Examining the impact Overturn of Roe v. Wade: Banning of Abortion on Infant Mortality Rates in the United States Using a Difference-in-Differences Approach*

The infant death rate increased by 0.26 deaths per 1,000 live births in states that became abortion-banned, compared to states where abortion remained legal.

Diana Shen

November 21, 2024

First sentence. Second sentence. Third sentence. Fourth sentence.

1 Introduction

The overturning of Roe v. Wade on June 24, 2022, through the Supreme Court's decision in Dobbs v. Jackson Women's Health Organization, ended nearly five decades of constitutional protection for abortion rights. This landmark ruling shifted authority over abortion regulations to individual states, creating a patchwork of laws nationwide. The decision reignited debates about reproductive rights and public health, highlighting disparities in access to abortion services and their effects on health outcomes, especially in states with restrictive policies.

Banning abortion usually has been linked to a rise in maternal mortality rates, with Texas experiencing a 56% increase from 2019 to 2022 (Chuck 2024). However, restrictive abortion laws can also impact infant mortality, as they often lead to more unintended or high-risk pregnancies among disadvantaged populations with limited access to healthcare. These pregnancies are associated with higher rates of preterm births, low birth weights, and delivery complications, all of which elevate infant mortality rates. Even before Roe v. Wade was overturned, researchers highlighted the impacts of restrictive abortion laws on mortality rates. Burdick et al. (2024) found that states with 11–12 restrictive laws during 2014–2018 had a 16% higher infant mortality rate, particularly in counties with higher proportions of Black populations,

*Code and data are available at: [https://github.com/DianaShen1224/Relationship-between-infant-mortality-rate-and-prohibited-abortion).

inadequate prenatal care, and maternal smoking. Similarly, (**Harper2023?**) showed that restrictive laws are linked to higher maternal, fetal, and infant mortality, states with moderate abortion legislation had significantly lower maternal mortality (25.79 per 100,000 live births) and infant mortality (20.56 per 1,000 live births) compared to restrictive states, underscoring the health benefits of less restrictive policies.

The estimand of our analysis is the change in the difference in infant death rates between states that banned abortion and states where abortion remained legal after June 24, 2022. The object of the estimation is to estimate the change in the difference in the infant death rate between the treatment and control groups based on observational data in the United States.

Instead of examining the impact of abortion laws over broader time periods like prior research, our study focuses on the effect of abortion bans on infant mortality rates in the United States specifically from 2021 to 2022, with an emphasis on changes following the 2022 Dobbs decision. Using a Difference-in-Differences (DID) approach, we assess the interaction between abortion legality in states and the timing of the injunction implemented to understand its influence on infant health outcomes. The study utilizes observational data from the Centers for Disease Control and Prevention's WONDER database, capturing infant mortality trends across states with varying levels of abortion restrictiveness. Key predictors include the mother's state of residence, race, age, and the year of infant death, enabling a nuanced examination of the effects of restrictive abortion laws on public health.

Our Results show that the overturning of Roe v. Wade and the implementation of abortion bans are linked to a rise in infant mortality rates in states where abortion became illegal. Through a Difference-in-Differences (DID) approach, we estimate that states enforcing abortion bans after the injunction saw an increase of 0.26 deaths per 1,000 live births compared to states where abortion remained legal. Furthermore, we find that maternal age and race are key determinants of infant mortality, with older maternal age groups associated with lower mortality rates and Black mothers experiencing significantly higher rates. These results emphasize the public health consequences of restrictive abortion laws and the urgent need for policies that mitigate these disparities.

This study is important as we highlight the public health implications of abortion bans, particularly their role in increasing infant mortality rates. By focusing on the 2022 Dobbs decision, we provide timely evidence on how restrictive abortion policies disproportionately affect vulnerable populations, emphasizing the need for data-driven policies to mitigate these adverse outcomes.

The structure of this paper is organized as follows: Section 2 provides details on the data sources and variables used in our analysis. **?@sec-model** explains the modeling approach, including the assumptions and specifications of our linear regression framework. In **?@sec-result**, we present our findings, emphasizing the key predictors of Harris's support. Finally, **?@sec-discussion** explores the implications of our results and suggests potential directions for future research. **?@sec-a** provides external data detail, **?@sec-b** provides model detail,

?@sec-c provides an datasheet, and **?@sec-d** provides an detailed introduction of surveys, sampling and observational data in our paper.

2 Data

2.1 Overview

We conduct our infant death rate observational data analysis using the R programming language (R Core Team 2023). Our dataset, obtained from CDC Wonder(Disease Control and Prevention 2024), based on data as of December 2022, provides a detailed overview of the Infant Death Rate across the USA. CDC WONDER (Wide-ranging ONline Data for Epidemiologic Research) is an online system providing public health professionals and the public with access to a wide range of the Centers for Disease Control and Prevention (CDC) epidemiologic data and resources. Following the guidelines outlined in Alexander (2023), we analyze the impact of abortion injunctions on infant death rates using the Difference-in-Difference method. This approach utilizes the interaction between the legality of abortion in the mother's state of residence and the timing of the infant's death as the predictor variable. Additionally, we include control variables such as infant year of death, the mother's state of residence, race, and age to account for potential confounding factors.

