Arsenic in Drinking Water and Infant Health

Laura Alcocer Quiñones
UC Davis

Wenjie Zhan
UC Davis

Mengfei Zhou
UC Davis

ASSA 2025 ASHEcon/HERO Health Economics Paper Session
San Francisco, CA
January 4, 2025

Arsenic is harmful

Arsenic (As): highly toxic; naturally occurring; no smell, color, or taste.

Drinking water: the major route of exposure (World Health Organization 2022).

Arsenic in drinking water $\mapsto$ worse birth outcomes and infant mortality (Bloom et al. 2014; Quansah et al. 2015).

EPA's new arsenic rule: 50 $\mu g/L \mapsto 10\ \mu g/L$

The maximum contaminant level (MCL) for arsenic since 1942: 50 $\mu g/L$.

Growing scientific evidence $\mapsto$ stricter MCL was needed.

On January 22, 2001, EPA: 50 $\mu g/L \mapsto 10\ \mu g/L$.

Full compliance: January 23, 2006.

Applies to community water systems; does not apply to private wells.

Existing literature

Scientific literature:

- Focuses on the relationship between arsenic levels and infant health
- Based on small prospective cohorts (Rahman et al. 2009; Fei et al. 2013); or
- Larger-scale but focuses on cross-sectional correlation (Shi et al. 2015; Almberg et al. 2017; C. M. Bulka et al. 2022).

Economic literature:

- Focuses on the relationship between arsenic levels and human capital outcomes in Bangladesh.
- Arsenic exposure $\mapsto$ lower cognitive ability (Asadullah and Chaudhury 2011; Pitt, Rosenzweig, and Hassan 2021), poorer labor outcomes (Carson, Koundouri, and Nauges 2011; Pitt, Rosenzweig, and Hassan 2021), and earlier marriage and childbirth (Chowdhury and Singh 2023).

This paper: how does EPA's new arsenic rule affect infant health?

This paper:

- Focuses on causal effects of EPA's new arsenic rule.
- Variations: (i) the rule change in 2001, (ii) estimated arsenic risk across counties, and (iii) unregulated private wells in 1990 across counties.
- Outcomes: birth weight, nearly all counties in contiguous U.S.

Main findings: the new rule seems to work!

EPA's new arsenic rule reduces the likelihood of detecting arsenic levels above 10 $\mu g/L$ in water systems.

EPA's new arsenic rule reduces the likelihood of low birth weight.

- The effect is primarily driven by counties with extremely high arsenic risk.
- High private well reliance mutes the effect.

Data

Arsenic testing data Map

- Six-Year Review 2 Contaminant Occurrence Data (SY2), EPA
- Tested arsenic concentration of water sample, 1998-2005.
- Water systems that the water sample was extracted from.
- We aggregate water-sample-level data to counties served (1043 counties).
- Reflects short-term arsenic levels, influenced by factors such as precipitation and measurement errors.

Data

Data on the probability of arsenic levels $> 10\ \mu g/L$ Map

- Estimated by Lombard et al. 2021.
- Boosted Regression Trees approach.
- Training data set: 20,450 private wells sampled by the U.S. Geological Survey (USGS) between 1970 and 2013.
- Predicting variables: 48 climatic, hydrological, and geological data from multiple sources.
- Model output:
  - Average precipitation 1981-2010 (16.48%);
  - P (5%), As (5.86%), Se (4.91%), Sb (4.88%), organic carbon (4.2%), and Be (4.06%) concentrations in the soil C horizon in 1974;
  - %Streamflow estimated to be baseflow in 2001 (5.03%);
  - Average annual groundwater recharge from 1981 to 2010 (4.04%).
- 1 km^2 grid $\mapsto$ population-weighted county mean.
- Reflects long-run risk of arsenic exposure.

