightest confidence intervals. These coefficients are, as expected, comparable to the long-run estimates from Table S4 (red dots in Fig. 3). For all other coefficients, we have far fewer observations and hence their confidence intervals are wider.

It is useful to interpret the increases in ion concentrations between 91–180 days after new well spuddings in the context of the well development and HF process. In our samples, the average time span between the spud date and well completion is 103 days, consistent with (31). Thus, the concentration increases occur after the average HF well is completed and during the early phases of production, when large amounts of flowback and produced water are collected. Although the temporal evidence alone does not identify the exact mechanism for the results, it links elevated concentrations to the unconventional O&G development process.

We conducted two tests that further explore the mechanism by focusing on the role of flowback and produced water at the beginning of production. First, we investigated whether the well-ion associations differ depending on the amount of water a well produces after the spud date. We coded two well count variables at the watershed level, (i) the number of wells with an above-median amount of produced water (*#wellsHUC10_High_Prod_Water*) and (ii) the number of wells with a below-median amount of produced water (*#wellsHUC10_Low_Prod_Water*). We defined the medians by sub-basin and year to account for regional differences as well as potential changes in HF technology over time. For watersheds where wells generate larger amounts of produced water, we found significant coefficients for Cl^- and Sr in all specifications, for Ba in Pennsylvania and for Br^- in one specification (Table S9, Column 1: Cl^- : 0.00057, $P = 0.010$; Ba: 0.00047, $P = 0.064$; Sr: 0.00041, $P = 0.012$; Column 2: Cl^- : 0.00071, $P = 0.031$; Ba: 0.00037, $P = 0.084$; Sr: 0.00041,

$P < 0.001$; Column 3: $\text{Br}^- = 0.00035$, $P = 0.085$; $\text{Cl}^- : 0.00057$, $P = 0.044$; Sr: 0.00041 , $P = 0.002$; Column 4: $\text{Cl}^- : 0.00058$, $P = 0.080$; Sr: 0.00037 , $P < 0.001$). These coefficients are larger in magnitude than the respective low-produced-water counterparts. Admittedly, the amount of produced water is likely correlated with other well and watershed characteristics and hence differences in coefficient magnitudes between the two groups do not reflect produced water levels alone. Even so, this evidence ties the elevated ion concentrations more closely to HF activity.

Second, we performed an analysis that exploits sub-basin variation in the regional geochemistry and the salinity of deep formations (1, 5, 40). The idea was to explore whether the estimated associations are stronger in areas where HF wells are expected to generate produced waters with higher salinity. We used data from the U.S. Geological Survey Produced Waters Geochemical Database and identified sub-basins for which produced waters in previous drillings exhibited total dissolved solids (TDS) concentrations above (below) the median, indicating higher (lower) salinity of the deep formations. We created two well count variables at the watershed level, (i) the number of wells in sub-basins with an above-median TDS level (*#wellsHUC10_High_Salinity*) and (ii) the number of wells in sub-basins with a below-median TDS level (*#wellsHUC10_Low_Salinity*). The coefficients in the high-salinity group are positive and larger in magnitude than the coefficients in the low-salinity group (Table S10, $\text{Br}^- : 0.00114$, $P = 0.189$; $\text{Cl}^- : 0.00120$, $P = 0.082$; Ba: 0.00064 , $P = 0.003$; Sr: 0.00040 , $P < 0.001$). Thus, the association between new HF wells and ion concentrations is most pronounced in sub-basins where deep formations exhibit higher levels of salinity. Importantly, natural brine seepage in high salinity areas cannot explain this result because the model uses the average concentration levels at the monitoring station as its own control. The result implies a (statistical) link between elevated concentrations and HF wells in areas where produced waters have higher salinity, which is consistent with produced water being part of the

mechanism.

As wells can be far from the closest monitor (average = 10.3 km, median = 8.2 km in our sample), we examined whether the associations between HF wells and ion concentrations are more pronounced when wells and monitors are closer together. However, the sparsity of water measurements makes such distance gradient analyses challenging. We therefore estimate results for different distance bins, [0–5 km], [0–10 km], and so on up to [0–30 km], using either Cl^- or pooling observations for all ions in one model. In both cases, we found a negative distance gradient (Fig. S5), meaning the coefficients are largest when we estimated them for wells within 5 km from a monitoring station, and then steadily decline when we expanded the widths of the distance bins and included wells that are further away. Due to sparse data, our confidence intervals are widest when we included only close-by wells, but then become smaller as statistical power increases with the expansion of the distance bins.

To address the data sparsity, and still tighten the analysis, we combined dimensions (such as time, distance, and direction of surface water flows) and then either partitioned or restricted the sample. To partition the sample, we defined watersheds as “high type” when there are relatively (i) more water measurements around new well spuddings, (ii) more monitors in close proximity to new wells, and (iii) more monitors that are likely downstream to new wells (37). We estimated an alternative version of eq. S1, in which we replace *#wellsHUC10* with two non-overlapping well counts, counting the cumulative number of wells spudded in high-type and low-type watersheds, respectively. The associations between HF wells and elevated ion concentrations stem mostly from the high-type watersheds (Table S11). This finding is reassuring because it shows that the results come from the types of watersheds that a priori have a higher chance of showing a relation between HF wells and ion concentrations if it exists (e.g., because monitors sit closer to HF wells).

We also conducted a separate analysis restricting the sample to most relevant water measurements. We paired all wells and monitoring stations within a watershed, so that we could make unique determinations for each well-monitor pair with respect to time, distance, and direction of water flows. We then restricted the sample to water measurements that are (i) either taken before or up to 360 days after the spud date, (ii) from monitors within a 15 km radius of the well, and (iii) from monitors for which the paired well likely sits upstream, and estimated eq. S2 using Weighted Least Squares regressions (42). The variable of interest, *Post Spud*, estimates the average change in the respective ion concentration after a new well spudding, using each well-monitor pair as its own control and after controlling for temperature, precipitation, and monthly variation in background ion concentrations at the sub-basin level. We found elevated ion concentrations in the year after new well spuddings (Fig. 4). In the ALL sample, we estimated positive and significant coefficients on *Post Spud* for three ions, but not for Br^- (Table S12, Panel C, Br^- : -0.00577, $P = 0.499$; Cl^- : 0.02044, $P = 0.045$; Ba: 0.01585, $P = 0.048$; Sr: 0.02014, $P = 0.098$). To gauge the magnitude of the estimates, we calculated the 360-day impact on the average watershed, multiplying the respective coefficient by the mean ion concentration and the average number of new wells spudded per year in the average watershed. For the ALL sample, the average 360-day concentration increases are 47,338.29 $\mu\text{g/L}$ for Cl^- , 49.17 $\mu\text{g/L}$ for Ba and 282.67 $\mu\text{g/L}$ for Sr (Fig. 4). These increases are much larger than the long-run increases (Fig. 2), but still well below the EPA maximum contaminant and health advisory levels (note to Fig. 2). The estimated coefficients are smaller using Ordinary Least Squares regressions (Table S12, Panel D) but even these are still an order of magnitude larger than the respective coefficients for the overall sample (Table S4). As sensitivity, we also estimated regressions using observations from “never-treated” monitors that are far or upstream from any well in the watershed as control group (Table S12, Panel E). The

results were similar to those in Fig. 4. Importantly, the well-monitor pair analysis (Fig. 4 and Table S12) showed that estimated well-ion associations strengthened when we focused on the most relevant water measurements.

In sum, our analysis reveals a robust association between new HF wells and elevated concentrations of Cl^- , Ba, and Sr in same-watershed surface waters. Our results for Br^- are generally weaker and often insignificant. The well-ion association was more pronounced when we estimated it using water measurements from monitors that are close to and likely downstream from wells and from the early phases of production when wells generate large amounts of flowback and produced water. Although this evidence is based on associations, and as such not causal, the regression analysis controls for a large number of observed and unobserved factors and models background ion concentrations in a flexible and comprehensive manner. Thus, we find it difficult to explain the results by other factors. The confluence of our findings is consistent with unconventional O&G development driving the anomalous surface water concentrations identified by our model.

Overall, our evidence complements prior studies on the impacts of unconventional O&G development on water quality (e.g., 2, 8) and extends studies using a similar large-scale approach for the Marcellus shale (18, 38). Our statistical approach does not reveal the exact mechanism for the association between new HF wells and concentrations increases. There are three potential channels that are particularly relevant in our context (2, 4, 8, 12, 17): (i) on-site accidents, leaks and spills of HF fluids, flowback and produced water, including those related to pits (overflow, unlined pits, tears in liners), brine trucking, and long-term leaching of soils after spills (4, 12, 20, 35, 37), (ii) direct disposal of untreated wastewater from HF operations (unauthorized or permitted) (4, 8, 12, 35), and (iii) inadequate treatment of wastewater at disposal sites (2, 36, 38).

We do not distinguish between these specific channels. However, the consistent increases in ion concentrations shortly after well completion, for wells with large amounts of produced water and for wells located in areas with high-salinity deep formations (Fig. 3, Tables S9, S10) point to the handling of flowback and produced waters, including leaks and spills, being part of the mechanism for the well-ion association we documented. We explored other explanations and found that our results for HF wells (i) do not reflect the presence of facilities accepting HF wastewater in a watershed (Table S13), (ii) are not observed for conventional wells (Table S14), and (iii) are still present when we control for a large number of previously documented spills (Table S15). Unknown or undetected spills and leaks could still explain our findings. Independent of the exact mechanism, our results are relevant because we show the association between unconventional O&G development and elevated ion concentrations extends to many watersheds over several U.S. shales.

We acknowledge that the long-run impact estimates we documented using all watershed monitors are very small in magnitude. However, not all wells are close to surface water and not all monitors are in locations where they could detect an effect. In addition, any impact likely declines over time. Consistent with these arguments, we estimated larger associations when we restricted the analysis to water measurements that are taken within a year of well spudding and from monitoring stations that are closer and likely downstream from wells. But even the 360-day concentration increases implied by these estimates do not exceed EPA limits or health advisory levels for the ions. We also note that water measurements are predominantly from rivers and hence dilution is another factor to consider when interpreting the magnitude of the associations we estimated.

Our large-scale, statistical approach was constrained by the ions included in public databases as well as the sparsity of water quality data (2, 4, 20, 27, 41, 44). The former means we could not

examine other analytes in HF fluids or related to unconventional O&G development that are potentially more dangerous than salts. The latter implies that we could not perform more granular analyses that would better identify the mechanism or a causal link. Both limitations of our analysis highlight that investigations of surface water impacts from unconventional O&G development would be greatly facilitated if there were more targeted water measurements of relevant analytes in close proximity to and timed around the development of new HF wells.

References and Notes:

1. E. Barbot, N. S. Vidic, K. B. Gregory, R. D. Vidic, Spatial and temporal correlation of water quality parameters of produced waters from devonian-age shale following hydraulic fracturing. *Envir. Sci. Technol.* **47**, 2562-2569 (2013). doi: [10.1021/es304638h](https://doi.org/10.1021/es304638h)
2. R.D. Vidic, S. L. Brantley, J. M. Vandebossche, D. Yoxheimer, J. D. Abad, Impact of Shale Gas Development on Regional Water Quality. *Science* **340**, 1235009 (2013). doi: [10.1126/science.1235009](https://doi.org/10.1126/science.1235009)
3. K. B. Gregory, R. D. Vidic, D. A. Dzombak, Water Management Challenges Associated with the Production of Shale Gas by Hydraulic Fracturing. *Elements* **7**, 181-186 (2011). doi: [10.2113/gselements.7.3.181](https://doi.org/10.2113/gselements.7.3.181)
4. S. L. Brantley *et al.*, Water resource impacts during unconventional shale gas development: The Pennsylvania experience. *Int. J. Coal Geo* **126**, 140-156 (2014). doi: [10.2113/gselements.7.3.181](https://doi.org/10.2113/gselements.7.3.181)
5. J. Rosenblum *et al.*, Temporal characterization of flowback and produced water quality from a hydraulically fractured oil and gas well. *Sci. Total Environ.* **596-597**, 369-377 (2017). doi: [10.1016/j.scitotenv.2017.03.294](https://doi.org/10.1016/j.scitotenv.2017.03.294)
6. U. S. Energy Information Agency (EIA), "Annual Energy Outlook 2018 with projections to 2050," (EIA, 2018; <https://www.eia.gov/outlooks/aeo/pdf/AEO2018.pdf>).
7. T. Colborn, C. Kwiatkowski, K. Schultz, M. Bachran, Natural Gas Operations from a Public Health Perspective. *Hum. Ecol. Risk Assess.* **17**, 1039-1056 (2011). doi: [10.1080/10807039.2011.605662](https://doi.org/10.1080/10807039.2011.605662)
8. A. Vengosh, R. B. Jackson, N. Warner, T. H. Darrah, A. Kondash, A Critical Review of the Risks to Water Resources from Unconventional Shale Gas Development and Hydraulic Fracturing in the United States. *Envir. Sci. Technol.* **48**, 8334-8348 (2014). doi: [10.1021/es405118y](https://doi.org/10.1021/es405118y)
9. Y. Kuwayama, S. Olmstead, A. Krupnick, Water Quality and Quantity Impacts of Hydraulic Fracturing. *Curr Sustainable Renewable Energy Rep* **2**, 17-24 (2015). doi: [10.1007/s40518-014-0023-4](https://doi.org/10.1007/s40518-014-0023-4)
10. J. urrie, M. Greenstone, K. Meckel, Hydraulic fracturing and infant health: New evidence from Pennsylvania. *Sci. Adv.* **3**, e1603021 (2017). doi: [10.1126/sciadv.1603021](https://doi.org/10.1126/sciadv.1603021)
11. E. G. Elliott, A. S. Ettinger, B. P. Leaderer, M. B. Bracken, N. C. Deziel, A systematic evaluation of chemicals in hydraulic-fracturing fluids and wastewater for reproductive and developmental toxicity. *J. Exposure Sci. Environ. Epidemiol.* **27**, 90-99 (2017). doi: [10.1038/jes.2015.81](https://doi.org/10.1038/jes.2015.81)
12. S. Entekin, M. Evans-White, B. Johnson, E. Hagenbuch, Rapid expansion of natural gas development poses a threat to surface waters. *Frontiers Ecol. Environ.* **9**, 503 (2011). doi: [10.1890/110053](https://doi.org/10.1890/110053)
13. L. M. McKenzie, R. Z. Witter, L. S. Newman, J. L. Adgate, Human health risk assessment of air emissions from development of unconventional natural gas resources. *Sci. Total Environ.* **424**, 79-87 (2012). doi: [10.1016/j.scitotenv.2012.02.018](https://doi.org/10.1016/j.scitotenv.2012.02.018)
14. Z. H. Bowen *et al.*, Assessment of surface water chloride and conductivity trends in areas of unconventional oil and gas development - Why existing national data sets cannot tell us what we would like to know. *Water Resour. Res.* **51**, 704-715 (2015). doi: [10.1002/2014WR016382](https://doi.org/10.1002/2014WR016382)

15. C. F. Mason, L. A. Muehlenbachs, S. M. Olmstead, The Economics of Shale Gas Development. *Annu. Rev. Resour. Econ.* **7**, 269-289 (2015). doi: [10.1146/annurev-resource-100814-125023](https://doi.org/10.1146/annurev-resource-100814-125023)
16. A. J. Kondash, N. E. Lauer, A. Vengosh, The intensification of the water footprint of hydraulic fracturing. *Sci. Adv.* **4**, eaar5982 (2018). doi: [10.1126/sciadv.aar5982](https://doi.org/10.1126/sciadv.aar5982)
17. U.S. Environmental Protection Agency (EPA), "Hydraulic Fracturing for Oil and Gas: Impacts from the Hydraulic Fracturing Water Cycle on Drinking Water Resources in the United States," (EPA, 2016).
18. E. Hill, L. Ma, Shale Gas Development and Drinking Water Quality. *Am. Econ. Rev.* **107**, 522-525 (2017). doi: [10.1257/aer.p20171133](https://doi.org/10.1257/aer.p20171133)
19. K. O. Maloney *et al.*, A detailed risk assessment of shale gas development on headwater streams in the Pennsylvania portion of the Upper Susquehanna River Basin, U.S.A. *Sci. Total Environ.* **610-611**, 154-166 (2018). doi: <https://doi.org/10.1016/j.scitotenv.2017.07.247>
20. A. Agarwal *et al.*, Assessing Contamination of Stream Networks near Shale Gas Development Using a New Geospatial Tool. *Envir. Sci. Technol.* **54**, 8632-8639 (2020). doi: [10.1021/acs.est.9b06761](https://doi.org/10.1021/acs.est.9b06761)
21. P. B. McMahon *et al.*, Methane and Benzene in Drinking-Water Wells Overlying the Eagle Ford, Fayetteville, and Haynesville Shale Hydrocarbon Production Areas. *Envir. Sci. Technol.* **51**, 6727-6734 (2017). doi: [10.1021/acs.est.7b00746](https://doi.org/10.1021/acs.est.7b00746)
22. P. B. McMahon *et al.*, Geochemistry and age of groundwater in the Williston Basin, USA: Assessing potential effects of shale-oil production on groundwater quality. *App. Geochem.* **125**, 104833 (2021). doi: [10.1016/j.apgeochem.2020.104833](https://doi.org/10.1016/j.apgeochem.2020.104833)
23. S. G. Osborn, A. Vengosh, N. R. Warner, R. B. Jackson, Methane contamination of drinking water accompanying gas-well drilling and hydraulic fracturing. *Proc. Natl. Acad. Sci. U.S.A.* **108**, 8172-8176 (2011). doi: [10.1073/pnas.1100682108](https://doi.org/10.1073/pnas.1100682108)
24. T. H. Darrah, A. Vengosh, R. B. Jackson, N. R. Warner, R. J. Poreda, Noble gases identify the mechanisms of fugitive gas contamination in drinking-water wells overlying the Marcellus and Barnett Shales. *Proc. Natl. Acad. Sci. U.S.A.* **111**, 14076-14081 (2014). doi: [10.1073/pnas.1322107111](https://doi.org/10.1073/pnas.1322107111)
25. R. B. Jackson *et al.*, Increased stray gas abundance in a subset of drinking water wells near Marcellus shale gas extraction. *Proc. Natl. Acad. Sci. U.S.A.* **110**, 11250-11255 (2013). doi: [10.1073/pnas.1221635110](https://doi.org/10.1073/pnas.1221635110)
26. G. T. Llewellyn *et al.*, Evaluating a groundwater supply contamination incident attributed to Marcellus Shale gas development. *Proc. Natl. Acad. Sci. U.S.A.* **112**, 6325-6330 (2015). doi: [10.1073/pnas.1420279112](https://doi.org/10.1073/pnas.1420279112)
27. T. Wen *et al.*, Big Groundwater Data Sets Reveal Possible Rare Contamination Amid Otherwise Improved Water Quality for Some Analytes in a Region of Marcellus Shale Development. *Envir. Sci. Technol.* **52**, 7149-7159 (2018). doi: [10.1021/acs.est.8b01123](https://doi.org/10.1021/acs.est.8b01123)
28. Z. Li *et al.*, Searching for anomalous methane in shallow groundwater near shale gas wells. *J. Contam. Hydrol.* **195**, 23-30 (2016). doi: [10.1016/j.jconhyd.2016.10.005](https://doi.org/10.1016/j.jconhyd.2016.10.005)
29. Z. Li *et al.*, Corrigendum to "Searching for anomalous methane in shallow groundwater near shale gas wells" [*J. Contam. Hydrol.* (195) (December 2016) 23–30]. *J. Contam. Hydrol.* **207**, 50-51 (2017). doi: [10.1016/j.jconhyd.2017.09.009](https://doi.org/10.1016/j.jconhyd.2017.09.009)