2.2 Measurement

The process of translating real-world events into our dataset requires a systematic approach to measurement and data gathering. In this research, we analyze infant mortality rates in the United States to understand the impact of abortion policies and demographic factors. The CDC's Linked Birth/Infant Death Records database provides comprehensive data, including maternal and infant characteristics, to support this study. Variables such as maternal age, race, education, birth weight, gestational age, and prenatal care are included to ensure a detailed examination of factors influencing infant health outcomes.

The data, sourced from birth and death certificates, undergoes a thorough cleaning and validation process to address inconsistencies and missing information, ensuring accuracy. Infant mortality rates are calculated per 1,000 live births, with additional indicators such as reliability labels for rates based on fewer than 20 deaths. To further enhance data accuracy, weights are applied to adjust for unlinked infant death records, ensuring that all deaths—linked and unlinked—are proportionally represented. Weights are calculated for each state and infant age-at-death cohort (<7 days, 7-27 days, and 28 days to 1 year) using the formula: $\text{Weight} = \frac{\text{Linked deaths}+\text{Unlinked deaths}}{\text{Linked deaths}}$ This adjustment means areas with complete linkages have a weight of 1.0, while areas with incomplete linkages are adjusted proportionally. Additionally, imputation techniques are employed to handle missing birthweight data, assigning values based on gestational period, race, sex, and plurality, further reducing potential bias in the analysis.

Once cleaned, the data is stratified by maternal and infant attributes, providing insights into trends and disparities across demographic groups. Statistical models are employed to estimate associations between state-level abortion policies and infant mortality, controlling for confounders such as maternal characteristics. This structured approach transforms complex datasets into actionable insights, shedding light on the health implications of varying abortion laws across the U.S. By leveraging comprehensive datasets and robust statistical techniques, this study offers a nuanced understanding of the intersection between policy, demographics, and infant health outcomes, informing public health strategies and legislative decisions.

2.3 Outcome variables

2.3.1 Death Rate

The `death rate` refers to the number of deaths of infants under one year of age per 1,000 live births in a specified population during a given time period. It serves as a key indicator of a population's overall health and access to healthcare. Rates labeled as "Unreliable" occur when there are fewer than 20 deaths in the numerator, as these figures do not meet the National Center for Health Statistics (NCHS) standards for reliability or precision. After cleaning the dataset, the "Unreliable" label is converted into a separate variable named `reliable`.

Figure 1 illustrates the distribution of infant death rates (per 1,000 live births) based on the available data. The majority of infant death rates cluster around lower values, with a peak observed between 5 and 10 deaths per 1,000 infants. As the death rate increases, the frequency of occurrences decreases, indicating that higher infant death rates are relatively rare. There are few instances where the rate exceeds 20 deaths per 1,000 infants, suggesting these are outliers. Overall, the distribution is right-skewed, highlighting the predominance of lower infant death rates in the dataset.

2.4 Predictor variables

2.4.1 Interaction Term: Abortion Legality and After Injunction

Abortion legality (named `abortion_illegal` in the dataset) is a newly created variable in our dataset, which refers to the legal status of abortion within a given state. It is represented as a binary variable, where 1 indicates that abortion is illegal in the state, and 0 indicates that abortion is legal. This classification reflects the policies and laws governing abortion access in each state.

Another created variable `after_injunction` indicates whether the date of infants' death falls after the implementation of the abortion injunctions following the overturn of Roe v. Wade