Data

Birth outcomes

- Restricted-use Vital Statistics Natality Data.
- Infant health: birth weight (gram), likelihood of low birth weight (birth weight ¡ 2500 grams).
- Maternal characteristics: %black, %Hispanic, %age $\leq$ 24, %edu $\leq$ high school, %first time, and %smoke.
- Singletons.
- We collapse birth-level data to county(of maternal residence)-by-year-(of birth)-by-race cells

Data

Private well data

- The 1990 decennial census: last census to include data on private well use.
- Private well density = the number of private wells / areas. Map

County characteristics

- Baseline characteristics: 1995 Intercensal Population Estimates and 1994 County Data Book.
- County-by-year unemployment rates: Local Area Unemployment Statistics (LAUS), Bureau of Labor Statistics (BLS).

Method: cohort difference-in-difference

Comparisons: before V.S. after the new arsenic rule; counties with high v.s. low estimated $P_c(As > 10\mu g/L)$.

Preferred specification:

$$Y_{crt} = \beta P_c(As > 10\mu g/L) \times Post_t + \gamma_c + \mu_r + \lambda_t + Char_c \times t + Unemp_{ct} + MC_{crt} + \varepsilon_{crt},$$

- Y_{crt}: Mean outcome of the cell defined by county c, race r, and year t.
- $P_c(\text{As} > 10\mu g/L)$: county-level probability of arsenic exceeding 10 $\mu g/L$ estimated by Lombard et al. 2021.
- $Post_t$: born after 2001.
- γ_c: county FEs; μ_r: race FEs; λ_t: year FEs.
- $Char_c \times t$: county baseline characteristics $\times$ linear time trend.
- $Unemp_{ct}$: county-by-year unemployment rate.
- MC_{crt}: cell-average maternal characteristics.

Average arsenic levels in drinking water decline after the new rule

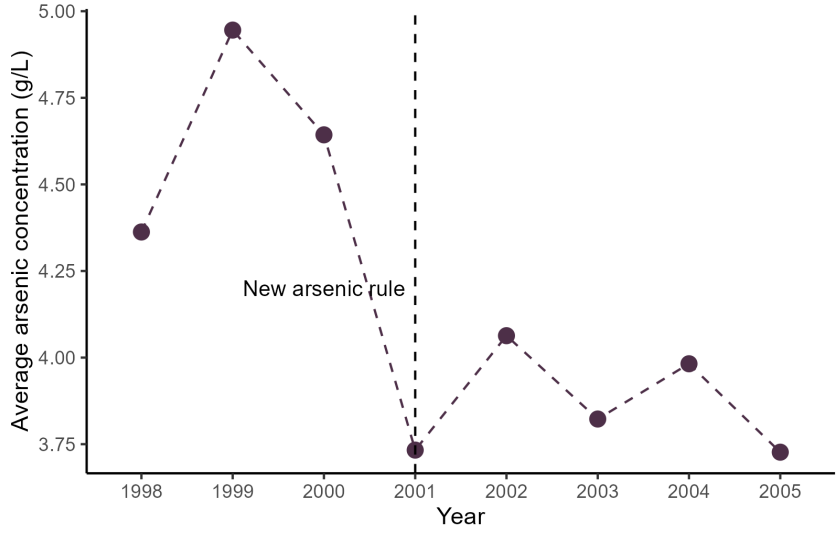


Figure: National trend of tested arsenic levels in drinking water (data: EPA's SY2)