30. N. R. Warner *et al.*, Geochemical and isotopic variations in shallow groundwater in areas of the Fayetteville Shale development, north-central Arkansas. *App. Geochem.* **35**, 207-220 (2013). doi: [10.1016/j.apgeochem.2013.04.013](https://doi.org/10.1016/j.apgeochem.2013.04.013)
31. O. A. Sherwood *et al.*, Groundwater methane in relation to oil and gas development and shallow coal seams in the Denver-Julesburg Basin of Colorado. *Proc. Natl. Acad. Sci. U.S.A.* **113**, 8391-8396 (2016). doi: [10.1073/pnas.1523267113](https://doi.org/10.1073/pnas.1523267113)
32. E. W. Boyer, B. R. Swistock, G. Clark, M. Madden, D. E. Rizzo, "The Impact of Marcellus Gas Drilling on Rural Drinking Water Supplies," (The Center for Rural Pennsylvania, 2012).
33. N. R. Warner *et al.*, Geochemical evidence for possible natural migration of Marcellus Formation brine to shallow aquifers in Pennsylvania. *Proc. Natl. Acad. Sci. U.S.A.* **109**, 11961-11966 (2012). doi: [10.1073/pnas.1121181109](https://doi.org/10.1073/pnas.1121181109)
34. K. O. Maloney *et al.*, Unconventional oil and gas spills: Materials, volumes, and risks to surface waters in four states of the U.S. *Sci. Total Environ.* **581-582**, 369-377 (2017).
35. L. A. Patterson *et al.*, Unconventional Oil and Gas Spills: Risks, Mitigation Priorities, and State Reporting Requirements. *Envir. Sci. Technol.* **51**, 2563-2573 (2017). doi: [10.1016/j.scitotenv.2016.12.142](https://doi.org/10.1016/j.scitotenv.2016.12.142)
36. N. R. Warner, C. A. Christie, R. B. Jackson, A. Vengosh, Impacts of shale gas wastewater disposal on water quality in western Pennsylvania. *Envir. Sci. Technol.* **47**, 11849-11857 (2013). doi: [10.1021/es402165b](https://doi.org/10.1021/es402165b)
37. N. E. Lauer, J. S. Harkness, A. Vengosh, Brine Spills Associated with Unconventional Oil Development in North Dakota. *Envir. Sci. Technol.* **50**, 5389-5397 (2016). doi: [10.1021/acs.est.5b06349](https://doi.org/10.1021/acs.est.5b06349)
38. S. M. Olmstead, L. A. Muehlenbachs, J. S. Shih, Z. Chu, A. J. Krupnick, Shale gas development impacts on surface water quality in Pennsylvania. *Proc. Natl. Acad. Sci. U.S.A.* **110**, 4962-4967 (2013). doi: [10.1073/pnas.1213871110](https://doi.org/10.1073/pnas.1213871110)
39. X. Niu *et al.*, Detecting the effects of coal mining, acid rain, and natural gas extraction in Appalachian basin streams in Pennsylvania (USA) through analysis of barium and sulfate concentrations. *Environ. Geochem. Health.* **40**, 865-885 (2018). doi: [10.1007/s10653-017-0031-6](https://doi.org/10.1007/s10653-017-0031-6)
40. M. S. Blondes, K. D. Gans, M. A. Engle, Y. K. Kharaka, M. E. Reidy, V. Saraswathula, J. J. Thordsen, E. L. Rowan and E. A. Morrissey, "U.S. Geological Survey National Produced Waters Geochemical Database v2.3" (US Geological Survey, 2018). doi: [10.5066/F7J964W8](https://doi.org/10.5066/F7J964W8)
41. D. A. Keiser, J. S. Shapiro, Consequences of the Clean Water Act and the demand for water quality. *Q. J. Econ.* **134**, 349-396 (2019). doi: [10.1093/qje/qjy019](https://doi.org/10.1093/qje/qjy019)
42. Materials and methods are available as supplementary materials at the Science website.
43. J. Mateo-Sagasta, S. Marjani Zadeh, H. E. Turrall, "More People, More Food, Worse Water? A Global Review of Water Pollution from Agriculture," (Food and Agriculture Organization of the United Nations, Water, Water, Land and Ecosystems Program of the CGIAR, International Water Management Institute, 2018; <http://www.fao.org/policy-support/resources/resources-details/en/c/1144303/>).
44. S. L. Brantley *et al.*, Engaging over data on fracking and water quality. *Science* **359**, 395-397 (2018). doi: [10.1126/science.aan6520](https://doi.org/10.1126/science.aan6520)

45. Pennsylvania Department of Environmental Protection (PA DEP), “Spud Data Report”. Retrieved from: <https://www.dep.pa.gov/DataandTools/Reports/Oil%20and%20Gas%20Reports/Pages/default.aspx> (accessed December 2019).
46. Environmental Protection Agency (EPA), “STORET Database”. Retrieved from: <https://www.waterqualitydata.us/> (accessed April 2017).
47. U.S. Geological Survey (USGS), “NWIS Database”. Retrieved from: <https://www.waterqualitydata.us/> (accessed April 2017).
48. S.L. Brantley, “Shale Network Database”, Consortium for Universities for the Advancement of Hydrologic Sciences, Inc. (CUAHSI) (2011). doi: [10.4211/his-data-shalenetwork](https://doi.org/10.4211/his-data-shalenetwork) (accessed December 2019).
49. W. Schlenker, “Daily Weather Data for Contiguous United States (1950-2017)”. Retrieved from: <http://www.columbia.edu/~ws2162/links.html> (accessed December 2019).
50. National Resource Conservation Service (NRCS), “Watershed Boundary Dataset”. Retrieved from: <https://www.nrcs.usda.gov/wps/portal/nrcs/main/national/water/watersheds/dataset/> (accessed December 2016).
51. Science for Nature and People Partnership (SNAPP) “Hydraulic Fracturing - Unconventional Oil and Gas Spills: Risks, Mitigation Priorities and State Reporting Requirements”. Knowledge network for biocomplexity. doi: [10.5063/FITB14VK](https://doi.org/10.5063/FITB14VK) (accessed November 2019).
52. P. Bonetti, C. Leuz, G. Michelon, “Public Repository for Large-Sample Evidence on the Impact of Unconventional Oil and Gas Development on Surface Water Quality” (2021). doi: [10.5281/zenodo.4937324](https://doi.org/10.5281/zenodo.4937324)
53. A. W. Bartik, J. Currie, M. Greenstone, C. R. Knittel, The Local Economic and Welfare Consequences of Hydraulic Fracturing. *Am. Econ. J. App. Econ.* **11**, 105-155 (2019). doi: [10.1257/app.20170487](https://doi.org/10.1257/app.20170487)
54. S. S. Kaushal *et al.*, Freshwater salinization syndrome on a continental scale. *Proc. Natl. Acad. Sci. U.S.A.* **115**, E574-E583 (2018). doi: [10.1073/pnas.1711234115](https://doi.org/10.1073/pnas.1711234115)
55. J. Moore, R. M. Fanelli, A. J. Sekellick, High-Frequency Data Reveal Deicing Salts Drive Elevated Specific Conductance and Chloride along with Pervasive and Frequent Exceedances of the U.S. Environmental Protection Agency Aquatic Life Criteria for Chloride in Urban Streams. *Environ. Sci. Technol.* **54**, 778-789 (2020). doi: [10.1021/acs.est.9b04316](https://doi.org/10.1021/acs.est.9b04316)
56. K. J. Black, S. J. McCoy, J. G. Weber, Fracking and indoor radon: Spurious correlation or cause for concern? *J. Environ. Econ. Manage.* **96**, 255-273 (2019). doi: [10.1016/j.jeem.2019.06.002](https://doi.org/10.1016/j.jeem.2019.06.002)
57. L. Sun, S. Abraham, Estimating dynamic treatment effects in event studies with heterogeneous treatment effects. *J. Econom.* (2020). doi: [j.jeconom.2020.09.006](https://doi.org/10.1016/j.jeconom.2020.09.006)
58. C. de Chaisemartin, X. D'Haultfœuille, Two-Way Fixed Effects Estimators with Heterogeneous Treatment Effects. *Am. Econ. Rev.* **110**, 2964-2996 (2020). doi: [10.1257/aer.20181169](https://doi.org/10.1257/aer.20181169)
59. A. Goodman-Bacon, Difference-in-Differences with Variation in Treatment Timing. *NBER Working Paper Series* **25018**, (2018; <https://www.nber.org/papers/w25018>).
60. A. Kondash, A. Vengosh, Water Footprint of Hydraulic Fracturing. *Environ. Sci. Technol. Lett.* **2**, 276-280 (2015) doi: [10.1021/acs.estlett.5b00211](https://doi.org/10.1021/acs.estlett.5b00211)

61. L. Hanson, S. Habicht, P. Daggupati, R. Srinivasan, P. Faeth, Modeling Changes to Streamflow, Sediment, and Nutrient Loading from Land Use Changes Due to Potential Natural Gas Development. *JAWRA J. Am. Water Resour. Assoc.* 53, 1293-1312 (2017). doi: [10.1111/1752-1688.12588](https://doi.org/10.1111/1752-1688.12588)

Acknowledgments:

We appreciate the helpful comments of Susan Brantley, Thomas Covert, James Evans, Michael Greenstone, Stephanie Grimshaw, Brent Grocholski (editor), Martin Herfort, Ryan Kellogg, John Powers, Joseph Shapiro, Raghavan Srinivasan, and Reika Yokochi. We also gratefully acknowledge excellent research assistance by Igor Kuznetsov, Eloy Lanau, Max Muhn, Fabian Nagel, Christopher Nance, Vicent Peris, Claudia Serra, and Chaoyuan She. **Funding:** Pietro Bonetti gratefully acknowledges research funding from IESE Business School. Christian Leuz gratefully acknowledges research funding from the Booth School of the University of Chicago. **Author contributions:** Authors are listed in alphabetical order. Pietro Bonetti conceived and designed the analysis, collected the data, contributed data or analysis tools, performed the analysis and wrote the paper. Christian Leuz conceived and designed the analysis, contributed data or analysis tools, reviewed results of the analyses and wrote the paper. Giovanna Michelin conceived and designed the analysis, collected data, contributed data or analysis tools, reviewed results of the analyses and wrote the paper; **Competing interests:** The authors declare no competing interests. **Data and materials availability:** Raw data were obtained from WellDatabase, Enverus, the Pennsylvania Department of Environmental Protection and the Pennsylvania Department of Conservation of Natural Resources (45), the Environmental Protection Agency (46), the USGS (40, 47), the Shale Network (48), Wolfram Schlenker's Daily Weather Data for Contiguous States (49), the National Resource Conservation Service (50), and the Science for Nature and People Partnership (51). Data from WellDatabase and Enverus were obtained from these companies by license agreements for non-commercial use. More details about the data are provided in the supplementary materials. The two merged datasets based on which the analyses are performed, along with Stata code, are available at (52).

Supplementary Materials:

Data Description

Research Design

Supplementary Analyses Supporting Statements in the Manuscript

Figures S1-S5

Tables S1-S15

References (53-61)

Fig. 1. Location of sample watersheds (HUC10s) with unconventional O&G development (*shaded in ocher*) and, superimposed, the distribution of HF wells (*red triangles*). Data on the location of wells come from WellDatabase, Enverus, the Pennsylvania Department of Environmental Protection and the Pennsylvania Department of Conservation of Natural Resources. Thin black lines outline HUC10 boundaries; thick black lines depict state boundaries. Maps with a close up of the main U.S. shales and location of water quality monitoring stations are shown in Fig. S2.

Fig. 2: HF Wells and Water Quality. This figure plots OLS coefficients and confidence intervals for the associations between ion concentrations and cumulative HF well counts ($\#wells_{HUC10}$), estimated using eq. S1 and two different model specifications, HUC4 and HUC8 (Table S4). We report results for treated watersheds (HUC10s) in Pennsylvania (PA) and for all treated U.S. watersheds (ALL). The last two columns report the cumulative impact in the average watershed (HUC10 Impact $\mu\text{g/L}$) implied by the coefficient estimates, obtained by multiplying the respective coefficient with the sample mean ion concentration and the cumulative number of wells in the average HUC10 over the sample period. Bold impact numbers are based on significant coefficients. The EPA maximum contaminant level (MCL) is 250,000 $\mu\text{g/L}$ for Cl^- and 2,000 $\mu\text{g/L}$ for Ba. The EPA does not provide a MCL for Br^- and Sr. Health advisory levels for one-day and lifetime exposure to Sr are 25,000 $\mu\text{g/L}$ and 4,000 $\mu\text{g/L}$, respectively.

Fig. 3. Temporal analysis of ion concentrations around well spud dates. **(A)** Br^- , **(B)** Cl^- , **(C)** Ba, **(D)** Sr. Panels A to D plot OLS coefficients for well counts calculated over fixed time intervals around the spud dates, together with the 95% confidence interval (see Table S8, Panel B, Columns (1)-(4), for the estimation of these coefficients). For comparison, the red dot marks the coefficient of #wellsHUC10 from Table S4, Panel A, Column (4), and its 95% confidence interval.

Fig. 4: HF Wells and Water Quality using Time, Distance and Well Position (Upstream). This figure plots WLS coefficients and confidence intervals for the associations between ion concentrations and an indicator for a new HF well, estimated using eq. S2 (Table S12, Panel C). For this analysis, we pair wells and monitors in a watershed. For each pair, we determine that well and monitor are within 15km and that the well is assigned as likely upstream of the monitor, and we only use water measurements taken up to 360 days after the spud date. We report results for treated watersheds (HUC10s) in Pennsylvania (PA) and for all treated U.S. watersheds (ALL). The last column reports the 360-day impact on the average watershed (HUC10 Impact $\mu\text{g/L}$) implied by the coefficient estimates, obtained by multiplying the respective coefficient with the sample mean ion concentration and the average number of new wells per year in the average HUC10. We computed the 360-day impact only for positive coefficients. For this reason, we do not report the mean ion concentration and average number of wells per year for Br^- . Bold impact numbers are based on significant coefficients. See Fig. 2 for EPA maximum contaminant and other health advisory levels.

Supplementary Materials for
Large-Sample Evidence on the Impact of Unconventional Oil and Gas
Development on Surface Waters

Pietro Bonetti, Christian Leuz, Giovanna Michelon
Correspondence to: cleuz@chicagobooth.edu

This PDF file includes:

Data Description
Research Design
Supplementary Analyses Supporting Statements in the Manuscript
Figs. S1 to S5
Tables S1 to S15
References 53-61

Data Description

HF Wells

We obtain data on the location and spud date of HF wells from three databases: (1) WellDatabase, (2) Enverus (formerly Drillinginfo), and, for Pennsylvania, (3) the Pennsylvania Department of Environmental Protection (PA DEP) and the Pennsylvania Department of Conservation of Natural Resources (PA DCNR) (45). WellDatabase and Enverus are widely used data sources in other empirical studies on the O&G industry (53). WellDatabase and Enverus collect oil and natural gas production information from various state agencies for each well. We complement WellDatabase and Enverus information with data from the Pennsylvania Department of Environmental Protection (PA DEP) and the Pennsylvania Department of Conservation of Natural Resources (PA DCNR). The databases provide information on the latitude/longitude of each well, the type of each well (horizontal vs. vertical), the production type of each well (gas, oil), and the spud date of each well. We use only wells with non-missing information on latitude/longitude, spud date, type of well, and type of production. We restrict the sample to O&G production wells. Fig. S1 shows the time trend in HF activity in the U.S. for our sample.

In our main analyses, we combine the three databases (i.e., WellDatabase, Enverus (formerly Drillinginfo), PA DEP/PA DCNR) to obtain a sample of HF wells that is as comprehensive as possible. If a well appears only in one of the three databases, we use the spud date recorded in the respective database. If a well appears in more than one database, but is recorded with different spud dates in the different databases, we give priority to the PA DEP and PA DCNR spud date if the well is covered in these databases, next we use the WellDatabase spud date, and finally assign the Enverus spud date if the well exists only in the latter database.¹

Surface Water Quality Measurement

Surface water quality is measured using the concentrations of four ions: Ba, Br⁻, Cl⁻, Sr. In the literature, these ions (or salts) are considered specific signatures of flowback and produced waters (1, 2, 5, 20, 40) and hence could indicate a contamination related to HF wells, if and when it exists. Produced water flows to the surface during production and consists of brines, i.e., water found in the deep formations holding oil and gas (or adjacent ones) as well as previously injected HF fluids (flowback). Deep formation brines mobilized by HF can contain highly elevated concentrations of all four ions (4, 8, 27, 33, 36, 39). Furthermore, these analytes have been measured reliably and tracked with reasonable frequency over a long time period in publicly available data, allowing us to estimate background ion concentrations.

Data on ion concentrations, downloaded with the HydroDesktop (www.cuahsi.org) software, come from the EPA (STORET data) (46), USGS (NWIS data) (47), the Shale Network (48), which includes data from the Susquehanna River Basin Commission, and from the PA DEP (SAC046). The STORET data and the NWIS data contribute by far the most observations to the sample. Surface-water observations include rivers, lakes, streams, and ponds. The data provide information on the latitude/longitude of each water monitoring station, the ion, the type of surface water (e.g., rivers, lakes), the sampling method, and it identifies the agency in charge of the monitoring station. Although the data include unique monitoring station identification codes, we follow (41) and

¹ We settled on this procedure for assigning spud dates after carefully reviewing the three databases. PA DEP and PA DCNR appear to be the most reliable sources, followed by the WellDatabase and then Enverus. In untabulated robustness tests, we also estimate eq. S1 using only one database at the time, rather than combining them. The results and inferences are very similar and are available upon request.

identify each monitoring station by its latitude and longitude because monitoring sites are often assigned different codes in the different repositories and their assigned names can also change over time.

For some water measurements, numerical values are reported as missing by the databases if the measured ion concentrations are: (i) “Present Below Quantification Limit”, (ii) “Not Detected”, or (iii) “Not Reported”. In the analyses, we treat these cases as follows:

- a) We replace a missing measurement value with the numerical value reported in the “Result Detection Condition Text,” if the text provides a quantification limit value (2), divided by 2 to set the value at the midpoint. There are only very few of these assignments in our sample. For Ba, we replace 7 missing values, for Cl^- we make 4 replacements, and for Br^- and Sr there are none;
- b) We assign a value of zero to any measurement, for which the “Result Detection Condition Text” shows “Not Detected”, as ignoring these measurements could bias in favor of finding an effect. This replacement is particularly relevant for Br^- (see p. 15 below);
- c) We keep the missing value, if the “Result Detection Condition Text” equals “Not Reported” and also if the text equals “Present Below Quantification Limit” but does not provide a numerical value (as otherwise condition a) above applies).