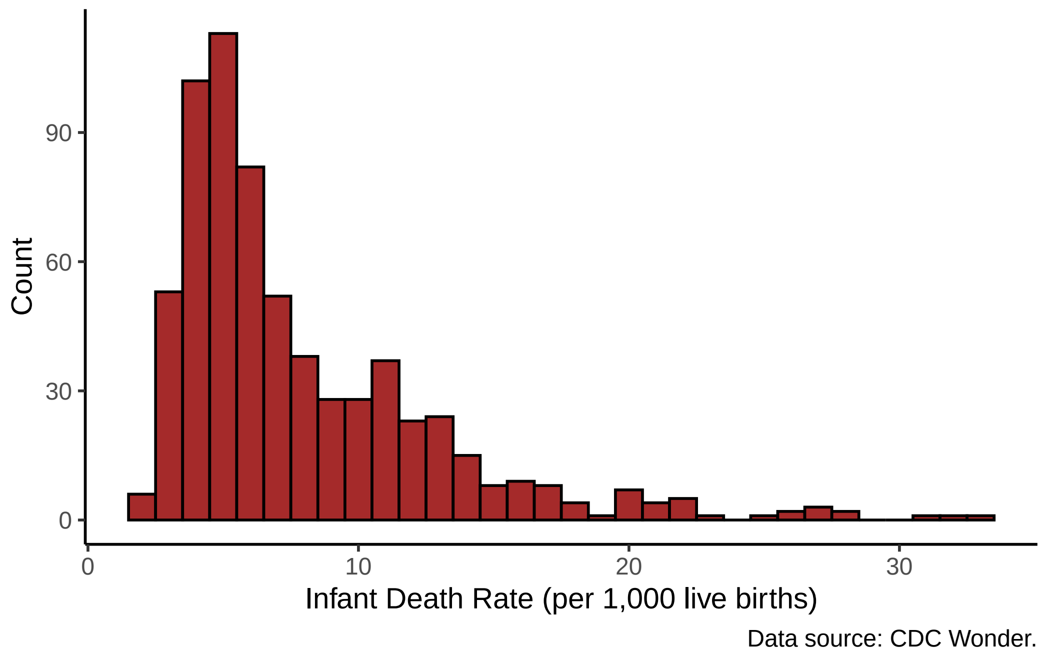


Figure 1: The distribution of infant death rate (per 1,000 live births), with a peak observed between 5 and 10 deaths per 1,000 infants.The distribution is right-skewed, highlighting the predominance of lower infant death rates in the dataset.

in June 2022. It is coded as a binary variable, where 1 represents dates after June 2022, signifying the post-injunction period, and 0 represents dates before June 2022, indicating the pre-injunction period.

To evaluate the impact of abortion injunctions on infant death rates, we utilize the Difference-in-Difference (DID) method. As part of this approach, we construct an interaction term defined as the product of `abortion_illegal` and `after_injunction`. The resulting interaction term equals 1 when abortion is illegal and the time is after the injunction, and 0 otherwise. This term is included in the analysis to evaluate the joint impact of these factors on infant death rates.

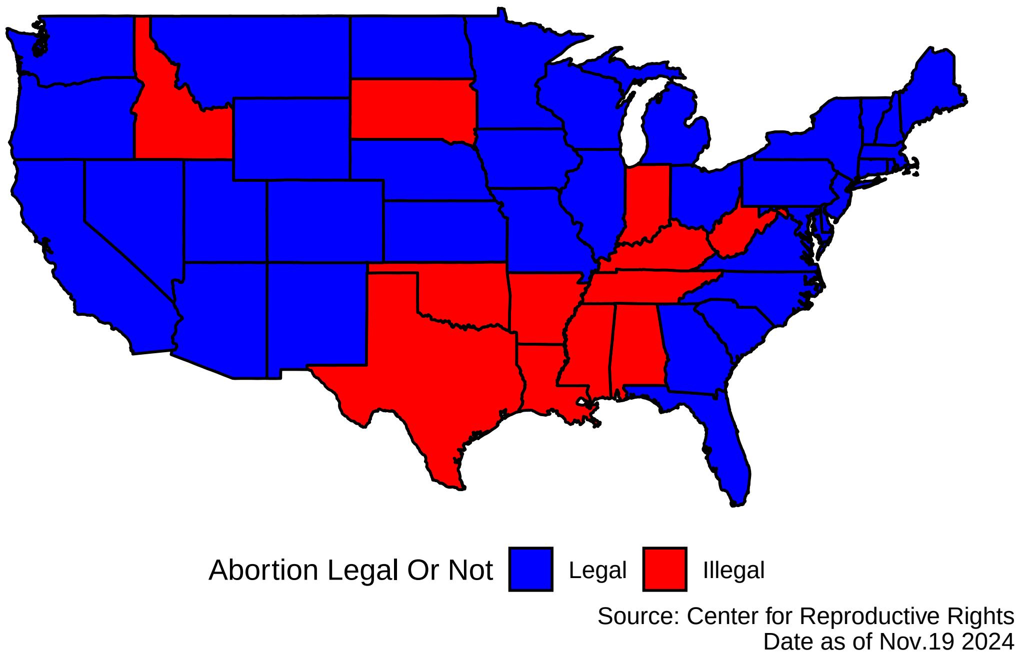


Figure 2: Abortion Legality by State. As of November 2024, abortion is illegal in Alabama, Arkansas, Idaho, Indiana, Kentucky, Louisiana, Mississippi, Missouri, North Dakota, Oklahoma, South Dakota, Tennessee, Texas, West Virginia, and Wisconsin.