The decline is larger in counties with a higher arsenic risk

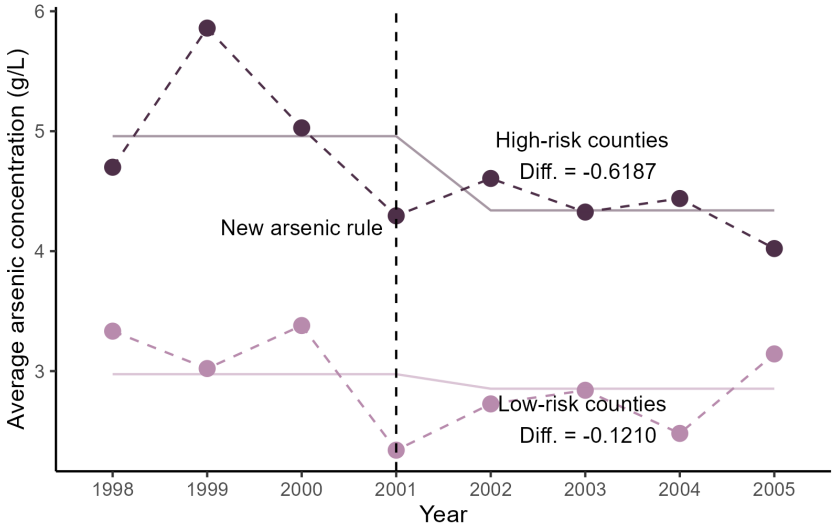


Figure: Higher arsenic risk: $\geq$ **median** $P_c(As > 10\mu g/L)$ (data: EPA's SY2)

New arsenic rule reduces the likelihood of detecting excess arsenic levels in drinking water

Event study plots

	Tested arsenic levels ($\mu g/L$)			Share of water systems with $P(As_{tested} > 10\mu g/L)$		
	(1)	(2)	(3)	(4)	(5)	(6)
$P_c(As > 10\mu g/L) \times Post_t$	0.0142			-0.2164		
	(0.0109)			(0.0725)***		
	⟨0.0128⟩			⟨0.0631⟩***		
$\mathbf{1}\{P_c(As > 10\mu g/L)$ at top 50$\} \times Post_t$		0.1914			-2.2991	
		(0.2427)			(1.2778)*	
		⟨0.3349⟩			⟨1.0954⟩**	
$\mathbf{1}\{P_c(As > 10\mu g/L)$ at top 25$\} \times Post_t$			0.0899			-3.3806
			(0.2980)			(1.7163)**
			⟨0.4214⟩			⟨1.7365⟩*
Observations	8,104	8,104	8,104	8,104	8,104	8,104
R^2	0.5692	0.5690	0.5689	0.6372	0.6309	0.6327
Dep. var. mean	3.42	3.42	3.42	3.16	3.16	3.16

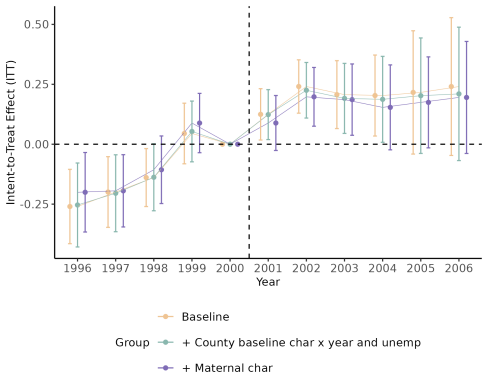
Notes: Standard errors are clustered at the county level (shown in parentheses) and the state level (in angle brackets). All regressions and the dependent variable mean are weighted by the number of births in each cell. ***, **, and * indicate that t-test are significant at the 1%, 5%, and 10% levels.

New arsenic rule reduces the likelihood of low birth weight

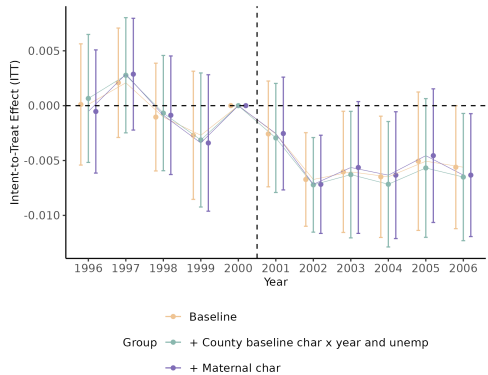
	Birth weight (grams)			Low birth weight (%)		
	(1)	(2)	(3)	(4)	(5)	(6)
$P_c(As > 10\mu g/L) \times Post_t$	0.3173	0.2875	0.2382	-0.0051	-0.0057	-0.0049
	(0.0644)***	(0.0703)***	(0.0714)***	(0.0015)***	(0.0016)***	(0.0014)***
	⟨0.0947⟩***	⟨0.0827⟩***	⟨0.0730⟩***	⟨0.0020⟩**	⟨0.0020⟩***	⟨0.0017⟩***
County FE	✓	✓	✓	✓	✓	✓
Year FE	✓	✓	✓	✓	✓	✓
Race FE	✓	✓	✓	✓	✓	✓
$Char_c \times$ year		✓	✓		✓	✓
$Unemp_{ct}$		✓	✓		✓	✓
MC_{crt}			✓			✓
Observations	75,414	70,263	66,085	75,414	70,263	66,085
R^2	0.9040	0.9067	0.9196	0.7980	0.8049	0.8205
Dep. var. mean	3334.25	3334.02	3332.68	6.13	6.13	6.17