Below, we gauge the robustness of our results to alternative ways of dealing with these measurements choices (see pp. 15-16 for details).²

We also examine the concentrations of three analytes that are not as specific to or indicative of HF-related effects: dissolved oxygen, phosphorus and fecal coliforms. Data on these concentrations come from the EPA (STORET data) (46) and the USGS (NWIS data) (47).

Precipitation and Temperature

Data on daily precipitation and temperature come from the Wolfram Schlenker’s Daily Weather Data for Contiguous States (49). The raw data files give daily minimum and maximum temperature as well as total precipitation on a 2.5x2.5 mile grid for the contiguous United States from 1900-2019. The data are based on the PRISM weather dataset. We compute the average temperature on the day of water measurement and the cumulative precipitation over the last three days including the day of water measurement for the 2.5x2.5 mile grid, in which each monitoring station is located. The use of Schlenker’s Daily Weather Data allows us to measure the local weather conditions at the time and location of water measurement with greater precision than we could with other databases (e.g., National Oceanic and Atmospheric Administration National Climatic Data).

Watershed Boundaries and Elevation model

Data on the watershed boundaries come in *shapefile* formats from the Watershed Boundary Dataset (WBD) provided by the Natural Resources Conservation Service (NRCS) (50) at the Geospatial Data Gateway (GDG). We use various hydrological unit levels, which are defined in terms of their hydrologic unit codes (HUC): HUC4 (sub-region), HUC8 (sub-basin), HUC10

² The USGS database reports a number of Ba concentrations at exactly 100 ppb. These observations are more likely to reflect measurements at the detection limit rather than “actual” measurements. We therefore check whether our results hold after treating these observations as reported at the detection limit, i.e., we set them to the midpoint of 50. The inferences remain unchanged. We thank an anonymous reviewer for this suggestion.

(watershed), HUC12 (sub-watershed). We determine the altitude of each well and each water monitoring station as well as assign wells as upstream or downstream of monitors using the HYDRO1k Digital Elevation Model provided by the USGS. We provide more details on this assignment below.

Watershed Regional Hydrogeology

Data on the regional hydrogeology, which we use to classify watersheds into those with high- and low-salinity deep formations, come from the USGS Produced Waters Geochemical Database (40). The dataset provides information on total dissolved solids (TDS) concentrations in produced waters during oil and gas extraction. We use these TDS concentrations as a proxy for the salinity of the deep formations. As a sensitivity analysis, we also use Cl^- concentrations instead of TDS (and obtain very similar results).

Sample

For the primary analysis, we only use observations from monitoring stations located in HUC10s with at least one active HF well over the sample period, focusing essentially on watersheds for which unconventional O&G development is relevant. We further require complete information on the latitude/longitude of each monitoring station, the sample date, the unit of measurement, the type of surface water the station monitors (e.g., rivers, lakes), the ion sampled, and the amount of the ion measured.³ Furthermore, we require at least two water measurements for each ion in each HUC8-month-year given the fixed effects in the model (described below). These sample requirements yield 60,783 surface water measurements over ten years (January 2006 to June 2016) from 2,038 unique monitoring stations and from 408 HUC10s with 46,479 HF wells, across 17 U.S. states, and 24 shales. Of the 408 watersheds, 105 are located in Pennsylvania and 303 are located outside Pennsylvania. This sample is comprised of 6,642 observations for Br^- , 27,628 for Cl^- , 13,578 for Ba and 12,935 for Sr.

Table S1 presents descriptive statistics for the four ion concentrations, both in level and logarithmic form. All the ions are reported in microgram per liter ($\mu\text{g/L}$). To limit the influence of outliers due to measurement or recording errors, we set extreme values above the 99th percentile to missing. We determine the 99th percentile per ion and HUC4 level to account for regional differences in ion concentrations. Most of our surface water observations come from rivers and streams: 97.80% for Br^- , 95.40% for Cl^- , 94.64% for Ba and 97.46% for Sr.

The distribution of ion observations across regions and shales is reported in Table S2 and their distribution across HUC10s is reported in Table S3. As shown in Table S2 and Panel A of Table S3, water measurements can be quite sparse. The most frequently measured ion is Cl^- , which is available for most shales, and the Marcellus shale has the most observations, for which there are data on all four ions. On average, there are 12 monitoring stations per HUC10 with active HF wells, ranging from an average of 6 for Br^- to 14 for Sr. The average number of measurements per ion in a HUC10 ranges from 46 for Br^- to 72 for Sr. Fig. S2 shows three maps of HUC10s with HF wells and the location of water monitoring stations in these treated HUC10s.

Research Design

A positive association between unconventional O&G development and ion concentrations in the same watershed can arise for many natural and anthropogenic reasons, even for ions considered

³ For comparability with prior research, we exclude measurements related to bed sediments or labeled as “suspended”.

signatures of HF flowback and produced waters (2, 33, 39). Thus, this study uses regression analysis with temperature and precipitation controls and an extensive set of “fixed effects”⁴ to construct local and time-varying baselines for background ion concentrations against which anomalous ion concentrations associated with HF wells in the same watershed are estimated (41). The fixed effects in the model create two baselines that essentially perform difference-in-differences comparisons.

The first baseline is monthly and, alternatively, at the sub-region (HUC4) or at the sub-basin (HUC8) level. This very flexible baseline controls for arbitrary regional differences in ion concentrations, but more importantly, for arbitrary *changes* in sub-region or sub-basin ion concentrations from month to month or year to year. This baseline is obtained by including either HUC4-month-year or HUC8-month-year fixed effects in the model. These fixed effects compute the average concentration for each HUC4 (HUC8) in each month and year, and then subtract the group-specific mean from each observation that falls into the specific HUC4 (HUC8) in the respective month. It controls, among other things, for seasonal changes, including the effects of road de-icing, agriculture, and general trends in salinity (38, 54-55), but also for the effects of economic development associated with the rise of unconventional O&G development.⁵

The second baseline is even more local and accounts for the average ion concentration at the monitoring station, using each monitoring station as its own control. This baseline is obtained using monitoring-station fixed effects, which compute the average ion concentration for each monitoring station and subtracts this group-specific mean from each observation of the respective monitor. It accounts for differences in concentrations levels across monitors and, among other things, for differences in the way water quality is measured, the type of monitor, the type of water body, the location of the monitor as well as location-specific concentration levels.⁶

We also control for local weather at the monitoring station at the time of water measurement by separately including temperature and precipitation readings from 2.5x2.5 mile grid in which each monitoring station is located. These local weather controls further mitigate concerns about the seasonal effects as well as the impact of snow storms and road de-icing.

The statistical association between HF and water quality is estimated exploiting (i) cross-sectional variation in the location of HF wells across watersheds (HUC10) within a sub-region (HUC4) or sub-basin (HUC8) and (ii) time-series variation in HF wells within a given watershed.

⁴ In a fixed effects model, each group mean is a group-specific fixed quantity. The fixed effects model computes the mean for a group (e.g., a monitor) and subtracts this mean from each observation when estimating the model. The advantage of this approach is that it controls for a confounding effect, even when it is unobservable and hence could not be directly measured. A fixed effect eliminates this influence by demeaning the variables using the “within transformation.”

⁵ For example, if a winter storm in an area leads to the use of road salts for de-icing and these salts wash into the area’s surface waters, salt concentrations should increase for many monitors in the area, regardless of whether there are newly spudded wells in the watershed. By subtracting a HUC8-specific monthly average (which can be different each year), we account for time-varying, seasonal effects in a flexible manner and hence mitigate the concern that winter storms or other unobserved events or trends (e.g., economic development, agriculture) confound our analysis and are mistaken as HF effects (especially as drilling has seasonal variation too).

⁶ For example, some monitors could be close to locations with natural brine migration or seepage and hence could show very high salt concentrations. If HF wells are close to such monitors, it is easy to confuse HF impact with natural brine seepage. By subtracting the fixed effect for each monitor, the model essentially uses the salt concentration at each monitor as its own control and estimates the effect “within” each monitor, which amounts to a before-after comparison. The regression then shows an increase related to new HF wells only if concentrations at a monitor increase over and above their normal (in this case naturally high) levels after a well is spudded.

Thus, the regression analysis tests whether variation in HF wells across watersheds or over time in a watershed is associated with elevated ion concentrations after controlling for: (i) arbitrary cross-sectional and time-series variation in background ion concentrations at the sub-region or sub-basin level, respectively, (ii) arbitrary variation in the average concentrations across water monitoring stations, and (iii) local weather at the time of water measurement. Specifically, we estimate the following regression model using ordinary least squares (OLS):

$$C_{ikd} = station_i + HUC_{kmy} + \alpha p_{ikd} + t_{ikd} + \beta \#wellsHUC10_{kd} + \varepsilon_{ikd} \quad [\text{Eq. S1}]$$

where C_{ikd} is the natural logarithm of ion concentration, measured at monitoring station i in HUC10 k on day d ;⁷ $station_i$ is the monitoring station fixed effect, which controls for the average ion concentration at the station; HUC_{kmy} is (alternatively) either a HUC4- or HUC8-month-year fixed effect, which controls flexibly for arbitrary changes in the average ion concentration in a HUC4 (HUC8) in a given month and year; p_{ikd} is the logarithm of the three-day cumulative precipitation registered on the day a water quality observation is drawn,⁸ t_{ikd} is the average temperature on the day a water quality measurement is drawn, in Celsius, controlled for in a non-parametric way (using five binary indicators for non-overlapping temperature ranges: [< -10], [$-10; 3$], [$3; 15$], [$15; 25$], [> 25]), and ε_{ikd} is the error term. By using five separate binary temperature indicators, this specification allows for a non-monotonic and non-linear relation between the respective ion concentration and temperature (e.g., higher concentrations when temperatures are close to freezing because of road salts and again higher concentrations on hot days because of evaporation). Standard errors are clustered at the HUC8 level to allow arbitrary serial- and cross-sectional correlations within a sub-basin.

The variable of interest, $\#wellsHUC10_{kd}$, measures the cumulative number of HF wells in a given HUC10 at a given point in time. To generate this variable, we first assign wells to HUC10s using QGIS geo software and then compute the cumulative number of HF wells spudded within the HUC10 k up to day d based on the spud date of each well.⁹ In other words, $\#wellsHUC10_{kd}$ increases by one when a new well is spudded on day d in watershed k . If new HF wells are associated with increases in the concentration of one of the four ions, we expect a positive coefficient for $\#wellsHUC10_{kd}$ in the regression for the respective ion. This approach (presented in Table S4) estimates a statistical link between newly spudded HF wells and ion concentrations in watersheds with exposure to HF wells. The coefficient estimates the average “impact” of an additional HF well, but a significant coefficient does not require or imply that each and every well has an effect on ion concentrations.

⁷ We take the natural logarithm of the ion concentration, C_{ikd} , measured at monitoring station i in HUC10 k on day d , to account for the positive skewness of the ion concentration distribution. There is no consensus in the literature on how to model ion concentrations in regressions. Concentrations have been modelled in levels (38, 41) or in logs (18, 41). The results are not materially different using levels as long as the positive skewness of ion concentrations is addressed by using some form of outlier treatment. See Table S5, Panel F.

⁸ The results are virtually unaffected if we instead model precipitation as the logarithm of the 2-day cumulative precipitation or as the logarithm of the measurement-day precipitation.

⁹ We use spud dates to determine the cumulative number of HF wells rather than well completion dates because the latter information is not in the PA DEP and PA DCNR database (45) and often missing in the WellDatabase and Enverus databases. There are on average 103 days between spudding and well completion in our sample (which is similar to the 80-day average reported in (38)). Using the spud date, rather than the well completion date, should not have a major influence in our long-run analysis and, if anything, biases the coefficient downwards.

Considering that Pennsylvania accounts for almost 41% of the sample (and given that many prior studies have focused on this state), we estimate our regression model separately for watersheds in Pennsylvania and for all U.S. watersheds with HF wells. In the manuscript, we report analyses only for (treated) watersheds with at least one active HF well over the sample period. As sensitivity analysis, we also conduct analyses (Table S5, Panel C) using in addition water measurements from watersheds without HF wells, as long as these (untreated) watersheds are located in a sub-region (HUC4) with some HF activity. The additional observations from untreated watersheds contribute only to the estimation of the monthly sub-region or sub-basin baselines.¹⁰

Analysis by Watershed “Type”

It is important to recognize that identifying an impact of unconventional O&G development in water chemistry data, even when it exists, is very challenging for multiple reasons. For example, an accident could be small. Moreover, spills could be diluted, biodegraded or obscured by other contaminants (16). Another important challenge is the availability of relevant water quality data. HF wells could be quite far from the closest water monitoring station. Moreover, the frequency of water measurement at many monitors is low and hence there may be few readings shortly before or after a new HF well is spudded. Similarly, (20) point out that sparsity of *spatial* and *temporal* water monitoring is a major impediment to spill detection. For these reasons, an analysis that uses water measurements from all monitoring stations in a watershed (as we do in Table S4) has low power. It also understates the effect if some of the monitors are upstream to or too far from a well and hence cannot detect an effect, even when it exists. A watershed-level analysis is still a useful first step because it does not presume a particular pathway and essentially just estimates if there is any detectable “impact” in watersheds with exposure to HF wells.

Given these challenges, we also conduct an analysis that recognizes that watersheds differ substantially along dimensions that would make easier or harder to detect a potential effect (e.g., frequency of water measurement, distances between HF wells and water monitors). For this analysis (presented in Table S11), we classify watersheds into those that have a higher chance of showing the relation between new HF wells and increased ion concentrations, if it exists, and those that have a lower chance. We define watersheds as “high type” when there are (relatively) (i) many water measurements around the time a new well is spudded; (ii) many monitors in reasonably close proximity to the new wells; (iii) the watershed has a higher fraction of monitors that are likely downstream to a well.

To make these determinations, we first compute by ion and for each new well in a HUC10, the fraction of water measurements that satisfy the following three conditions (relative to total number of water measurements in the HUC10): (i) a measurement takes places *within 360 days* before or after the spud date; (ii) the *spatial* distance between the well and the monitoring station of a measurement is within 15 km; (iii) a measurement is from a monitor that is likely *downstream* from the well. Then, we compute, for each ion and HUC10, the average for this fraction over all wells. In other words, the resulting ion-HUC10 averages are the *average percentage* of ion

¹⁰ When we estimate eq. S1, the sample size falls slightly from column 1 to column 2, and from column 3 to column 4 because in columns 2 and 4 employ a finer fixed effect structure (HUC8-month-year) relative to the fixed effect structure employed in columns 1 and 3 (HUC4-month-year). A finer fixed effect structure is more stringent and results in more singletons, which are then dropped from the sample to make sure that we are not overstating statistical significance.

measurements in a HUC10 that satisfy the three conditions. The higher this fraction is for a given HUC10, the higher is the chance that water measurements in this HUC10 could detect effects around new HF wells, if and when they exist.

The distance between a well and a monitoring station (in km) is computed using Vincenty's formula for calculating the distance between two points on a sphere. To assign wells as (likely) up- or downstream from a monitor within the respective watershed, we sub-divide each watershed into cells of constant size (1 km²) and first identify flow direction and flow accumulation by computing flow direction codes (1 to 255) and flow accumulation values for each cell. We then apply the flow-length routine in ArcGIS to these square-km cells to assign upstream/downstream monitors.¹¹ We make flow length comparisons only if monitor and well are in the same sub-watershed (same HUC12) to further reduce noise in the assignment of monitors.

Next, we classify watersheds into two "types" based on the ion-HUC10 averages, splitting by the sample median of these averages. For example, a value of 0.22 for Cl⁻ in a certain HUC10 means that, for wells in this HUC10, on average 22% of the Cl⁻ measurements are within 360 days from the spud date, within 15km from the wells, and likely downstream to the wells. If the median over all watersheds is 20 percent, then a watershed with 22 (12) percent is classified as a high (low) type. The partitioning based on the sample median guarantees that the *high*- and *low*-type groups have roughly the same number of watersheds, and so inferences cannot be driven by an uneven number of HUC10s in the two groups. We then estimate eq. S1 but replace *#wellsHUC10* with two non-overlapping cumulative well count variables, counting the cumulative number of wells spudded in high (*#wellsHUC10_high_type*) and low (*#wellsHUC10_low_type*) type watersheds, respectively.

Analysis Pairing Wells and Monitors

To exploit time, distance, and flow direction in the analysis *at the monitor-well level*, we use an alternative research design that structures the data in a different way. For this analysis, each well is paired with all water monitoring stations in the same HUC10. For each well-monitor pair, we determine the distance between the well and the monitoring station (in km, as describe above) as well as whether the well likely sits up- or downstream from the monitoring station (as describe above). We then estimate the following model:

$$C_{ijkd} = pair_{ij} + HUC_{kmy} + \alpha p_{ijkd} + t_{ijkd} + \beta Post\ Spud_{ijkd} + \varepsilon_{ijkd} \quad [\text{Eq. S2}]$$

where C_{ijkd} is the natural logarithm of ion concentration, measured at monitoring station i paired with well j in HUC10 k on day d ; $pair_{ij}$ is the monitoring station-well pair fixed effect,

¹¹ Flow length is the distance travelled from any cell along the surface flow network to an outlet. This approach essentially computes the flow length from any cell in a watershed to the closest headwater or stream outlet of that watershed. Thus, if the flow length of the monitor cell is smaller than the flow length of the well cell, the monitor is assigned as "downstream" to the well. While this algorithm uses information about water direction and goes beyond using elevation data, it is clearly an approximation only. For example, it could look like a well is upstream to a monitor based on this routine, but it sits in a different catchment separated by a ridge. For this reason, we make up- and downstream assignments only within the same sub-watershed. While this restriction is an improvement, it is not perfect. Thus, we view our assignment of monitors more from a likelihood perspective: monitors with a shorter flow length are more likely to show effects if there is impact by a well with a higher flow length. Note, however, that this admittedly imperfect assignment of monitors should not bias in favor of finding results. If anything, including monitors in the analysis that actually are upstream from the wells should bias against finding significant coefficients.

which controls for the average ion concentration of the well-monitor pair; HUC_{kmy} is the HUC8-month-year fixed effect; p_{ijkd} is the cumulative three-day precipitation registered the day a water quality observation is drawn, t_{ijkd} is the average temperature on the day a water quality measurement is drawn, in Celsius, controlled for in a non-parametric way (using five binary indicators for non-overlapping temperature ranges: $[-10]$, $[-10; 3]$, $[3; 15]$, $[15; 25]$, $[> 25]$), and ε_{ijkd} is the error term. Thus, the model expressed in eq. S2 has the same fixed effects structure as eq. S1, except that it includes well-monitor pair fixed effects, which means that we benchmark any potential HF effect against the average ion concentration computed over the relevant water measurements for the respective well-monitor pair (see water sample restrictions below). Standard errors are clustered at the HUC8 level.