Notes: Standard errors are clustered at the county level (shown in parentheses) and the state level (in angle brackets). All regressions and the dependent variable mean are weighted by the number of births in each cell. ***, **, and * indicate that t-test are significant at the 1%, 5%, and 10% levels.

New arsenic rule reduces the likelihood of low birth weight



(a) Birth weight (grams)



(b) Low birth weight (%)

Effects of the new arsenic rule are driven by a small share of counties with extremely high arsenic risk

Distribution of $P_c(As > 10\mu g/L)$ | Event study plots

$P_c(As > 10\mu g/L)$	Low birth weight (%) ≥ 10% (1)	≥ 15% (2)	≥ 20% (3)	≥ 25% (4)	≥ 30% (5)
$\mathbf{1}\{\text{High arsenic}\} \times Post_t$	-0.0421	-0.0737	-0.1002	-0.1841	-0.2278
	(0.0438)	(0.0468)	(0.0577)*	(0.0500)***	(0.0528)***
	⟨0.0527⟩	⟨0.0506⟩	⟨0.0593⟩*	⟨0.0529⟩***	⟨0.0451⟩***
Observations	66,085	66,085	66,085	66,085	66,085
R^2	0.8205	0.8205	0.8205	0.8205	0.8205
Dep. var. mean	6.17	6.17	6.17	6.17	6.17
Number of high-risk counties	464	266	188	141	107
Share of high-risk counties	15.25%	8.74%	6.18%	4.63%	3.52%
Average As ($\mu g/L$) in high-risk counties before 2001	5.75	7.33	8.36	9.27	10.45

Notes: Standard errors are clustered at the county level (shown in parentheses) and the state level (in angle brackets). All regressions and the dependent variable mean are weighted by the number of births in each cell. ***, **, and * indicate that t-test are significant at the 1%, 5%, and 10% levels.

High private well density mutes the effects of the new arsenic rule

Event study plots

Private well density in 1990	Low birth weight (%)			
	Low density		High density	
	$<$ 75th pct	$<$ 50th pct	$\geq$ 75th pct	$\geq$ 50th pct
	(1)	(2)	(3)	(4)
$P_c(As > 10\mu g/L) \times Post_t$	-0.0046	-0.0050	-0.0026	-0.0031
	(0.0015)***	(0.0016)***	(0.0037)	(0.0023)
	⟨0.0018⟩**	⟨0.0019⟩**	⟨0.0027⟩	⟨0.0018⟩*
Observations	48,865	33,120	17,220	32,965
R^2	0.8483	0.8724	0.5804	0.6676
Dep. var. mean	6.18	6.21	6.09	6.03

Notes: Standard errors are clustered at the county level (shown in parentheses) and the state level (in angle brackets). All regressions and the dependent variable mean are weighted by the number of births in each cell. ***, **, and * indicate that t-test are significant at the 1%, 5%, and 10% levels.

Robustness check

Unbalanced panel.

Sample selection: only keep cells with no fewer than 25 births.

Alternative exposure: 1 km^2 grid $\mapsto$ population-weighted county value at 90th percentile.

Conclusion

EPA's new arsenic rule reduces the likelihood of detecting arsenic levels above 10 $\mu g/L$ in water systems by 22 percentage points.