$Post\ Spud_{ijkd}$ is the variable of interest. It is a binary indicator that is equal to one when a water measurement on day d at monitor i from HUC10 k is drawn after the well j spud date, and zero otherwise. This variable indicates an association between new HF wells and the ion concentration within a watershed (HUC10). As such, it estimates the average “impact” of a HF well on the respective ion concentration.

To focus on the most relevant water measurements in terms of time, distance and flow direction, we restrict the sample to: (i) water measurements that are either taken before or up to 360 days after the spud date;¹² (ii) monitors that are within 15km from a well; (iii) monitors that are assigned as downstream to a well. Because our model uses each well-monitor pair as its own control, we require at least two measurements before and after well spudding per pair and ion concentration. Table S12, Panel A and B report surface water ion concentrations for all U.S. (Pennsylvania) watersheds for monitors that satisfy (i)-(iii) above.

Due to the sample restrictions combined with the model’s demanding fixed-effects structure and the conservative clustering of standard errors at the HUC8 level, statistical power is an issue in this analysis. We therefore estimate eq. S2 with weighted-least squares (WLS) regressions, using the number of HUC8-month-year observations as weights. This approach gives more weight to observations when there are more data to compute the monthly sub-basin baselines for the respective ion concentrations. For comparison, we also estimate eq. S2 with OLS. In addition, we use randomization to compute bootstrapped p-values for the Table S12, Panel C, coefficients, rather than assuming a particular clustering structure.

Supplementary Analyses Supporting Statements in the Manuscript

Distance Gradient Analysis— Cumulative Well Counts by Distance

We conduct a distance analysis in the spirit of (18) and (56) to examine whether the positive association between HF wells and ion concentrations is stronger for wells that are closer to monitors. However, the sparsity of water measurement presents a challenge for this analysis. The availability of water measurements in close proximity of a well depends on the distance between the well and the closest water body, say a river, and also on the existence of a monitoring station for that river. We therefore use *cumulative* distance bins with increasing distance to estimate a

¹² We use all available water measurements in the pre-spud date period, as doing so improves the estimation of the baseline. If we instead strictly limit the sample to observations that are taken within 360 days before and 360 days after the well spudding, the estimated magnitudes are even larger. In our assessment, the latter estimates are statistically less reliable, as the number of pre-period water measurements used to establish the baseline is much smaller. Thus, we prefer to report the results that are smaller in magnitude but in our view more reliable.

distance gradient, rather separate bins for certain non-overlapping distances. Specifically, we estimate eq. S1 and include two separate well counts: $\#wellsHUC10 [0km; Xkm]$, counting the number of wells spudded within X km of monitors within a given HUC10, with $X = 5, 10, 15, 20, 25, 30$; and $\#wellsHUC10 [30km; Max]$, counting the number of wells spudded that are at least 30 km from monitors within a given HUC10. For each value of X , we estimate a separate regression with both counts, i.e., nearby wells and wells beyond 30 km. We then plot the coefficient for the nearby-well count, $\#wellsHUC10 [0km; Xkm]$, for different values of X to show a distance gradient, i.e., how the association changes as X increases and hence the nearby-well count includes wells that are further and further away. Given the sparsity of water measurements, we *stack* observations for Ba, Br⁻, Cl⁻ and Sr in one model. That is, we estimate one model for all ions simultaneously. To account for the fact that the model uses observations from different ions, we include interactions for the ion indicator and the control variables and fixed effects, so that these coefficients are specific to each ion.¹³ We also estimate a model for Cl⁻, which is the ion with the most data. Fig. S5 plots the estimated coefficients and shows that the coefficient is largest when wells are within 5km from the monitoring station. Coefficients steadily decline as the counts include wells that are further and further away. However, the widths of the confidence intervals for different values of X also illustrates how data sparsity affects power. Using only wells that are very close, the model has much less power. When we expand distance bin width, i.e., permit more and more wells into the estimation, the confidence intervals become much tighter. The coefficient also declines (as one would expect). Overall, Fig. S5 shows that the association between HF wells and ion concentrations is stronger when wells are closer to water monitors.

Pennsylvania vs. Other U.S. States

To provide separate results for PA and non-PA watersheds, we re-estimate eq. S1 for the U.S. dataset replacing $\#wellsHUC10$ with two non-overlapping cumulative well count variables, one for Pennsylvania and one for all the other U.S. states. The results are shown in Table S5, Panel A. In many specifications, and except for Ba, the coefficients for PA and non-PA are remarkably similar in magnitude, suggesting that the results are not limited to Pennsylvania.¹⁴

Estimates for Different Time Periods

We investigate whether our results are concentrated in a particular time period, for instance, in the early years of the sample period when HF regulations have been less strict than in later years.¹⁵ We estimate eq. S1 for different time periods. Specifically, we replace $\#wellsHUC10$ with three non-overlapping $\#wellsHUC10$ variables, one for the period between 2006 and 2010, one for the period between 2011 and 2013, and one for the period between 2014 and 2016. Results are reported in Table S5, Panel B. The results are not limited to a particular period and fairly consistent through time. The significance levels of the results are higher, and the standard errors are smaller in the later periods (except for Ba and Br⁻), presumably because the frequency of water measurements increases over time. The coefficient magnitudes are quite large in the first time period (even though they are insignificant) compared to the next two time periods. This decline could reflect that unconventional O&G extraction practices have improved over time.

¹³ This model is akin to running a SUR (seemingly unrelated regression) model. The model produces an average estimate over all chemicals (similar to (18)), but by pooling all chemicals we harness more power.

¹⁴ Some watersheds cross state boundaries, which raises the issue of assigning them to PA or non-PA. We assign all watersheds that cross the Pennsylvania state boundaries to Pennsylvania.

¹⁵ For example, before 2011, discharging brines directly into streams was not prohibited in Pennsylvania.

Using Observations from Untreated HUC10s within Sub-Regions with HF Wells

The main analysis is restricted to observations from HUC10s with HF wells (treated). A sub-region with HF wells can have also untreated HUC10s. As a sensitivity check, we also use water measurements from these untreated HUC10s to estimate background ion concentrations. The sample now includes water quality observations from all HUC10s that are within HUC4s with at least one HF well. The additional water measurements contribute only to the estimation of the monthly sub-region or sub-basin baselines, as they are from untreated watersheds without HF wells. The results, reported in Table S5, Panel C, corroborate the Table S4 results and inferences. However, the coefficients are slightly smaller and less significant, possibly because the use of untreated HUC10s results in less relevant baselines for background ion concentrations. Sub-regions can be large and heterogeneous in their hydrogeology as well as include areas that are not suitable for unconventional O&G development.

Controlling for Snow and Road Salts

Winter storms and the resulting application of road salts for de-icing could be a confound when analyzing salts in surface waters. To control for such influences, the model includes temperature and precipitation controls as well as fixed effects that account for the monthly variation in any given HUC8 and across different HUC8s, related to any observed or unobserved factors. If a winter storm over a HUC8 leads to the discharge of road salts for de-icing and these salts wash into the HUC8's surface waters, salt concentrations should increase at many monitors within the sub-basin. The average of these increases is taken out by the HUC8-month-year fixed effect, which should control for the effect of road salts on surface waters. Nevertheless, we run a robustness test including an additional control variable for snow in eq. S1. We create a dummy variable that equals to one if precipitation and temperature measurements at the local monitoring station indicate that it could have snowed in the last three days leading up to the water measurement (using a conservative definition for snow: cumulative precipitation above 0.5mm and average temperatures below 4 degrees Celsius). Results are reported in Table S5, Panel D. The coefficients for the snow indicator are generally positive and, in some cases, quite sizeable, which is consistent with the notion that snow and de-icing can increase salt concentrations in surface waters. However, the snow indicator is significant only in the HUC4 specification for Cl^- in PA. Importantly, the estimated coefficients for *#wellsHUC10* remain essentially unchanged when we include the snow indicator as an additional control, suggesting that the documented associations between HF wells and ion concentrations are not driven by road salts and also that the model already sufficiently controls for snow and road salts (especially when including the finer HUC8-month-year fixed effects).

Controlling for Within-HUC10 Seasonality and Excluding the Winter Months

We control for within-HUC10 seasonality in background ion concentrations by adding HUC10-quarter fixed effects to eq. S1. This modification controls flexibly for arbitrary weather or seasonal fluctuations at the watershed level. The results, reported in Table S5, Panel E, Columns (1) and (2), corroborate the results in Table S4. As a further (untabulated) robustness test, we estimate eq. S1 excluding observations during the winter season (Dec, Jan, and Feb). The HF associations are very similar to those in Table S4, showing that the results do not stem solely from the winter months when de-icing occurs.

WLS Estimates for Eq. 1

In some sub-regions and watersheds water quality data are sparse. To account for this data sparsity in some locations, we also estimate the model with WLS using the number of water quality observations within a given HUC8-month-year as weights. This approach gives more weight to sub-regions and months for which there are more data and therefore our model provides more accurate surface water baselines. The results, reported in Table S5, Panel E, Columns (3) and (4), corroborate the results in Table S4 and inferences in the manuscript, except for *Sr* in Column (4).

OLS Estimates for Eq. S2

The well-monitor pair analysis restricts the sample to more relevant water measurements, which can result in data sparsity for some locations. For this reason, we estimate eq. S2 with weighted-least squares (WLS) regressions, using the number of HUC8-month-year observations as weights. This approach gives more weight to observations when there are more data to compute the monthly sub-basin baselines for the respective ion concentrations. For comparison, Table S12, Panel D, reports the results estimating eq. S2 using OLS. These estimates are smaller and, as expected, have lower significance levels. Nevertheless, the coefficients exhibit similar patterns as those reported in Table S12, Panel C. Moreover, the OLS estimates in Table S12, Panel D, are still substantially larger in magnitude than those in Table S4.

Results for Eq. S2 with “Never-Treated Control” Monitors

A recent literature in econometrics (57-59) highlights that difference-in-differences (DiD) analyses with two-way fixed effects (one for time and one for group) can produce biased estimates in the presence of heterogeneous treatment effects. With staggered treatments, the problem arises because DiD estimates based on two-way fixed effects are essentially weighted averages of many comparisons, including those that use post-treatment observations from earlier treatments as controls for later-treated observations, and vice versa.

Our analysis of eq. S2 uses DiD with staggered well treatments and two-way fixed effects.¹⁶ Moreover, we expect our setting to exhibit heterogeneous effects (e.g., surface water impacts likely differ across wells when they exist). To gauge this issue and circumvent the comparison problem, we conduct an analysis that uses observations from “never-treated” water monitors as controls. Such monitors are far and upstream from HF wells in the watershed and hence unlikely to be ever treated. Specifically, we identify monitors that are more than 15km from any well and assigned as upstream in all well-monitor pairings. We pair these “control” monitors with all HF wells in the same watershed and add their water measurements to the sample of water measurements that come from monitors that are potentially treated by HF wells (i.e., satisfy the sample criteria for the analysis presented in Table S12, Panel C). We then augment eq. S2 with a binary variable, *Treatment_Sample*, that indicates for each water measurement whether it comes from a potentially treated monitor (= 1) or a control monitor (= 0). The coefficient of interest is the interaction of this binary indicator and the post-spud indicator. Table S12, Panel E, reports the results from this analysis using WLS estimation. The results are quite comparable to those in Table S12, Panel C.

¹⁶ As our analysis of eq. S1 also uses two-way fixed effects, we gauge the influence of negative weights on our estimates (given the coefficients on *#wellsHUC10* are positive), as suggested in (58). In untabulated analyses, we find only a small proportion of HUC8-month-year cells with negative weights and the sum of these negative weights is close to zero, all of which is reassuring.

Bootstrapping p-Values and Statistical Inferences for Eq. S2 Estimates

In all analyses, we cluster standard errors at the HUC8 level, which allows for arbitrary correlations in water measurements within an entire sub-basin. Clustering at such a high level could be conservative (as noted above OLS estimates in Table S12, Panel D have low significance levels). To assess the extent to which our statistical inferences in Table S12, Panel C are conservative, we bootstrap the p-values using the actual distribution of water measurements. Towards this end, we randomly assign: (i) spud dates between January 1, 2006 to December 31, 2016; (ii) distances between wells and monitoring stations, randomizing the distances between 16 and 30,000 meters so that in expectation distances are evenly distributed around 15,000 meters (15,000 meters is roughly the median distance between wells and monitors in our sample); (iii) upstream/downstream assignments of wells to monitoring stations, randomizing the differences in the flow direction scores between wells and monitors over the range $[-3,000,000; 7,000,000]$ so that in expectation 30 (70) percent of the observations are from monitors that are assigned as likely upstream (downstream) from wells, which roughly corresponds to the proportions in our sample. We repeat the joint random assignment of spud dates, distances, and flow directions 1,000 times and, in each iteration, estimate eq. S2 and then store the coefficient on the variable of interest. Given the randomization, the average coefficient is expected to be small and close to zero. More importantly, the resulting distribution of coefficient estimates, generated by the 1,000 draws, allows us to compute p-values for estimated coefficients in Table S12, Panel C, essentially counting how many of the randomized coefficients exceed the actual coefficient. The results of this exercise are reported in Panel F of Table S12. Using bootstrapped p-values, the estimated coefficients tabulated in Table S12, Panel C for Ba, Cl^- and Sr are all highly significant (at 1% level or better), indicating that our clustering at the HUC8 level is likely conservative.

Controlling for HUC10 Size

Although HUC10s are generally fairly similar in size, there are some differences in watershed size and well density. Therefore, we scale the main variable of interest, *#wellsHUC10*, by the size of the HUC10s (in km^2) and re-estimate eq. S1. This approach follows (38) and accounts for HUC10 size and heterogeneity in the density of fracture wells across watersheds. The results, reported in Table S5, Panel E, Columns (5) and (6), corroborate the results in Table S4 and the inferences in the manuscript. Due to scaling by HUC10 size, the regression coefficients reported in this analysis denote the percentage increase in the respective ion concentration if the HF well density in a watershed increases by 1, which is a very large increase. Hence, the coefficient should be interpreted relative to the average well density.¹⁷

Analyses Using Alternative Transformations of the Ion Concentrations

As a robustness check, we model ion concentrations using other transformations than the natural logarithm. In our main analysis, we use the natural logarithm of the respective ion

¹⁷ There are on average 0.083 wells per km^2 , highlight that an increase in HF well density by 1 would be very large. Therefore, it is better to gauge the magnitude of the estimated effects by computing the percentage increases in ion concentrations if the HF well density in a watershed moves from the 25th percentile to the 75th percentile. Using the interquartile range, the respective percentage increases for the significant coefficients in Table S5, Panel E, Column 5, are as follows: Ba 1.37% (0.542 $\mu\text{g/L}$), Cl^- 3.06% (605.329 $\mu\text{g/L}$), Sr 1.74% (2.112 $\mu\text{g/L}$), which are comparable to Table S4.

concentration to account for the substantial skewness in the observed ion concentrations.¹⁸ However, there are other ways to account for positive skewness. We use three alternative specifications. In the first specification, we estimate the relation between the untransformed ion concentrations and HF wells using robust regressions (using weights from the “rreg” procedure in STATA to reduce the influence of influential observations). In the second specification, we convert ion concentrations to decile ranks to “normalize” values. We assign ranks within HUC4 to account for regional differences. For each HUC4 and ion, we assign water quality observations to ten equally populated groups, from highest to lowest and then assign use the rank of the group, rather than the concentration level. In the third specification, we transform ion concentrations into a binary variable, which is equal to one if a water quality observation is above the 75th percentile of each ion’s concentration distribution of the respective HUC4 and year. The results, reported in Table S5 Panel F, are generally consistent with those reported in Table S4 using the natural logarithm, suggesting that the results are robust to alternative ways of dealing with positive skewness. We note that some of the coefficients of interest are not significant, e.g., when we use ion concentration in decile ranks or in binary form. However, those two approaches also eliminate (meaningful) variation in ion concentrations, which likely reduces power. The robust regressions (without logging) produce significant results that are very similar to those reported in Table S4.

Robustness Tests Related to Variable Measurement

We conduct a number of analyses to check whether the results and inferences in the manuscript (reported in Table S4 and Table S12) are robust to alternative variable measurement choices. These robustness tests below are not tabulated but available upon request.

First, we test whether the results and inferences are robust to alternative truncation choices for the ion concentrations, which we apply to limit the influence of outliers or recording errors. In all tabulated analyses, we truncate the water quality observations at the 99th percentile, computed per ion at the HUC4 level to allow for regional variation in ion concentrations. In additional (untabulated) analyses, we consider alternative truncation choices. Specifically, we (i) truncate ion concentrations at the 1st and 99th percentile per ion by HUC4; (ii) truncate ion concentrations at the 5th and 95th percentile per ion by HUC4; (iii) truncate ion concentrations at the 95th percentile per ion by HUC4; (iv) truncate the residuals after first applying the fixed effects, which accounts for even more regional variation before truncation. That is, we first estimate eq. S1 or eq. S2 without any truncation of the ion concentrations. We store the residuals from the estimation and then truncate the residuals at the 1st and 99th percentiles. The results show that our inferences are not affected by a particular truncation strategy for the ion concentrations.

Second, we check whether the results for eq. S1 and S2 and the corresponding inferences are robust to alternative modelling choices for the weather controls. We consider non-linear relations. Specifically, we add additional controls to eq. S1 or S2: precipitation and temperature in quadratic form [precipitation $\times$ precipitation, temperature $\times$ temperature], and precipitation and temperature in cubic form [precipitation³, temperature³]. We further model temperature as a continuous variable, rather than in categorical form. The results are robust to these variations and very similar to those reported in the manuscript.

¹⁸ Given the use of the natural logarithm, we add one to all ion concentrations, so that zero concentrations are still included and start at zero after applying the logarithm. We checked that adding one to the ion concentration does not materially affect our analysis.

Third, we restrict the sample used to estimate eq. S1 to HUC4s with at least 50 water quality observations for each analyte, in order to check if sub-regions with sparse data unduly influence the results. They do not. We also restrict the sample to water quality observations drawn from routine monitoring activity only. Again, the results are very similar and the inferences the same.

Fourth, the water quality data at times come with text flags indicating measurement issues (e.g., that an ion was present but below the quantification limit). In these cases, the numerical value is typically missing. There are very few observations in our sample with such measurement issues, except for Br^- . Our starting sample (before making the assignments described below) has the following number of observations with measurement text flags:

- a. Ba has 203 observations marked as “Present Below Quantification Limit” and 864 marked as “Not Detected.” In total, these are 4.9% of all Ba observations.
- b. Br^- has 2,208 observations marked as “Present Below Quantification Limit” and 4,464 marked as “Not Detected.” In total, these are 62% of all Br^- observations.
- c. Cl^- has 266 observations marked as “Present Below Quantification Limit” and 616 marked as “Not Detected.” In total, these are 2.5% of all Cl^- observations.
- d. Sr has 74 observations marked as “Present Below Quantification Limit” and 10 marked as “Not Detected.” In total, these are 0.58% of all Sr observations.

We create three alternative measurement versions, each dealing with the text flags differently. We call them V1, V2 and V3:

In the first version (V1), we assign missing values to any measurement that has a non-missing “Result Detection Condition Text.” This approach basically eliminates these ion concentrations, which is similar to using only “uncensored” data, as discussed in (39). While this approach avoids the use of ambiguous measurements, it could work in favor of finding a positive association, as it ignores instances when ion concentrations are very low or essentially zero. This is why it is not our preferred version. We use V1 only to gauge the sensitivity of our results to different ways of dealing with observations with text flags indicating measurement issues.

In the second version (V2), we use additional information in the “Result Detection Condition Text” to assign values, even when the measurement value is at first missing. This approach mostly adds zero values when the text states that an ion was “not detected.” It is our preferred version and the one we use in all reported analyses. The Data Description (on p. 3 above) provides the details on this measurement version.¹⁹

In the third version (V3), we also include readings where the database indicates that the ion was present but below the detection limit and code them as zeros. This approach is the most inclusive. Specifically:

- a) We replace a missing measurement value with the numerical value reported in the “Result Detection Condition Text,” if that variable provides a quantification limit value, divided by 2 to set the value at the midpoint (same as V2);
- b) We assign a value of zero to any measurement, for which the “Result Detection Condition Text” indicates “Not Detected” (as in V2) or “Present Below Quantification Limit” but without providing a numerical value (in V2 those observations are set to missing);

¹⁹ We also set to missing measurements, for which the “Result Detection Condition Text” equals “NA”, “Present Above Quantification Limit” and “Systematic Contamination.” But these cases do not end up in our final sample.

- c) We keep the missing value if the “Result Detection Condition Text” flag equals to “Not Reported.”

We estimate results for eq. S1 and eq. S2 for all three versions. For Cl^- , Ba, and Sr, the results are not materially different for the three versions and hence not affected by the way of dealing with the text flags. For Br^- , we obtain significantly positive associations for *#wellsHUC10* if we use V1. Thus, our choice of V2 for the analyses presented in the manuscript is conservative.

Analysis of the Timing of Water Quality Sampling

The timing of water quality sampling is likely not random. One concern is that such timing could be systematically related to HF activity. We therefore check that the results are not unduly affected by non-random timing of water quality measurements. First, we plot the frequency of water quality measurements by calendar month and do the same for frequency of new HF wells. Figs. S3 and S4 document that water measurements are more commonly taken in spring and summer. HF spudding occurs more commonly in the summer and fall.

Next, to test the relation more formally, we re-shape the data at the monitoring station-monthly level, and code up a binary variable that is equal to one when a monitoring station draws a water quality reading (for any of the four analytes) in a given month, and equal to zero otherwise. We then regress this binary variable on *#wellsHUC10* to examine if there is a systematic relation between the timing of water measurement and new HF wells. Table S6 reports the results and shows that the relation is insignificant, i.e., new HF wells do not predict water quality monitoring in a systematic way, which is reassuring for our analysis. In additional analyses, available upon request, we reach the same conclusion when we split *#wellsHUC10* in two non-overlapping cumulative well count variables, one counting the number of wells spudded within 15 km from monitors within the respective HUC10 and one counting wells that are farther away. This model does not find that a systematic relation between water sampling and the distance between the monitor and new HF wells.

Alternative Time Windows around Well Spud Dates in the Temporal Analysis

As a robustness check, we re-estimate the model presented in Fig. 3 and Table S8 using alternative windows for the time-specific well counts. First, we define counts for the following windows in days: [-360, -151], [-150, -51], [-50, 0], [1, 50], [51, 150], [151, 360], [> 360]. In this specification, we find that increases in the ion concentrations occur at 51–150 days after the spud date, consistently for all four ions (not tabulated, but available upon request). Second, we define counts for the following windows in days: [-360, -181], [-180, -121], [-120, -61], [-60, 0], [1, 60], [61, 120], [121, 180], [181, 360], [> 360]. In this specification, we find that increases in the ion concentrations occur at 121–180 days after the spud date, again consistently for all four ions (not tabulated, but available upon request).

Overall, these findings show that the temporal analysis presented in Fig. 3 and Table S8 is robust to changes in time-window definitions. We find substantially elevated ion concentrations in surface waters shortly after the spud date of a new well. The analyses with different time windows suggest that increases most likely occur between 90 and 150 days of a new spudding. The consistency in the patterns across ions and in the timing of the increases is remarkable and further supports the interpretation of the results provided in the manuscript. Although the temporal evidence does not identify the mechanism for the documented HF well-ion association, it

statistically links elevated concentrations to the unconventional O&G development process and makes alternative explanations unlikely.

Controlling for Facilities Accepting HF Wastewater

Elevated Cl^- concentrations downstream from treatment facilities accepting HF wastewater in Pennsylvania have been documented in (38). Increased Cl^- and Br^- concentrations downstream of effluents from wastewater treatment plants have been found in Western Pennsylvania (29). We therefore control for the presence of water treatment facilities in the same watershed in order to separate the potential impact of new HF wells from the impact of water treatment facilities. We collect data on treatment facilities in Pennsylvania from the PA DEP Oil and Gas Reports. We re-estimate eq. S1 adding the cumulative number of wastewater facilities accepting shale O&G waste (*#wastewater facilities*) as an additional control variable. This variable captures over-time changes in the number of wastewater facilities accepting shale oil and gas waste within a given watershed. Thus, if the results were driven by wastewater treatment facilities, the coefficients on *#wellsHUC10* should be attenuated once the control for treatment facilities is added to the model. The results, reported in Table S13, show that, for one Sr specification, the coefficient on *#wastewater facilities* is significantly positive. More importantly, for Cl^- , Ba and Sr, the *#wellsHUC10* coefficients are not attenuated and continue to be significantly associated with higher ion concentrations, even after controlling for wastewater facilities accepting shale oil and gas waste. The results (not tabulated) are very similar when the *#wastewater facilities* count is defined for upstream wastewater facilities only. Thus, the results suggest that the association between HF wells and ion concentrations is not driven by facilities accepting shale O&G wastewater.

Conventional Wells and Water Quality

To shed further light on whether the results in the manuscript are specific to HF wells, we also examine the associations between *conventional* wells and ion concentrations. Although conventional O&G development could also affect surface waters, HF and conventional wells differ in their potential water impacts. For this reason, the results for conventional wells should not necessarily be the same as for HF wells. For Pennsylvania, the wells recorded in the PA DEP and PA DCNR databases are identified as conventional or unconventional wells. For the other U.S. states, we follow the prior literature (60) and code *vertical* wells as conventional wells. Note that measurement error in the assignment of conventional wells likely biases results in the direction of a positive association, if some vertical wells are actually unconventional ones (and unconventional ones have positive associations as our results show). We emphasize that the conventional well analysis is not a placebo test, at least not as far as Br^- is concerned. Although the literature considers Br^- a signature of HF-related contamination, bromides are also and arguably more extensively used in *conventional* drilling. They are important components of “clear drilling muds,” i.e., water-based muds, which represent the most common muds in conventional drilling. These clear drilling muds are used throughout the conventional drilling cycle, which could make water impacts for Br^- relatively more likely for conventional wells compared to HF wells.

We estimate an analogous model for conventional wells using eq. S1, but now the variable of interest is *#wellsHUC10_conventional*, which measures the cumulative number of *conventional* wells spudded within the HUC10 up to the day of water measurement. We estimate the model for all HUC10s with conventional wells. Table S14 reports the results. We do not find that

conventional wells are associated with increases in ion concentrations, except in one specification for Br^- , which could reflect the use of clear drilling muds, as discussed.

HF Wells and Water Quality – Controlling for Known Spills

We check that our results do not simply reflect known spills related to unconventional O&G development. To explore this possibility, we estimate two alternative versions of eq. S1.

First, we add as an additional control the cumulative number of spills within HUC10 k in the past year ($\#Spill[0;1\text{-Year}]$). The variable $\#Spill[0;1\text{ Year}]$ counts the number of known spills within a HUC10 in the past year. In most cases, this variable increases by one at the spill date, stays at one in the year after the spill date, and then goes back to zero. If there is another spill in the same HUC10 of the first spill, and the spill date is within a year from the first spill, then the variable goes to two, and so on. Essentially, the coefficient on this variable estimates the average increase in the respective ion concentration for the year after the spill.

Second, to account for heterogeneity across spills, we add two separate control variables for known spills. We add the cumulative number of spills with an above the median spill amount within HUC10 k in the past year ($\#Spill_h[0;1\text{-Year}]$) and the cumulative number of spills with a below the median spill impact within HUC10 k in the past year ($\#Spill_l[0;1\text{-Year}]$).

In both versions, we also include the cumulative, long-run spill count, $\#Spill[>1\text{ Year}]$, which counts the number of spills within a HUC10 beyond one year. This variable accounts for long-run changes in ion concentrations due to spills and ensures that the model properly distinguishes between pre- and post-spill observations.

Spills data come from (35, 51). They collect a large sample of unconventional O&G spills in Colorado, North Dakota, New Mexico, and Pennsylvania. We merge this sample with the water quality sampled used in Table S4, Columns (3) and (4). The resulting sample contains 3,531 spills. We deliberately use a large and fairly comprehensive sample to control for known spills. We recognize that not all of these spills are likely relevant for analyses using Br^- , Cl^- , Ba, and Sr, but including all spills as controls is conservative.²⁰

Table S15 reports the results. For brevity we report only the coefficient on the cumulative number of HF wells, $\#wellsHUC10$, which is the coefficient of interest. If a positive HF well-ion association were to reflect the effects of known spills, we would expect the $\#wellsHUC10$ coefficient to decrease in magnitude or in statistical significance, once we control for the spills. We still find significant $\#wellsHUC10$ coefficients, even after controlling for known spills (except in one specification for Ba in PA). Moreover, we see little attenuation in the $\#wellsHUC10$ coefficients, if anything, some get larger. We perform an analogous analysis for Table S12,

²⁰ As robustness test, we use two alternative spill samples and focus on more relevant spills. First, we use spills from Table 5 in (4), which lists 32 major spills related to unconventional shale gas development in Pennsylvania. We use only 18 spills involving saltwater/flowback and HF fluids, which we merge with the water quality sample used in Table S4, Columns (1) and (2), resulting in a loss of four spills, for which we do not have relevant water measurements. The results, available upon request, confirm the inferences from Table S15 that the HF results go beyond known spills. Second, we consider a subset of spills from (35), involving saltwater/flowback and HF fluids, which we merge with the water quality sample used in Table S4, Columns (3) and (4), resulting in a sample of 1,240 spills. The results, available upon request, confirm the inferences from Table S15.

controlling for spills from (35). Inferences for the coefficient of interest, *Post Spud*, do not change. Thus, the results presented in the manuscript go beyond known spills.²¹

Sensitivity Checks

The following analyses are not tabulated, but available upon request:

1. Rivers versus lakes

The sample includes surface water observations coming from rivers and lakes. The group ‘Lakes’ includes lakes, ponds, and reservoirs. The group ‘Rivers’ includes rivers, streams, and canals. However, water observations from rivers dominate the sample: 97.80% for Br^- , 95.40% for Cl^- , 94.64% for Ba and 97.46% for Sr. Thus, there are not enough observations to estimate HF associations separately for different surface water types (e.g., lakes vs. rivers). However, we can examine the association between HF wells and ion concentrations, restricting the water quality observations to ‘Rivers’ only. The results for the restricted sample are very similar. The coefficients of interest continue to be statistically significant and are very similar in magnitude.

2. Regional differences and increases in water sampling over time

It is possible that over time more water quality samples are taken in areas, where new HF wells are drilled, and that these areas could be naturally more saline, e.g., because of natural brine migration. In this case, a positive association between new HF wells and ion concentrations could be spurious if new watersheds enter the analysis and the model does not account for regional differences in water measurements over time. In our model, the inclusion of the monthly sub-region or sub-basin fixed effects, which are both regional and flexible over time, mitigate this possibility. But we nevertheless conduct two additional analyses to gauge this concern. The results of both tests suggest that the findings reported in Table S4 are not driven by changes in the watershed composition and regional differences in water sampling over time.

First, we re-estimate eq. S1 keeping the sample of HUC10s constant. Specifically, we retain a watershed only if it has at least five years of data with at least one observation in each quarter within each year, starting with the first water quality observation for this HUC10. The results using this constant sample are very similar and the inferences essentially the same.

Second, we re-estimate eq. S1 collapsing the dataset at the ion-HUC10-month-year level. Specifically, we take the ion-HUC10-month-year average of all the variables entering in eq. S1, plus the end-of-the month value for *#wellsHUC10*. By averaging observations at the HUC10-month-year level for each ion, we generate a sample that is more balanced over time, even when there are increases in the frequency of water quality measurement over time. The results for this more balanced sample still show a significantly positive association between HF wells and elevated ion concentrations.

²¹ An alternative way to show that our results are not driven by detected spills is to exclude water quality observations from the sample that stem from locations with spills, rather than to control for spills. As there is at least one spill in all the watersheds with HF activity, we cannot simply eliminate all watersheds. Instead, we exclude observations from the Table S4 sample that are: a) taken within 90 days from a spill date and b) drawn from monitoring stations located in the same HUC12 as a spill. We consider all spills involving HF chemicals, saltwater, and with water impact or with likely water impact. The results (untabulated) are similar to those reported in Table S15.

3. Well-monitor pair analysis for other water quality analytes not specific to HF

As reported in the manuscript, Table S7 shows that three other water quality analytes that are not as specific to HF (dissolved oxygen, phosphorus and fecal coliforms) do not exhibit significantly positive associations with new HF wells. This result stands in contrast to Table S4, but it is not a perfect placebo test. Phosphorus could increase due to shale gas development, mostly because of land disturbance for HF infrastructure releasing phosphorus that built up in the soil due to agriculture (61). Moreover, these proxies could be related to other economic activities such as agriculture (43), which could be indirectly related to O&G development. Thus, this analysis also serves to gauge how well our models control for the impact of regional economic activity on water quality, which is important when estimating the effect of HF wells. Here, we repeat this analysis of the three analytes for the well-monitor pairs. We re-estimate eq. S2. As in Table S7, we do not find that new HF wells are associated with significant increases in the three analytes.

4. Multiple Wells and Well Pad Analysis

In eq. S2, the dependent variable (i.e., the ion concentration at a given monitor on a given day) could be influenced by several wells in the vicinity of the respective monitor and in that sense the dependent variable is measured at a higher level of aggregation than the independent variable (i.e., *Post Spud*), which is set at the individual well level. Imagine two wells in an area are spudded in sequence and assume further that both have impact. In this case, the same elevated water measurements are post-spud observations for the first well, but pre-spud observations for the second well. This situation could result in a negative estimate for the second well, if the impact of the first well is larger but declines after the spud date of the second well. The converse is possible too, if the spud dates of multiple wells are very close in time, e.g., when wells are drilled on the same pad. Because our analysis pairs wells with monitors, wells that are spudded on the same pad within a short period of time are evaluated using essentially the same water measurements. In this situation, the model could show effects for each well on the pad even when only one well has surface water impact, overstating the magnitude of the estimated impact. As it is not uncommon to drill multiple wells on the same pad, we gauge the relevance of this issue for our estimates. We identify wells that are likely on the same well pad and then keep only the first well that is spudded on the pad, removing all others that are spudded within 180 days of the first well. We do not have well pad information and therefore assume that wells located within 1.11 km² from each other are on the same pad (i.e., these are wells that share the same latitude/longitude up to 2-decimal digits). Eliminating wells from the sample obviously reduces statistical power. We nevertheless find similar results and inferences (untabulated). Clustering of wells on a well pad could also overstate the calculated HUC10 impact reported in Fig. 4. If we compute the 360-day impact for the average watershed by multiplying the respective estimated coefficient with the sample mean ion concentration and the average number of new pads per year and watershed, we obtain 19,269.27 µg/l for Cl⁻ (17,623.08 µg/l in Pennsylvania), 22.12 µg/l for Ba (21.54 µg/l in Pennsylvania), and 159.28 µg/l for Sr (91.49 µg/l in Pennsylvania).

Fig. S1. Time trends in HF Activity in the U.S. The plot shows the total number of new HF wells in the U.S. per month. New HF wells are assigned to a year and month by their spud date. Data on HF wells come from the WellDatabase, Enverus (formerly Drillinginfo), the Pennsylvania Department of Environmental Protection (PA DEP) and the Pennsylvania Department of Conservation of Natural Resources (PA DCNR).

A

B

C

Fig. S2. Location of watersheds with HF wells and water monitoring stations across the U.S. Figure shows watersheds (HUC10s) with HF wells (*in ocher*) and water quality monitoring stations recording Br^- , Cl^- , Ba, and Sr (*blue dots*) in these watersheds for the North East region (A), the South West region (B), and the Mountain and North West regions (C). Data on the location of water monitoring stations come from the EPA (STORET data), USGS (NWIS data), the Shale Network (2), Susquehanna River Basin Commission, and from the PA DEP (SAC046). Thin black lines depict HUC10s boundaries; thick black lines depict state boundaries.

Fig. S3. Timing of water sampling throughout the year [total number of water readings]. Total number of water quality readings are plotted by month of the year, using water measurements in treated HUC10s for all sample years.

Fig. S4. Timing of well spudding throughout the year [total number of spuds]. Total number of HF well spuddings are plotted by month of the year, using all sample wells and all sample years.

A

B

Fig. S5. HF Wells and Water Quality – Cumulative Well Counts by Distance. (A) Cl^- and (B) All ions. Panels A and B show the OLS coefficients from the estimation of eq. S1 in which we replace #wellsHUC10 with two separate well counts: #wellsHUC10 [0km; Xkm], counting the number of wells spudded within X km of monitors within a given HUC10, with X = 5, 10, 15, 20, 25, 30, and #wellsHUC10 [30km; Max], counting the remaining number of wells that are at least 30 km from monitors within a given HUC10. For each value of X, we run a separate regression with both counts included and plot the nearby-well count coefficient, #wellsHUC10 [0km; Xkm] for the different values of X to show a distance gradient. The coefficient on #wellsHUC10 [30km; Max] is always included in the model but not reported. We report the estimated coefficients together with the respective 90% confidence intervals.