EPA's new arsenic rule reduces the likelihood of low birth weight by 0.0049 percentage points overall and by 0.1 percentage points in counties with a high arsenic risk (defined as $P_c(As > 10\mu g/L) \geq 20\%$).

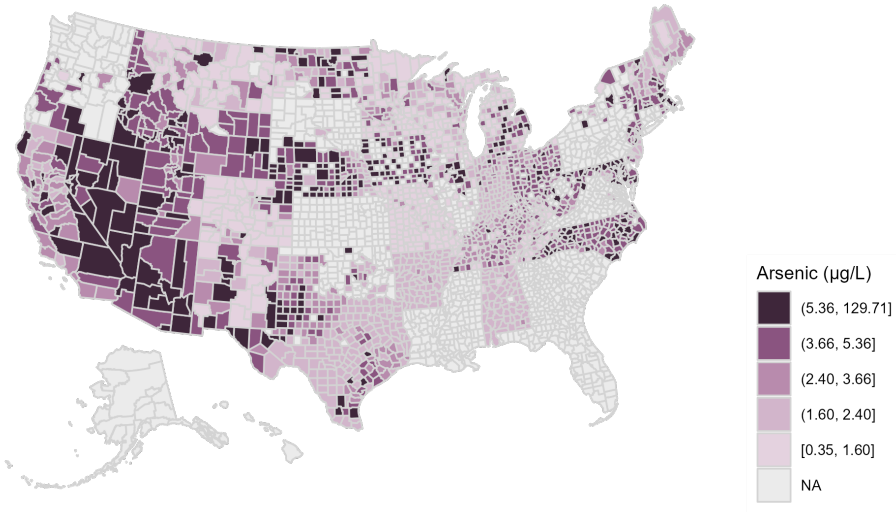
Policy implications:

- Regulating even a single toxic contaminant can lead to better health.
- The natural environment can contribute to health disparities, a factor that is often overlooked.

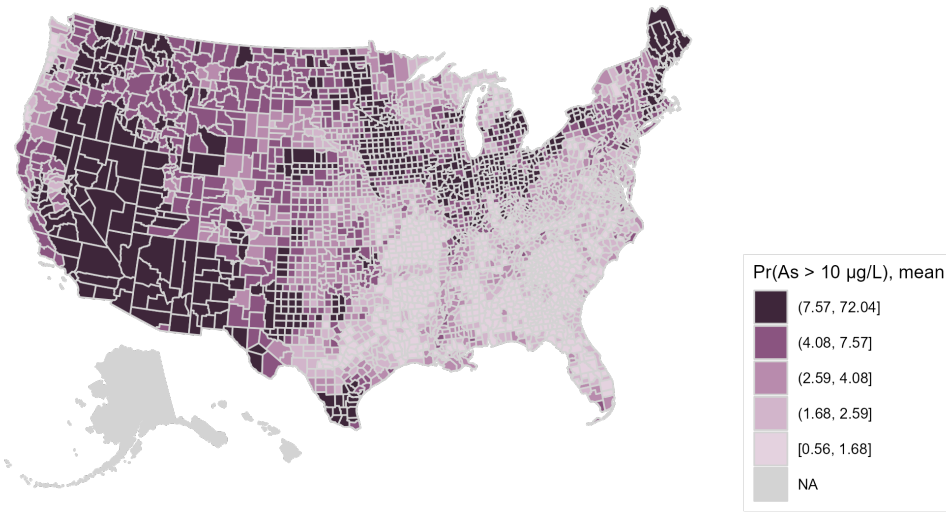
Thank you for your attention! Please reach out to me via wjzhan@ucdavis.edu for any comments/questions.

Arsenic testing data from EPA

Back

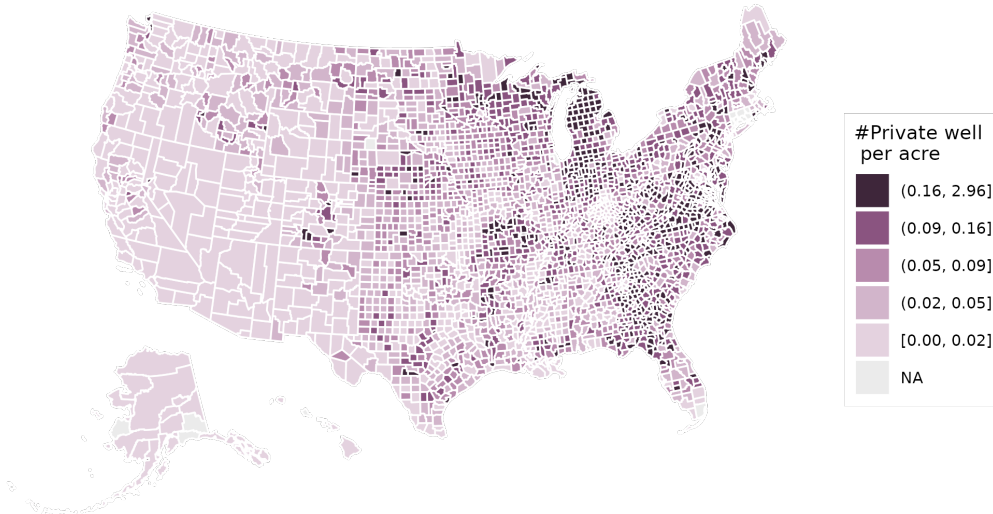


Probability of arsenic levels $> 10\ \mu g/L$ from Lombard et al. 2021 Back



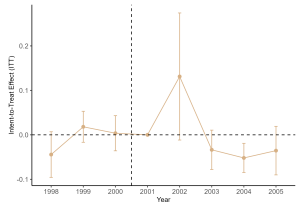
Private well density in 1990

Back

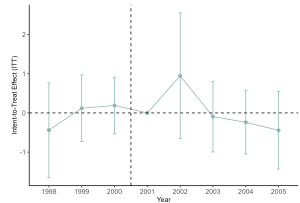


New arsenic rule reduces the likelihood of detecting excess arsenic levels in drinking water

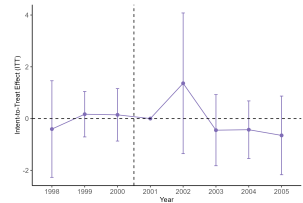
Back



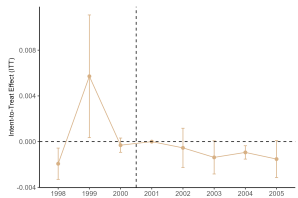
(a) Arsenic ($\mu g/L$), $P_c(As > 10\mu g/L)$ as exposure



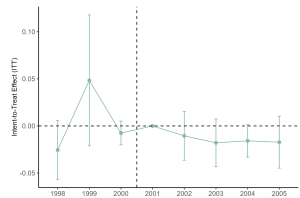
(b) Arsenic ($\mu g/L$), top 50 arsenic risk as treatment var.



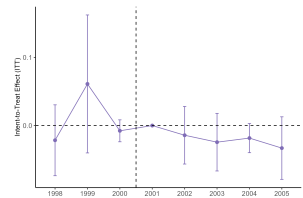
(c) Arsenic ($\mu g/L$), top 75 arsenic risk as treatment var.