Table S1. Descriptive Statistics for Surface Water Quality Measurements (µg/L)*Panel A – Treated HUC10s (all U.S. Watersheds)*

Bromide	N	Mean	P25	P50	P75	SD
Concentration	6,642	97.384	0.000	0.000	73.000	433.748
Ln(Concentration)	6,642	1.979	0.000	0.000	4.304	2.345
Chloride						
Concentration	27,628	47,522.550	5,600.000	14,600.000	38,300.000	186,007.700
Ln(Concentration)	27,628	9.490	8.631	9.589	10.553	1.907
Barium						
Concentration	13,578	59.374	31.700	47.000	70.000	94.615
Ln(Concentration)	13,578	3.734	3.487	3.871	4.263	1.007
Strontium						
Concentration	12,935	330.165	50.000	143.000	310.000	597.155
Ln(Concentration)	12,935	4.944	3.932	4.970	5.740	1.269

Panel B – Treated HUC10s (Pennsylvania only)

Bromide	N	Mean	P25	P50	P75	SD
Concentration	4,213	134.311	0.000	19.918	100.000	536.552
Ln(Concentration)	4,213	2.444	0.000	3.041	4.615	2.408
Chloride						
Concentration	9,843	19,772.050	4,510.000	10,200.000	21,010.000	35,613.940
Ln(Concentration)	9,843	9.205	8.414	9.230	9.953	1.157
Barium						
Concentration	3,491	39.655	27.000	36.000	45.000	22.005
Ln(Concentration)	3,491	3.601	3.332	3.611	3.829	0.440
Strontium						
Concentration	7,070	121.520	30.000	59.000	136.000	208.190
Ln(Concentration)	7,070	4.228	3.434	4.094	4.920	0.980

Panel C – Non-treated HUC10s within treated HUC4s

Bromide	N	Mean	P25	P50	P75	SD
Concentration	18,731	100.947	0.000	0.000	18.300	1,301.598
Ln(Concentration)	18,731	1.227	0.000	0.000	2.960	2.040
Chloride						
Concentration	120,537	362,030.100	5,730.000	17,000.000	41,300.000	1,864,807.000
Ln(Concentration)	120,537	9.614	8.654	9.741	10.629	2.392
Barium						
Concentration	33,920	58.176	31.800	48.000	72.000	51.925
Ln(Concentration)	33,920	3.761	3.490	3.891	4.290	0.994
Strontium						
Concentration	16,569	578.917	110.000	270.000	580.000	1,061.629
Ln(Concentration)	16,569	5.513	4.710	5.602	6.365	1.372

Panels A, B, and C present descriptive statistics for surface water quality measurements. Panel A reports surface water ion concentrations for all U.S. watersheds (HUC10) with HF wells during our sample period. Panel B reports surface water ion concentrations for watersheds (HUC10) in Pennsylvania with HF wells. Panel C reports surface water ion concentrations for watersheds (HUC10) that do not have HF wells during our sample period but that are located within HUC4 with HF wells. The table reports statistics for the raw ion concentrations and after applying the natural logarithm (ln).

Table S2. Number of Water Quality Measurements by Region and Shale

Region	Shale	Bromide	Chloride	Barium	Strontium
North East Region	Antrim		79	39	39
	Devonian (Ohio)	32	1,949	1,685	1,701
	Marcellus	4,726	14,077	7,037	10,481
Mountain and North West Regions	Bakken	63	1,644	1,106	161
	Heath	4	11	13	
	Hermosa	9	282	159	
	Hilliard-Baxter-Mancos	32	280	148	
	Lewis	16	127	26	8
	Mancos		1,108	602	
	Manning Canyon		66	8	
	Mowry		4	4	4
	Niobrara	57	2,229	859	208
	Niobrara- Mowry		763	556	
	Pierre -Niobrara		2		
	Three Forks		61	47	
	Total				
South West Region	Abo-Yeso	2	27	2	
	Barnett		540	23	
	Bone Spring		44		
	Eagle Ford	4	14	14	14
	Excello-Mulky	232	568	243	243
	Fayetteville	1,453	2,060	929	
	Tuscaloosa		28		
	Woodford	12	911	78	76
	Woodford-Caney		754		
	Total	6,642	27,628	13,578	12,935
This table presents the total number of water quality observations in the sample by region and shale for each ion. Shales are broadly grouped by U.S. regions.					

Table S3. Descriptive Statistics on Surface Water Quality Observations and HF Wells

Panel A – Distribution of Surface Water Quality Observations and HF Wells across Treated HUC10s (all U.S. Watersheds)

Treated HUC10	Unique watersheds	N				
	408	60,783				
Unique treated HUC10 by state	N	Mean	P25	P50	P75	SD
	408	50	20	44	55	34
Unique monitoring stations by HUC10	N	Mean	P25	P50	P75	SD
	2,038	12	4	9	17	10
Unique monitoring stations by HUC10/ion	N	Mean	P25	P50	P75	SD
Bromide	459	6	2	4	6	6
Chloride	1,984	11	4	9	17	9
Barium	1,285	13	4	10	20	11
Strontium	1,072	14	5	13	20	11
Water quality observations by HUC10/ion	N	Mean	P25	P50	P75	SD
Bromide	6,642	46	15	30	61	45
Chloride	27,628	68	13	41	95	80
Barium	13,578	47	9	28	64	55
Strontium	12,935	72	14	56	108	72
Distribution of the well count variable by ion (HF wells)		Mean	P25	P50	P75	SD
Bromide		116	5	26	127	234
Chloride		48	1	5	29	139
Barium		46	1	6	27	135
Strontium		43	2	12	41	80
Average distance between wells and monitoring station [km]		Mean	P25	P50	P75	SD
Bromide		15.813	8.532	14.092	21.411	9.418
Chloride		14.688	7.854	13.184	20.214	8.808
Barium		15.069	8.179	13.739	20.450	8.929
Strontium		14.853	7.898	13.282	20.096	9.067

Panel B – Distribution of Surface Water Quality Observations and HF Wells across Treated HUC10s (Pennsylvania only)

Treated HUC10	Unique watersheds	N				
	105	24,628				
Unique monitoring stations by HUC10	N	Mean	P25	P50	P75	SD
	406	6	3	5	7	5
Unique monitoring stations by HUC10/ion	N	Mean	P25	P50	P75	SD
Bromide	205	3	2	3	4	2
Chloride	393	6	3	5	7	5
Barium	204	4	2	3	5	2
Strontium	275	4	2	4	6	2
Water quality observations by HUC10/ion	N	Mean	P25	P50	P75	SD
Bromide	4,213	53	26	42	64	44
Chloride	9,853	96	15	74	145	85
Barium	3,492	40	13	27	55	33
Strontium	7,070	78	17	59	111	66
Distribution of the well count variable by ion (HF wells)		Mean	P25	P50	P75	SD
Bromide		66	5	22	83	98
Chloride		48	2	13	52	82
Barium		60	5	21	82	86
Strontium		62	5	22	78	91
Average distance between wells and monitoring station [km]		Mean	P25	P50	P75	SD
Bromide		16.034	8.826	14.462	21.916	9.187
Chloride		14.372	7.718	12.987	20.000	8.468
Barium		15.388	8.513	14.350	21.238	8.640
Strontium		15.485	8.419	14.163	21.301	8.946

Panel A presents summary statistics on the distribution of surface water quality observations and HF wells across all U.S. HUC10s. Panel B provides the same summary statistics on the distribution of surface water quality observations and HF wells across HUC10s but for Pennsylvania only.

Table S4. HF Wells and Water Quality

	(1) PA	(2) PA	(3) ALL	(4) ALL
Bromide (µg/L)				
<i>#wellsHUC10</i>	0.00074 [0.00060]	0.00019 [0.00107]	0.00035* [0.00019]	0.00020 [0.00040]
Observations	4,213	4,194	6,642	6,636
R-squared	0.886	0.926	0.881	0.912
Chloride (µg/L)				
<i>#wellsHUC10</i>	0.00059*** [0.00021]	0.00071** [0.00032]	0.00054*** [0.00019]	0.00056* [0.00030]
Observations	9,843	9,792	27,628	27,613
R-squared	0.924	0.953	0.862	0.895
Barium (µg/L)				
<i>#wellsHUC10</i>	0.00051** [0.00025]	0.00038* [0.00021]	0.00022 [0.00017]	0.00023 [0.00015]
Observations	3,491	3,450	13,578	13,571
R-squared	0.801	0.853	0.831	0.863
Strontium (µg/L)				
<i>#wellsHUC10</i>	0.00042*** [0.00014]	0.00041*** [0.00007]	0.00041*** [0.00013]	0.00036*** [0.00008]
Observations	7,070	7,044	12,935	12,930
R-squared	0.934	0.961	0.967	0.977
Monitoring station FE	Yes	Yes	Yes	Yes
Weather controls	Yes	Yes	Yes	Yes
HUC4*Month*Year FE	Yes	No	Yes	No
HUC8*Month*Year FE	No	Yes	No	Yes

This table reports coefficients from OLS estimation of eq. S1 for treated watersheds (i.e., HUC10s with at least one active HF well over the sample period). *#wellsHUC10* measures the cumulative number of HF wells in a given HUC10 at a given point in time. Columns (1) and (2) present results for Pennsylvania only and Columns (3) and (4) present results for all treated U.S. watersheds.

Standard errors clustered by sub-basin (HUC8) are reported below the coefficients. *, **, *** indicate significance at the two-tailed 10%, 5%, and 1% levels, respectively.

Table S5. HF Wells and Water Quality – Extensions to Table S4 Analysis

Panel A – HF Wells and Water Quality – Pennsylvania vs. other U.S. states

	(1)	(2)
	ALL	ALL
Bromide (µg/L)		
#wellsHUC10-PA	0.00064 [0.00056]	0.00015 [0.00104]
#wellsHUC10-non PA	0.00018* [0.00010]	0.00023 [0.00017]
Observations	6,642	6,636
R-squared	0.881	0.912
Chloride (µg/L)		
#wellsHUC10-PA	0.00056*** [0.00021]	0.00063** [0.00031]
#wellsHUC10-non PA	0.00052* [0.00028]	0.00050 [0.00045]
Observations	27,628	27,613
R-squared	0.862	0.895
Barium (µg/L)		
#wellsHUC10-PA	0.00049** [0.00024]	0.00037* [0.00021]
#wellsHUC10-non PA	0.00005 [0.00016]	0.00015 [0.00017]
Observations	13,578	13,571
R-squared	0.831	0.863
Strontium (µg/L)		
#wellsHUC10-PA	0.00041*** [0.00014]	0.00035*** [0.00008]
#wellsHUC10-non PA	0.00041** [0.00017]	0.00041** [0.00017]
Observations	12,935	12,930
R-squared	0.967	0.977
Monitoring station FE	Yes	Yes
Weather controls	Yes	Yes
HUC4*Month*Year FE	Yes	No
HUC8*Month*Year FE	No	Yes

Panel A reports coefficients from the OLS estimation of eq. S1 for treated HUC10s. We replace the #wellsHUC10 variable with two non-overlapping, cumulative well counts, one for Pennsylvania and one for all the other U.S. states, respectively.

Reported below the coefficients are standard errors clustered at HUC8-level. *, **, *** indicate significance at the two-tailed 10%, 5%, and 1% levels, respectively

Panel B – HF Wells and Water Quality – Estimates for Different Time Periods

	(1) PA	(2) PA	(3) ALL	(4) ALL
Bromide (µg/L)				
#wellsHUC10[2006; 2010]	0.00070 [0.00178]	-0.00054 [0.00147]	0.00037 [0.00073]	-0.00019 [0.00069]
#wellsHUC10[2011; 2013]	0.00103 [0.00079]	0.00034 [0.00110]	0.00029 [0.00032]	-0.00013 [0.00044]
#wellsHUC10[2014; 2016]	0.00079 [0.00063]	0.00017 [0.00107]	0.00033 [0.00023]	0.00011 [0.00040]
Observations	4,213	4,194	6,642	6,636
R-squared	0.886	0.926	0.881	0.912
Chloride (µg/L)				
#wellsHUC10[2006; 2010]	0.00209 [0.00149]	0.00221 [0.00190]	0.00152* [0.00088]	0.00184 [0.00120]
#wellsHUC10[2011; 2013]	0.00058* [0.00032]	0.00083* [0.00046]	0.00073*** [0.00026]	0.00069* [0.00040]
#wellsHUC10[2014; 2016]	0.00067*** [0.00023]	0.00076* [0.00039]	0.00059*** [0.00020]	0.00061* [0.00033]
Observations	9,843	9,792	27,628	27,613
R-squared	0.924	0.953	0.862	0.895
Barium (µg/L)				
#wellsHUC10[2006; 2010]	0.00052 [0.00031]	0.00002 [0.00034]	-0.00043 [0.00037]	-0.00036 [0.00038]
#wellsHUC10[2011; 2013]	0.00051** [0.00021]	0.00050*** [0.00015]	0.00024 [0.00019]	0.00017 [0.00019]
#wellsHUC10[2014; 2016]	0.00051** [0.00024]	0.00034* [0.00020]	0.00020 [0.00017]	0.00020 [0.00015]
Observations	3,491	3,450	13,578	13,571
R-squared	0.801	0.853	0.831	0.863
Strontium (µg/L)				
#wellsHUC10[2006; 2010]	0.00124** [0.00057]	0.00088** [0.00035]	0.00077 [0.00062]	0.00017 [0.00053]
#wellsHUC10[2011; 2013]	0.00065*** [0.00021]	0.00074*** [0.00015]	0.00058*** [0.00020]	0.00060*** [0.00015]
#wellsHUC10[2014; 2016]	0.00047*** [0.00013]	0.00042*** [0.00007]	0.00043*** [0.00012]	0.00034*** [0.00007]
Observations	7,070	7,044	12,935	12,930
R-squared	0.934	0.961	0.967	0.977
Monitoring station FE	Yes	Yes	Yes	Yes
Weather controls	Yes	Yes	Yes	Yes
HUC4*Month*Year FE	Yes	No	Yes	No
HUC8*Month*Year FE	No	Yes	No	Yes

Panel B reports coefficients from the OLS estimation of eq. S1 for treated HUC10s. We replace the #wellsHUC10 variable with non-overlapping #wellsHUC10 counts for three time periods: 2006-2010, 2011-2013, 2014-2016. Columns (1) and (2) present results for Pennsylvania and Columns (3) and (4) present results for all treated U.S. watersheds.

Reported below the coefficients are standard errors clustered at HUC8-level. *, **, *** indicate significance at the two-tailed 10%, 5%, and 1% levels.

Panel C – All HUC10s within HUC4s with HF Wells

	(1) PA	(2) PA	(3) ALL	(4) ALL
Bromide (µg/L)				
#well HUC10	0.00019 [0.00070]	0.00026 [0.00105]	0.00056* [0.00033]	0.00032 [0.00039]
Observations	7,367	6,856	23,795	23,251
R-squared	0.869	0.917	0.752	0.835
Chloride (µg/L)				
#well HUC10	0.00044** [0.00018]	0.00068** [0.00030]	0.00024 [0.00017]	0.00059** [0.00029]
Observations	22,165	22,037	138,470	138,160
R-squared	0.908	0.936	0.876	0.906
Barium (µg/L)				
#well HUC10	0.00037* [0.00021]	0.00040* [0.00020]	0.00001 [0.00015]	0.00027* [0.00015]
Observations	7,041	5,963	43,574	42,398
R-squared	0.804	0.842	0.787	0.820
Strontium (µg/L)				
#well HUC10	0.00035** [0.00015]	0.00036*** [0.00005]	0.00032** [0.00013]	0.00032*** [0.00006]
Observations	10,144	10,101	27,353	27,285
R-squared	0.935	0.960	0.952	0.964
Monitoring station FE	Yes	Yes	Yes	Yes
Weather controls	Yes	Yes	Yes	Yes
HUC4*Month*Year FE	Yes	No	Yes	No
HUC8*Month*Year FE	No	Yes	No	Yes
Panel C reports coefficients from the OLS estimation of eq. S1 for treated HUC10s plus all untreated HUC10s (i.e., no HF wells over the sample period) that are located in HUC4s with at least one treated HUC10. Columns (1) and (2) present results for Pennsylvania only and Columns (3) and (4) present results for all U.S. watersheds. Reported below the coefficients are standard errors clustered at HUC8-level. *, **, *** indicate significance at the two-tailed 10%, 5%, and 1% levels, respectively.				

Panel D – HF Wells and Water Quality – Controlling for Snow and Road Salts

	(1) PA	(2) PA	(3) ALL	(4) ALL
Bromide (µg/L)				
<i>#wellsHUC10</i>	0.00074 [0.00059]	0.00016 [0.00104]	0.00035* [0.00019]	0.00019 [0.00039]
<i>Snow</i>	0.03434 [0.13147]	0.10989 [0.18073]	0.00423 [0.10633]	0.08275 [0.13914]
Observations	4,213	4,194	6,642	6,636
R-squared	0.886	0.926	0.881	0.912
Chloride (µg/L)				
<i>#wellsHUC10</i>	0.00058*** [0.00021]	0.00070** [0.00032]	0.00054*** [0.00019]	0.00056* [0.00030]
<i>Snow</i>	0.09957** [0.03853]	0.07641 [0.06025]	-0.00204 [0.04573]	0.00567 [0.04280]
Observations	9,843	9,792	27,628	27,613
R-squared	0.924	0.953	0.862	0.895
Barium (µg/L)				
<i>#wellsHUC10</i>	0.00050** [0.00025]	0.00038* [0.00022]	0.00023 [0.00017]	0.00023 [0.00015]
<i>Snow</i>	0.00794 [0.02273]	-0.00496 [0.02504]	-0.01874 [0.02144]	-0.00248 [0.03011]
Observations	3,491	3,450	13,578	13,571
R-squared	0.801	0.853	0.831	0.863
Strontium (µg/L)				
<i>#wellsHUC10</i>	0.00041*** [0.00015]	0.00041*** [0.00007]	0.00041*** [0.00013]	0.00035*** [0.00008]
<i>Snow</i>	0.03754 [0.02974]	0.03680 [0.03698]	0.03845 [0.02950]	0.04747 [0.03265]
Observations	7,070	7,044	12,935	12,930
R-squared	0.934	0.961	0.967	0.977
Monitoring station FE	Yes	Yes	Yes	Yes
Weather controls	Yes	Yes	Yes	Yes
HUC4*Month*Year FE	Yes	No	Yes	No
HUC8*Month*Year FE	No	Yes	No	Yes

Panel D reports coefficients from the OLS estimation of eq. S1 for treated HUC10s including a snow indicator as control. Columns (1) and (2) present results for Pennsylvania only and Columns (3) and (4) present results for all treated U.S. watersheds. Snow is a binary variable that equals one if on day d and at monitor k the average temperature is equal or below 4 degrees Celsius and the cumulative 3-day precipitation up to day d is equal or above 0.5 mm.

Reported below the coefficients are standard errors clustered at HUC8-level. *, **, *** indicate significance at the two-tailed 10%, 5%, and 1% levels, respectively.

Panel E – Sensitivity Analyses

	Controlling for <i>HUC10</i> seasonality		WLS		Controlling for <i>HUC10</i> size	
	(1) PA	(2) ALL	(3) PA	(4) ALL	(5) PA	(6) ALL
Bromide (µg/L)						
<i>#wellsHUC10</i>	-0.00016 [0.00074]	0.00018 [0.00039]	0.00020 [0.00067]	0.00036 [0.00032]	0.04102 [0.68441]	0.10360 [0.28265]
Observations	3,883	6,222	4,194	6,636	4,194	6,636
R-squared	0.938	0.924	0.925	0.923	0.926	0.912
Chloride (µg/L)						
<i>#wellsHUC10</i>	0.00072** [0.00031]	0.00035* [0.00020]	0.00087** [0.00039]	0.00060* [0.00032]	0.47478** [0.19907]	0.40098* [0.22104]
Observations	9,532	26,393	9,792	27,613	9,792	27,613
R-squared	0.957	0.905	0.955	0.908	0.953	0.895
Barium (µg/L)						
<i>#wellsHUC10</i>	0.00033** [0.00015]	0.00038** [0.00017]	0.00039* [0.00022]	0.00010 [0.00017]	0.23830* [0.13846]	0.14950 [0.09603]
Observations	3,114	12,488	3,450	13,571	3,450	13,571
R-squared	0.865	0.885	0.829	0.841	0.853	0.863
Strontium (µg/L)						
<i>#wellsHUC10</i>	0.00045*** [0.00009]	0.00045*** [0.00009]	0.00051*** [0.00004]	0.00016 [0.00036]	0.25825*** [0.05098]	0.22487*** [0.05446]
Observations	6,910	12,447	7,044	12,930	7,044	12,930
R-squared	0.965	0.979	0.961	0.968	0.961	0.977
Monitoring station FE	Yes	Yes	Yes	Yes	Yes	Yes
Weather controls	Yes	Yes	Yes	Yes	Yes	Yes
HUC8*Month*Year FE	Yes	Yes	Yes	Yes	Yes	Yes
HUC10*Quarter FE	Yes	Yes	No	No	No	No

Panel E, Columns (1) and (2) report coefficients from the OLS estimation of eq. S1 for treated HUC10s adding HUC10*Quarter fixed effects to the model. Columns (3) and (4) report coefficients from the WLS estimation of eq. S1 for treated HUC10s using the number of water quality observations by HUC8-Month-year as weights. Columns (5) and (6) report OLS coefficients estimated for treated HUC10s scaling *#wellsHUC10* by the size of HUC10s in km². Columns (1), (3), and (5) present results for Pennsylvania only and Columns (2), (4), and (6) present results for all treated U.S. watersheds.

Reported below the coefficients are standard errors clustered by HUC8. *, **, *** indicate significance at the two-tailed 10%, 5%, and 1% levels, respectively.

Panel F – Alternative Transformations of the Ion Concentrations (to Address Skewness)

	Robust regressions		Concentrations in decile ranks		Concentrations in binary form	
	(1) PA	(2) ALL	(3) PA	(4) ALL	(5) PA	(6) ALL
Bromide (µg/L)						
#wellsHUC10	0.00331** [0.00125]	0.00030 [0.00029]	0.00095 [0.00084]	0.00067** [0.00031]	0.00079*** [0.00023]	0.00027 [0.00020]
Observations	2,888	4,064	4,194	6,634	4,194	6,636
R-squared	0.999	0.999	0.916	0.880	0.815	0.857
Chloride (µg/L)						
#wellsHUC10	4.37440 [2.73740]	3.04962* [1.62405]	0.00081 [0.00112]	0.00059 [0.00056]	0.00054** [0.00020]	0.00026** [0.00011]
Observations	8,325	22,769	9,792	27,613	9,792	27,613
R-squared	0.995	0.995	0.938	0.896	0.773	0.719
Barium (µg/L)						
#wellsHUC10	0.00710** [0.00306]	0.00105 [0.00292]	0.00308*** [0.00059]	0.00030 [0.00121]	0.00022* [0.00011]	0.00007 [0.00016]
Observations	3,276	12,335	3,450	13,562	3,450	13,571
R-squared	0.946	0.997	0.805	0.809	0.648	0.637
Strontium (µg/L)						
#wellsHUC10	0.01726* [0.01003]	0.03287* [0.01739]	0.00174*** [0.00054]	0.00230** [0.00095]	0.00037 [0.00022]	0.00034* [0.00018]
Observations	5,979	10,492	7,044	12,930	7,044	12,930
R-squared	0.991	0.998	0.939	0.926	0.796	0.749
Monitoring station FE	Yes	Yes	Yes	Yes	Yes	Yes
Weather controls	Yes	Yes	Yes	Yes	Yes	Yes
HUC8*Month*Year FE	Yes	Yes	Yes	Yes	Yes	Yes

Panel F reports coefficients from the OLS estimation of eq. S1 for treated HUC10s using transformations of the ion concentrations other than the natural logarithm. In Columns (1) and (2), we use robust regressions with untransformed (or raw) ion concentration measured at monitoring i in HUC10 k on day d . We first estimate weights for each observation based on the rreg procedure in STATA and then apply these weights in an OLS regression so as to appropriately cluster the standard errors. In Columns (3) and (4), we convert ion concentrations into decile ranks computed by HUC4. In Columns (5) and (6), we transform the ion concentration measured at monitoring i in HUC10 k on day d into a binary variable that is equal to one if the concentration is above the 75th percentile of the respective ion concentration distribution by HUC4-Year. Columns (1), (3), and (5) present results for Pennsylvania and Columns (2), (4), and (6) present results for all treated U.S. watersheds.

Reported below the coefficients are standard errors clustered by HUC8. *, **, *** indicate significance at the two-tailed 10%, 5%, and 1% levels, respectively.

Table S6. Time Pattern in Water Quality Sampling relative to New HF Wells

	(1) Bromide measurement	(2) Chloride measurement	(3) Barium measurement	(4) Strontium measurement
<i>#wellsHUC10</i>	0.00005 [0.00004]	0.00006 [0.00004]	0.00005 [0.00003]	0.00001 [0.00001]
Observations	674,880	674,880	674,880	674,880
R-squared	0.429	0.418	0.402	0.386
Sample	ALL	ALL	ALL	ALL
Monitoring station FE	Yes	Yes	Yes	Yes
HUC8*Month*Year FE	Yes	Yes	Yes	Yes

This table reports coefficients from the OLS estimation of the following equation for treated HUC10s:

$$Measurement_{ikm} = station_i + HUC_{kmy} + \alpha \#wellsHUC10_{km} + \varepsilon_{ikm}$$

The dependent variable, *Measurement*, is a binary variable that equals one if an ion concentration has been measured at monitoring station *i* in HUC10 *k* and month *m*. *#wellsHUC10* is the cumulative number of HF wells spudded within HUC10 *k* up to month *m*. The model includes monitoring-station fixed effects and HUC8-month-year fixed effects.

Reported below the coefficients are standard errors clustered by HUC8. *, **, *** indicate significance at the two-tailed 10%, 5%, and 1% levels, respectively

Table S7. HF Wells and Other Water Quality Proxies not Specific to HF

	(1) PA	(2) PA	(3) ALL	(4) ALL
Dissolved oxygen (µg/L)				
<i>#wellsHUC10</i>	-0.00002 [0.00007]	0.00006 [0.00005]	0.00009 [0.00009]	0.00014 [0.00012]
Observations	4,678	4,583	21,991	21,988
R-squared	0.481	0.610	0.567	0.640
Phosphorus (µg/L)				
<i>#wellsHUC10</i>	0.00053 [0.00042]	0.00106 [0.00065]	-0.00046 [0.00038]	-0.00020 [0.00056]
Observations	9,981	9,922	25,940	25,929
R-squared	0.723	0.805	0.776	0.824
Fecal coliforms (CFU/100ml)				
<i>#wellsHUC10</i>	0.00069 [0.00332]	0.00151 [0.01610]	-0.00061 [0.00262]	-0.00164 [0.00617]
Observations	3,354	3,352	7,520	7,520
R-squared	0.655	0.765	0.643	0.704
Sample	Only Treated HUC10s	Only Treated HUC10s	Only Treated HUC10s	Only Treated HUC10s
Monitoring station FE	Yes	Yes	Yes	Yes
Weather controls	Yes	Yes	Yes	Yes
HUC4*Month*Year FE	Yes	No	Yes	No
HUC8*Month*Year FE	No	Yes	No	Yes
<p>This table reports coefficients from the OLS estimation of eq. S1 for treated HUC10s for three water quality analytes that are <u>not</u> as specific to or indicative of HF-related effects on surface water. Columns (1) and (2) presents results for Pennsylvania only and Columns (3) and (4) present results for all treated U.S. watersheds.</p> <p>Reported below the coefficients are standard errors clustered by HUC8. *, **, *** indicate significance at the two-tailed 10%, 5%, and 1% levels, respectively.</p>				

Table S8. HF Wells and Water Quality – Temporal Analysis

<i>Panel A – Watersheds in Pennsylvania</i>				
	(1)	(2)	(3)	(4)
<i>#wellsHUC10</i> [-180, -91]	0.00535 [0.00369]	-0.00432 [0.00305]	-0.00122 [0.00178]	-0.00184** [0.00073]
<i>#wellsHUC10</i> [-90, 0]	0.00152 [0.00228]	-0.00161 [0.00099]	-0.00168 [0.00198]	-0.00092 [0.00160]
<i>#wellsHUC10</i> [1, 90]	-0.00499* [0.00290]	0.00338* [0.00174]	-0.00012 [0.00063]	0.00082*** [0.00029]
<i>#wellsHUC10</i> [91, 180]	0.01184* [0.00651]	0.00352** [0.00156]	0.00491*** [0.00139]	0.00347*** [0.00105]
<i>#wellsHUC10</i> [181,360]	0.00267 [0.00241]	-0.00043 [0.00103]	-0.00051 [0.00082]	-0.00071 [0.00081]
<i>#wellsHUC10</i> [>360]	-0.00107 [0.00108]	0.00049* [0.00026]	0.00021 [0.00046]	0.00033*** [0.00009]
Observations	4,194	9,792	3,450	7,044
R-squared	0.927	0.954	0.854	0.961
Contaminant	Bromide	Chloride	Barium	Strontium
Weather controls	Yes	Yes	Yes	Yes
Monitoring station FE	Yes	Yes	Yes	Yes
HUC8*Month*Year FE	Yes	Yes	Yes	Yes
<i>Panel B – All U.S. Watersheds</i>				
	(1)	(2)	(3)	(4)
<i>#wellsHUC10</i> [-180, -91]	0.00417 [0.00309]	0.00149 [0.00413]	0.00118 [0.00197]	-0.00042 [0.00123]
<i>#wellsHUC10</i> [-90, 0]	0.00145 [0.00216]	-0.00204* [0.00108]	-0.00281** [0.00139]	-0.00077 [0.00144]
<i>#wellsHUC10</i> [1, 90]	-0.00358 [0.00350]	0.00223 [0.00190]	-0.00014 [0.00089]	-0.00042 [0.00098]
<i>#wellsHUC10</i> [91, 180]	0.01095** [0.00510]	0.00401** [0.00174]	0.00347** [0.00143]	0.00289** [0.00116]
<i>#wellsHUC10</i> [181,360]	0.00218 [0.00204]	0.00031 [0.00156]	-0.00088 [0.00064]	-0.00063 [0.00081]
<i>#wellsHUC10</i> [>360]	-0.00063 [0.00076]	0.00044* [0.00024]	0.00023 [0.00023]	0.00039*** [0.00009]
Observations	6,636	27,613	13,571	12,930
R-squared	0.912	0.895	0.863	0.977
Contaminant	Bromide	Chloride	Barium	Strontium
Weather controls	Yes	Yes	Yes	Yes
Monitoring station FE	Yes	Yes	Yes	Yes
HUC8*Month*Year FE	Yes	Yes	Yes	Yes

This table reports coefficients from the OLS estimation of eq. S1 for treated HUC10s, in which we replace *#wellsHUC10* with separate time-period specific HF wells counts, with windows defined over the following days: [-180; -91], [-90; 0], [1; 90], [91; 180], [181; 360]; [>360] from the spud date. For instance, *#wellsHUC10* [91, 180] counts all wells in the respective HUC10 that were spudded between 91 and 180 days ago, relative to the date of the respective ion concentration measurement. The [<-180] window is the omitted category and therefore serves as a benchmark. Panel A presents results for Pennsylvania only and Panel B presents results for all treated U.S. watersheds.

Reported below the coefficients are standard errors clustered by HUC8. *, **, *** indicate significance at the two-tailed 10%, 5%, and 1% levels, respectively.

Table S9. HF Wells and Water Quality – Split by Produced Water

	(1) PA	(2) PA	(3) ALL	(4) ALL
Bromide(µg/L)				
<i>#wellsHUC10_High_Prod_Water</i>	0.00059 [0.00057]	0.00016 [0.00106]	0.00035* [0.00020]	0.00015 [0.00039]
<i>#wellsHUC10_Low_Prod_Water</i>	-0.00385 [0.00289]	-0.00287 [0.00292]	0.00012 [0.00022]	-0.00023 [0.00046]
Observations	4,213	4,194	6,642	6,636
R-squared	0.886	0.926	0.881	0.912
Chloride (µg/L)				
<i>#wellsHUC10_High_Prod_Water</i>	0.00057** [0.00021]	0.00071** [0.00032]	0.00057*** [0.00020]	0.00058* [0.00033]
<i>#wellsHUC10_Low_Prod_Water</i>	-0.00020 [0.00099]	0.00055 [0.00103]	0.00027 [0.00016]	-0.00006 [0.00029]
Observations	9,843	9,792	27,628	27,613
R-squared	0.924	0.953	0.862	0.895
Barium (µg/L)				
<i>#wellsHUC10_High_Prod_Water</i>	0.00047* [0.00025]	0.00037* [0.00021]	0.00024 [0.00018]	0.00023 [0.00015]
<i>#wellsHUC10_Low_Prod_Water</i>	-0.00021 [0.00045]	-0.00031 [0.00041]	0.00000 [0.00027]	0.00023 [0.00021]
Observations	3,491	3,450	13,578	13,571
R-squared	0.801	0.853	0.831	0.863
Strontium (µg/L)				
<i>#wellsHUC10_High_Prod_Water</i>	0.00041** [0.00016]	0.00041*** [0.00007]	0.00041*** [0.00013]	0.00037*** [0.00008]
<i>#wellsHUC10_Low_Prod_Water</i>	0.00026 [0.00069]	0.00079 [0.00070]	0.00009 [0.00045]	0.00027 [0.00038]
Observations	7,070	7,044	12,935	12,930
R-squared	0.934	0.961	0.967	0.977
Monitoring station FE	Yes	Yes	Yes	Yes
Weather controls	Yes	Yes	Yes	Yes
HUC4*Month*Year FE	Yes	No	Yes	No
HUC8*Month*Year FE	No	Yes	No	Yes

This table reports coefficients from the OLS estimation of eq. S1 for treated HUC10s. We replace the *#wellsHUC10* variable with two non-overlapping, cumulative well counts: one counting the cumulative number of HF wells spudded within the HUC10 *k* up to day *d* with an above-median HUC8-year amount of produced water (*#wellsHUC10_High_Prod_Water*) and one counting the cumulative number of HF wells spudded within the HUC10 *k* up to day *d* with a below-median HUC8-year amount of produced water (*#wellsHUC10_Low_Prod_Water*). Columns (1) and (2) present results for Pennsylvania only and Columns (3) and (4) present results for all treated U.S. watersheds.

Reported below the coefficients are standard errors clustered by HUC8. *, **, *** indicate significance at the two-tailed 10%, 5%, and 1% levels, respectively.

Table S10. HF Wells and Water Quality – Split by Salinity of Deep Formations

	(1) Bromide (µg/L)	(2) Chloride (µg/L)	(3) Barium (µg/L)	(4) Strontium (µg/L)
<i>#wellsHUC10_High_Salinity</i>	0.00114 [0.00086]	0.00120* [0.00069]	0.00064*** [0.00021]	0.00040*** [0.00010]
<i>#wellsHUC10_Low_Salinity</i>	0.00001 [0.00064]	0.00023 [0.00017]	0.00003 [0.00016]	0.00033*** [0.00012]
Observations	6,636	27,613	13,571	12,930
R-squared	0.912	0.895	0.863	0.977
Sample	ALL	ALL	ALL	ALL
Monitoring station FE	Yes	Yes	Yes	Yes
Weather controls	Yes	Yes	Yes	Yes
HUC8*Month*Year FE	Yes	Yes	Yes	Yes

This table reports coefficients from the OLS estimation of eq. S1 over all treated U.S. watersheds, as this analysis exploits variation in the salinity of the deep formations across shales and regions. We replace the *#wellsHUC10* variable with two non-overlapping, cumulative well counts: one counting the cumulative number of HF wells spudded within the HUC10 *k* up to day *d* localized in HUC8s with high salinity (*#wellsHUC10_High_Salinity*) and one counting the cumulative number of HF wells spudded within the HUC10 *k* up to day *d* localized in HUC8s with low salinity (*#wellsHUC10_Low_Salinity*). To measure the salinity of the deep formations, we use TDS concentrations of brine samples from previous O&G drilling (see Data Description for details). We obtain similar results if we measure the salinity of the deep formations using Cl⁻ concentrations of brine samples from previous O&G drilling.

Reported below the coefficients are standard errors clustered by HUC8. *, **, *** indicate significance at the two-tailed 10%, 5%, and 1% levels, respectively.

Table S11. HF Wells and Water Quality by HUC10 Type

	(1) PA	(2) PA	(3) ALL	(4) ALL
Bromide (µg/L)				
<i>#wellsHUC10_high_type</i>	0.00074 [0.00060]	0.00018 [0.00108]	0.00034* [0.00019]	0.00020 [0.00040]
<i>#wellsHUC10_low_type</i>	0.00029 [0.00099]	0.00038 [0.00084]	0.00049 [0.00035]	0.00029 [0.00043]
Observations	4,213	4,194	6,642	6,636
R-squared	0.886	0.926	0.881	0.912
Chloride (µg/L)				
<i>#wellsHUC10_high_type</i>	0.00061*** [0.00022]	0.00071** [0.00031]	0.00064*** [0.00023]	0.00064* [0.00034]
<i>#wellsHUC10_low_type</i>	0.00046 [0.00033]	0.00077 [0.00051]	0.00016 [0.00024]	-0.00028 [0.00048]
Observations	9,843	9,792	27,628	27,613
R-squared	0.924	0.953	0.862	0.895
Barium (µg/L)				
<i>#wellsHUC10_high_type</i>	0.00050** [0.00024]	0.00037* [0.00021]	0.00020 [0.00016]	0.00023 [0.00015]
<i>#wellsHUC10_low_type</i>	0.00065 [0.00058]	0.00051 [0.00034]	0.00087** [0.00040]	0.00024 [0.00025]
Observations	3,491	3,450	13,578	13,571
R-squared	0.801	0.853	0.831	0.863
Strontium (µg/L)				
<i>#wellsHUC10_high_type</i>	0.00041*** [0.00014]	0.00040*** [0.00006]	0.00041*** [0.00013]	0.00036*** [0.00007]
<i>#wellsHUC10_low_type</i>	0.00049 [0.00038]	0.00065 [0.00048]	0.00050 [0.00036]	0.00045 [0.00043]
Observations	7,070	7,044	12,935	12,930
R-squared	0.934	0.961	0.967	0.977
Monitoring station FE	Yes	Yes	Yes	Yes
Weather controls	Yes	Yes	Yes	Yes
HUC4*Month*Year FE	Yes	No	Yes	No
HUC8*Month*Year FE	No	Yes	No	Yes

This table reports coefficients from OLS estimation of eq. S1 for treated watersheds (i.e., HUC10s with at least one active HF well over the sample period). We replace *#wellsHUC10* with two non-overlapping well count variables that count the cumulative number of wells spudded in watersheds that are “high type” or “low type”, respectively, indicating watersheds that have a higher or lower chance of detecting an effect, if it exists. See “Research Design” section of Supplementary Materials for details on coding watershed types. Columns (1) and (2) present results for Pennsylvania only and Columns (3) and (4) present results for all treated U.S. watersheds. Standard errors clustered by sub-basin (HUC8) are reported below the coefficients. *, **, *** indicate significance at the two-tailed 10%, 5%, and 1% levels, respectively.