(d) $P(As_{tested} > 10\mu g/L)$, $P_c(As > 10\mu g/L)$ as exposure

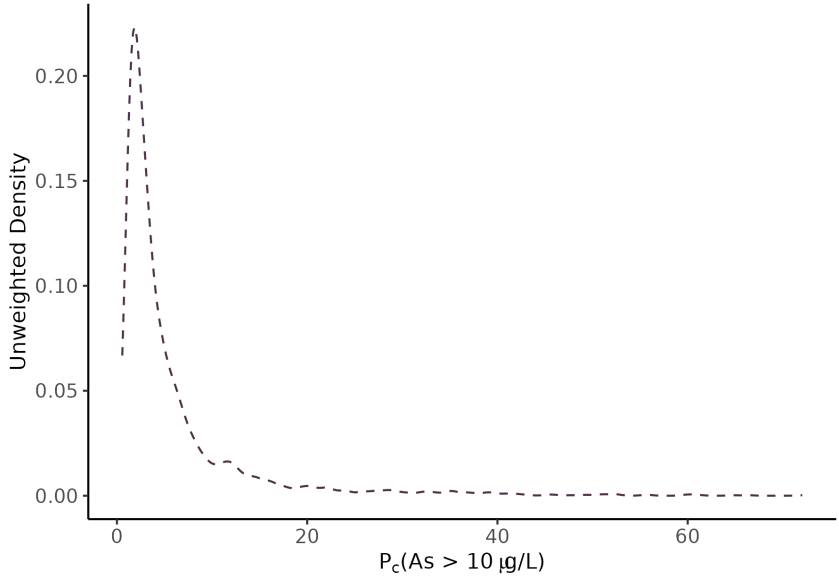


(e) $P(As_{tested} > 10\mu g/L)$, top 50 arsenic risk as treatment var.



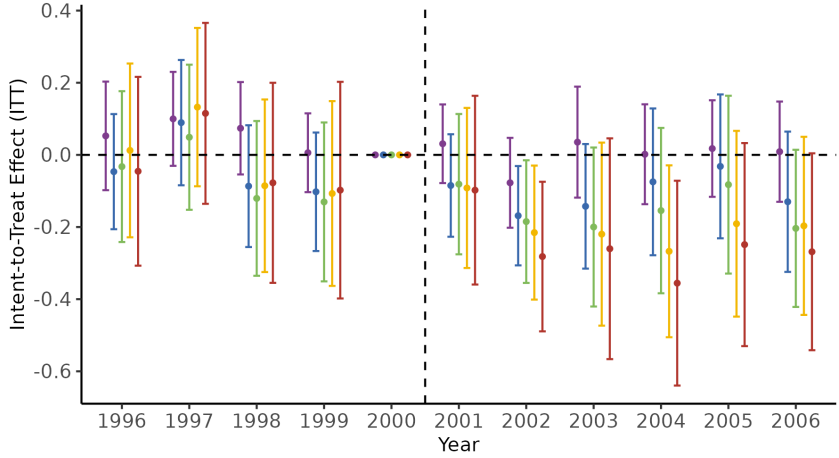
(f) $P(As_{tested} > 10\mu g/L)$, top 75 arsenic risk as treatment var.

Distribution of $P_c(As > 10\mu g/L)$ Back



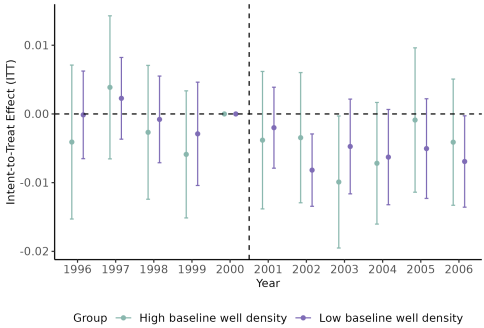
Effects of the new arsenic rule are driven by a small share of counties with extremely high arsenic risk

Back

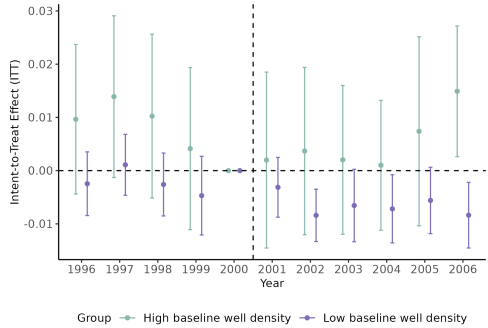


High private well density mutes the effects of the new arsenic rule

Back



(a) Low birth weight, $\geq$ 50th pct as high



(b) Low birth weight, $\geq$ 75th pct as high