Table S12. HF Wells and Water Quality using Time, Distance and Well Position (Upstream)

<i>Panel A – Descriptive Statistics for Restricted Sample of Surface Water Quality Measurements (µg/L) – all U.S. Watersheds</i>						
Bromide	N	Mean	P25	P50	P75	SD
Concentration	4,128	120.175	0.000	0.000	100.000	489.223
Ln(Concentration)	4,128	2.106	0.000	0.000	4.615	2.448
Chloride						
Concentration	13,986	44,783.720	6,010.000	15,200.000	42,270.000	124,748.800
Ln(Concentration)	13,986	9.545	8.701	9.629	10.652	1.909
Barium						
Concentration	6,598	58.391	29.000	42.000	62.000	109.965
Ln(Concentration)	6,598	3.686	3.401	3.761	4.143	0.970
Strontium						
Concentration	6,301	306.114	50.000	120.000	271.000	554.275
Ln(Concentration)	6,301	4.890	3.932	4.796	5.606	1.207
<i>Panel B – Descriptive Statistics for Restricted Sample of Surface Water Quality Measurements (µg/L) – Pennsylvania Watersheds</i>						
Bromide	N	Mean	P25	P50	P75	SD
Concentration	2,863	166.850	0.000	36.460	100.000	580.553
Ln(Concentration)	2,863	2.741	0.000	3.623	4.615	2.494
Chloride						
Concentration	5,905	24,936.290	5,500.000	13,350.000	25,800.000	42,650.030
Ln(Concentration)	5,905	9.423	8.613	9.499	10.158	1.203
Barium						
Concentration	2,333	42.174	28.000	38.000	48.000	24.696
Ln(Concentration)	2,333	3.653	3.367	3.664	3.892	0.452
Strontium						
Concentration	4,341	146.133	40.000	70.000	160.000	244.536
Ln(Concentration)	4,341	4.426	3.714	4.263	5.081	0.964

Panel C – Estimating Eq. S2 with Weighted Least Squares (Main Specification)

<i>Restricted Sample: Within 15km & Upstream & Close to Spud Date</i>								
	Bromide (µg/L)		Chloride (µg/L)		Barium (µg/L)		Strontium (µg/L)	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
<i>Post Spud</i>	-0.00464	-0.00577	0.02656**	0.02044**	0.01674	0.01585**	0.02103*	0.02014*
	[0.01295]	[0.00845]	[0.01265]	[0.01005]	[0.01915]	[0.00788]	[0.01227]	[0.01188]
Observations	109,362	190,514	207,121	394,798	77,936	181,866	157,056	184,035
R-squared	0.885	0.943	0.936	0.954	0.834	0.907	0.956	0.961
Weights	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Sample	PA	ALL	PA	ALL	PA	ALL	PA	ALL
Weather controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Monitoring-Well Pair FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
HUC8*Month*Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Panel D – Estimating Eq. S2 with Ordinary Least Squares

<i>Restricted Sample: Within 15km & Upstream & Close to Spud Date</i>								
	Bromide (µg/L)		Chloride (µg/L)		Barium (µg/L)		Strontium (µg/L)	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
<i>Post Spud</i>	-0.01497	-0.00456	0.01542*	0.01255*	0.00677	0.00851	0.01067	0.00867
	[0.01290]	[0.00975]	[0.00818]	[0.00682]	[0.01338]	[0.00635]	[0.00926]	[0.00854]
Observations	109,362	190,514	207,121	394,798	77,936	181,866	157,056	184,035
R-squared	0.917	0.941	0.948	0.959	0.868	0.930	0.961	0.973
Weights	No	No	No	No	No	No	No	No
Sample	PA	ALL	PA	ALL	PA	ALL	PA	ALL
Weather controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Monitoring-Well Pair FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
HUC8*Month*Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Panel E – Using “Never-Treated” Monitors as Controls in a Difference-in-Differences Analysis

<i>Restricted Sample of Treatment Observations (Within 15km & Upstream & Close to Spud Date) plus Controls</i>								
	Bromide (µg/L)		Chloride (µg/L)		Barium (µg/L)		Strontium (µg/L)	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
<i>Post Spud x Treatment_Sample</i>	-0.02812 [0.02142]	-0.01927 [0.01167]	0.03515** [0.01471]	0.03091** [0.01320]	0.01625 [0.02265]	0.01050 [0.01143]	0.01678** [0.00618]	0.01726*** [0.00615]
Observations	314,158	604,792	632,704	1,757,384	260,699	811,223	460,869	715,357
R-squared	0.860	0.941	0.909	0.891	0.740	0.867	0.940	0.962
Weights	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Sample	PA	ALL	PA	ALL	PA	ALL	PA	ALL
Weather controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Monitoring-Well Pair FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
HUC8*Month*Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Panel F – Bootstrapping p-Values using Randomization

	PA	ALL
Chloride		
Coefficient in Table S12, Panel C	0.02656	0.02044
Average coefficient over all random draws	0.00289	-0.00294
No. of randomized coefficients > coefficient in Table S12, Panel C	0	1
Bootstrapped p-value (given 1,000 random draws)	0.001	0.002
Barium		
Coefficient in Table S12, Panel C	0.01674	0.01585
Average coefficient over all random draws	-0.00461	-0.00230
No. of coefficients that exceed the coefficient in Table S12, Panel C	0	1
Bootstrapped p-value (given 1,000 random draws)	0.001	0.002
Strontium		
Coefficient in Table S12, Panel C	0.02103	0.02014
Average coefficient over all random draws	0.00197	0.00178
No. of coefficients that exceed the coefficient in Table S12, Panel C	0	13
Bootstrapped p-value (given 1,000 random draws)	0.001	0.014

For the analyses presented in this table, we pair all wells with all water monitoring stations in a given watershed (HUC10). The pairing of specific wells and monitors allows us to uniquely determine time between well spudding and measurement, well-monitor distance, and water flow direction at the monitor relative to the well. That is, for each well-monitor pair, we determine the distance between the well and the monitoring station (in km) as well as whether the well likely sits up- or downstream from the monitoring station. See “Research Design” section of Supplementary Materials for details on coding. We restrict the sample such that the well and the monitor are within 15km and that the well is assigned as likely upstream of the monitor. We only use water measurements taken before or up to 360 days after the spud date. We require at least two water measurements before and after a well spud date for each pair. Panel A reports surface water ion concentrations for the restricted sample of all U.S. watersheds (HUC10). Panel B reports surface water ion concentrations for the restricted sample of watersheds (HUC10) in Pennsylvania. Panel C table reports coefficients from the WLS estimation of Eq. S2 over treated watersheds (HUC10s) using the number of HUC8-month-year observations as weights. *Post Spud* is an indicator equal to one for all water measurements taken after the well in the pair is spudded, and zero otherwise. Panel D provides the same analysis using OLS. For Panel E, we expand the sample and add water measurements from “never-treated” water monitoring stations. We identify observations from treated and control monitors with a binary indicator, *Treatment_Sample*, which is one for observations entering the analysis in Table S12, Panel C and zero for observations from the control monitors and perform a difference-in-differences analysis. Panel E reports the coefficient on the interaction between *Post Spud* and *Treatment_Sample* from the WLS estimation of Eq. S2 over the augmented sample. See the description in the “Supplementary Analyses” for more details on how we define never-treated monitors. Panel F reports bootstrapped p-values for the coefficient estimates in Table S12, Panel C, using randomization with 1,000 repetitions. Specifically, we use the following formula: $p\text{-value} = \frac{[\text{No. of coefficients that exceed the respective coefficient in Table S12, Panel C}] + 1}{[\text{No. random draws} + 1]}$. See description in the “Supplementary Analyses” for more details on the randomization.

Reported below the coefficients are standard errors clustered by HUC8. *, **, *** indicate significance at the two-tailed 10%, 5%, and 1% levels, respectively.

Table S13. HF Wells and Water Quality – Controlling for Wastewater Facilities

	(1) PA	(2) PA	(3) PA	(4) PA
Bromide(µg/L)				
<i>#wellsHUC10</i>			0.00072 [0.00060]	0.00018 [0.00110]
<i>#wastewater facilities</i>	0.00061 [0.00047]	0.00014 [0.00075]	0.00053 [0.00047]	0.00010 [0.00087]
Observations	4,213	4,194	4,213	4,194
R-squared	0.886	0.926	0.886	0.926
Chloride (µg/L)				
<i>#wellsHUC10</i>			0.00060*** [0.00021]	0.00071** [0.00032]
<i>#wastewater facilities</i>	-0.00006 [0.00026]	0.00003 [0.00012]	-0.00009 [0.00022]	-0.00004 [0.00012]
Observations	9,843	9,792	9,843	9,792
R-squared	0.923	0.953	0.924	0.953
Barium (µg/L)				
<i>#wellsHUC10</i>			0.00051** [0.00025]	0.00038* [0.00021]
<i>#wastewater facilities</i>	-0.00004 [0.00011]	0.00010 [0.00017]	-0.00010 [0.00013]	0.00002 [0.00019]
Observations	3,491	3,450	3,491	3,450
R-squared	0.800	0.853	0.801	0.853
Strontium (µg/L)				
<i>#wellsHUC10</i>			0.00041*** [0.00015]	0.00037*** [0.00008]
<i>#wastewater facilities</i>	0.00020 [0.00028]	0.00056* [0.00031]	0.00013 [0.00026]	0.00046 [0.00029]
Observations	7,070	7,044	7,070	7,044
R-squared	0.934	0.961	0.934	0.961
Monitoring station FE	Yes	Yes	Yes	Yes
Weather controls	Yes	Yes	Yes	Yes
HUC4*Month*Year FE	Yes	No	Yes	No
HUC8*Month*Year FE	No	Yes	No	Yes

This table reports coefficients from the OLS estimation of Eq. S1 for treated HUC10s. We add the cumulative number of wastewater facilities accepting shale oil and gas waste within the HUC10 k up to day d (*#wastewater facilities*) as an additional control.

Reported below the coefficients are standard errors clustered by HUC8. *, **, *** indicate significance at the two-tailed 10%, 5%, and 1% levels, respectively.

Table S14. Conventional Wells and Water Quality

	(1) PA	(2) PA	(3) ALL	(4) ALL
Bromide (µg/L)				
<i>#wellsHUC10_conventional</i>	-0.00015 [0.00073]	-0.00006 [0.00089]	0.00110** [0.00051]	0.00069 [0.00063]
Observations	4,045	4,024	18,781	18,760
R-squared	0.883	0.922	0.790	0.860
Chloride (µg/L)				
<i>#wellsHUC10_conventional</i>	-0.00014 [0.00016]	-0.00010 [0.00024]	0.00011 [0.00010]	0.00009 [0.00015]
Observations	10,271	10,223	77,874	77,838
R-squared	0.928	0.955	0.870	0.900
Barium (µg/L)				
<i>#wellsHUC10_conventional</i>	0.00011 [0.00007]	-0.00020 [0.00013]	-0.00010 [0.00010]	-0.00004 [0.00011]
Observations	3,522	3,488	34,916	34,892
R-squared	0.809	0.854	0.839	0.866
Strontium (µg/L)				
<i>#wellsHUC10_conventional</i>	-0.00004 [0.00046]	-0.00018 [0.00046]	0.00008 [0.00010]	0.00007 [0.00011]
Observations	6,995	6,974	27,547	27,526
R-squared	0.933	0.959	0.960	0.971
Monitoring station FE	Yes	Yes	Yes	Yes
Weather controls	Yes	Yes	Yes	Yes
HUC4*Month*Year FE	Yes	No	Yes	No
HUC8*Month*Year FE	No	Yes	No	Yes
<p>This table reports coefficients from the OLS estimation of Eq. S1 for HUC10s with conventional wells. We use wells identified as conventional by the PA DEP/PACNR in Pennsylvania and those identified as vertical wells outside of PA. Instead of the HF well count, we use a well count for the cumulative number of conventional wells spudded in HUC10 k up to day d (<i>#wellsHUC10_conventional</i>). Columns (1) and (2) present results for Pennsylvania only and Columns (3) and (4) present results for all U.S. watersheds.</p> <p>Reported below the coefficients are standard errors clustered by HUC8. *, **, *** indicate significance at the two-tailed 10%, 5%, and 1% levels, respectively.</p>				

Table S15. HF Wells and Water Quality – Controlling for Known Spills

	(1) PA	(2) PA	(3) ALL	(4) ALL
Bromide (µg/L)				
(i) #wellsHUC10 & control for spills	0.00067 [0.00068]	-0.00025 [0.00133]	0.00032 [0.00022]	0.00006 [0.00052]
(ii) #wellsHUC10 & controls for big/small spills	0.00105* [0.00056]	0.00060 [0.00090]	0.00043** [0.00020]	0.00034 [0.00029]
Observations	4,213	4,194	6,642	6,636
Chloride (µg/L)				
(i) #wellsHUC10 & control for spills	0.00063*** [0.00021]	0.00074** [0.00031]	0.00050*** [0.00016]	0.00051** [0.00025]
(ii) #wellsHUC10 & controls for big/small spills	0.00066** [0.00026]	0.00076* [0.00038]	0.00044*** [0.00015]	0.00046** [0.00022]
Observations	9,843	9,792	27,628	27,613
Barium (µg/L)				
(i) #wellsHUC10 & control for spills	0.00050* [0.00028]	0.00034 [0.00028]	0.00021 [0.00018]	0.00020 [0.00016]
(ii) #wellsHUC10 & controls for big/small spills	0.00054* [0.00027]	0.00038 [0.00026]	0.00021 [0.00018]	0.00021 [0.00015]
Observations	3,491	3,450	13,578	13,571
Strontium (µg/L)				
(i) #wellsHUC10 & control for spills	0.00041** [0.00016]	0.00036*** [0.00007]	0.00041*** [0.00014]	0.00032*** [0.00009]
(ii) #wellsHUC10 & controls for big/small spills	0.00048*** [0.00014]	0.00041*** [0.00007]	0.00048*** [0.00013]	0.00036*** [0.00008]
Observations	7,070	7,044	12,935	12,930
#spill [>1-Year]	Yes	Yes	Yes	Yes
Monitoring station FE	Yes	Yes	Yes	Yes
Weather controls	Yes	Yes	Yes	Yes
HUC4*Month*Year FE	Yes	No	Yes	No
HUC8*Month*Year FE	No	Yes	No	Yes

This table reports coefficients from the OLS estimation of Eq. S1 for treated HUC10s. For each ion, we estimate two alternative versions of Eq. S1: (i) adding the cumulative number of spills within HUC10 k in the past year ($\#Spill[0;1\text{-Year}]$) as control variable, (ii) adding two control variables: the cumulative number of spills with an above the median spill amount within HUC10 k in the past year ($\#Spill_{hf}[0;1\text{-Year}]$), and the cumulative number of spills with a below the median spill impact (or with a missing spill impact) within HUC10 k in the past year ($\#Spill_{lf}[0;1\text{-Year}]$). These coefficients are not tabulated for brevity. To properly distinguish between pre- and post-spill observations, the models also include a long-run count variable for the cumulative number of spills in HUC10 k beyond one year, $\#Spill[>1\text{-Year}]$. (35) collect a large sample of unconventional O&G spills in Colorado, North Dakota, New Mexico, and Pennsylvania. We use the entire spill database and merge it with the water quality sample used in Table S4, resulting in a sample of 3,531 spills. Using all spills from (35) is likely conservative, but makes sense given the purpose is to control for known spills. The results for (i) are similar when we use only spills of saltwater/flowback and HF fluids. Reported below the coefficients are standard errors clustered by HUC8. *, **, *** indicate significance at the two-tailed 10%, 5%, and 1% levels, respectively.

References

53. A. W. Bartik, J. Currie, M. Greenstone, C. R. Knittel, The Local Economic and Welfare Consequences of Hydraulic Fracturing. *Am. Econ. J. App. Econ.* **11**, 105-155 (2019). doi: [10.1257/app.20170487](https://doi.org/10.1257/app.20170487)
54. S. S. Kaushal *et al.*, Freshwater salinization syndrome on a continental scale. *Proc. Natl. Acad. Sci. U.S.A.* **115**, E574-E583 (2018). doi: [10.1073/pnas.1711234115](https://doi.org/10.1073/pnas.1711234115)
55. J. Moore, R. M. Fanelli, A. J. Sekellick, High-Frequency Data Reveal Deicing Salts Drive Elevated Specific Conductance and Chloride along with Pervasive and Frequent Exceedances of the U.S. Environmental Protection Agency Aquatic Life Criteria for Chloride in Urban Streams. *Environ. Sci. Technol.* **54**, 778-789 (2020). doi: [10.1021/acs.est.9b04316](https://doi.org/10.1021/acs.est.9b04316)
56. K. J. Black, S. J. McCoy, J. G. Weber, Fracking and indoor radon: Spurious correlation or cause for concern? *J. Environ. Econ. Manage.* **96**, 255-273 (2019). doi: [10.1016/j.jeem.2019.06.002](https://doi.org/10.1016/j.jeem.2019.06.002)
57. L. Sun, S. Abraham, Estimating dynamic treatment effects in event studies with heterogeneous treatment effects. *J. Econom.* (2020). doi: [j.jeconom.2020.09.006](https://doi.org/j.jeconom.2020.09.006)
58. C. de Chaisemartin, X. D'Haultfœuille, Two-Way Fixed Effects Estimators with Heterogeneous Treatment Effects. *Am. Econ. Rev.* **110**, 2964-2996 (2020). doi: [10.1257/aer.20181169](https://doi.org/10.1257/aer.20181169)
59. A. Goodman-Bacon, Difference-in-Differences with Variation in Treatment Timing. *NBER Working Paper Series* **25018**, (2018; <https://www.nber.org/papers/w25018>).
60. A. Kondash, A. Vengosh, Water Footprint of Hydraulic Fracturing. *Environ. Sci. Technol. Lett.* **2**, 276-280 (2015). doi: [10.1021/acs.estlett.5b00211](https://doi.org/10.1021/acs.estlett.5b00211)
61. L. Hanson, S. Habicht, P. Daggupati, R. Srinivasan, P. Faeth, Modeling Changes to Streamflow, Sediment, and Nutrient Loading from Land Use Changes Due to Potential Natural Gas Development. *JAWRA J. Am. Water Resour. Assoc.* **53**, 1293-1312 (2017). doi: [10.1111/1752-1688.12588](https://doi.org/10.1111/1752-1688.12588)