Figure 2 illustrates the legality of abortion across U.S. states. Based on Reproductive Rights (2024), as of November 2024, abortion has been illegal in Alabama, Arkansas, Idaho, Indiana, Kentucky, Louisiana, Mississippi, Missouri, North Dakota, Oklahoma, South Dakota, Tennessee, Texas, West Virginia, and Wisconsin following the June 2022 decision.

Figure 3 illustrates trends in infant death rates (per 1,000 live births) over time in states where abortion is illegal versus legal. In states with abortion bans, infant mortality shows a slight increase after June 2022, marked by the dashed line representing the overturn of Roe v. Wade. In contrast, states where abortion remains legal exhibit relatively stable trends. This suggests a potential link between abortion bans and rising infant mortality rates, warranting further

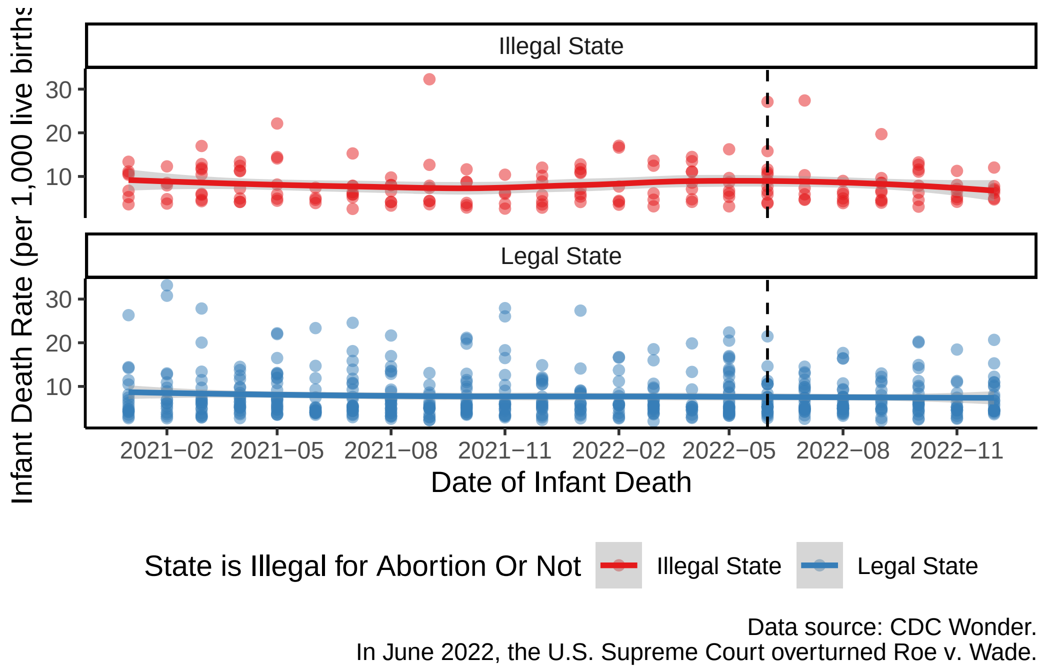


Figure 3: Infant death rates (per 1,000 live births) over time, with a slight increase after June 2022 in states where abortion is illegal, while rates remain stable in states where abortion is legal.

investigation.

Table 1: Average Infant Death Rate (per 1,000 live births) by Interaction Between Abortion Legality and Injunction Timing. States with "Illegal Abortion and After Injunction" (denoted as '1') have a higher average infant death rate compared to states with "Legal Abortion and Before/After Injunction or Illegal Abortion and Before Injunction" (denoted as '0').

illegal_after_injunction	Average Infant Death Rate (per 1,000 live births)
0	7.9
1	8.3

Note:
Data Source: CDC Wonder

Table 1 highlights average infant death rates (per 1,000 live births) based on abortion legality and injunction timing. States categorized as '1', where abortion became illegal after the injunction, show a higher average infant death rate of 8.3 compared to 7.9 in '0' states, where abortion remained legal or abortion used to be legal before the injunction. This suggests that restrictive abortion policies and their enforcement may contribute to higher infant mortality, potentially due to increased unintended or high-risk pregnancies, reduced access to maternal healthcare, and exacerbated social or economic disparities.

2.5 Control Variables

2.5.1 Year of Death

The Year of Death variable represents the calendar year in which the infant death occurred, specifically focusing on the years 2021 and 2022 in this study. This variable is essential for analyzing temporal trends and identifying patterns or shifts in infant mortality rates over this critical period. By controlling the year of death, we can closely examine the impact of significant policy changes, such as the overturning of Roe v. Wade in 2022, as well as other temporal factors on infant health outcomes.

2.5.2 State

The state of the mother's legal residence at the time of birth. This data is categorized by the mother's state of residence and excludes territories.

Figure 4 is a choropleth map visualizes the average infant death rate per 1,000 live births across U.S. states, with data categorized by color intensity. States with higher infant death rates are shown in lighter shades (yellow), while states with lower rates are represented in

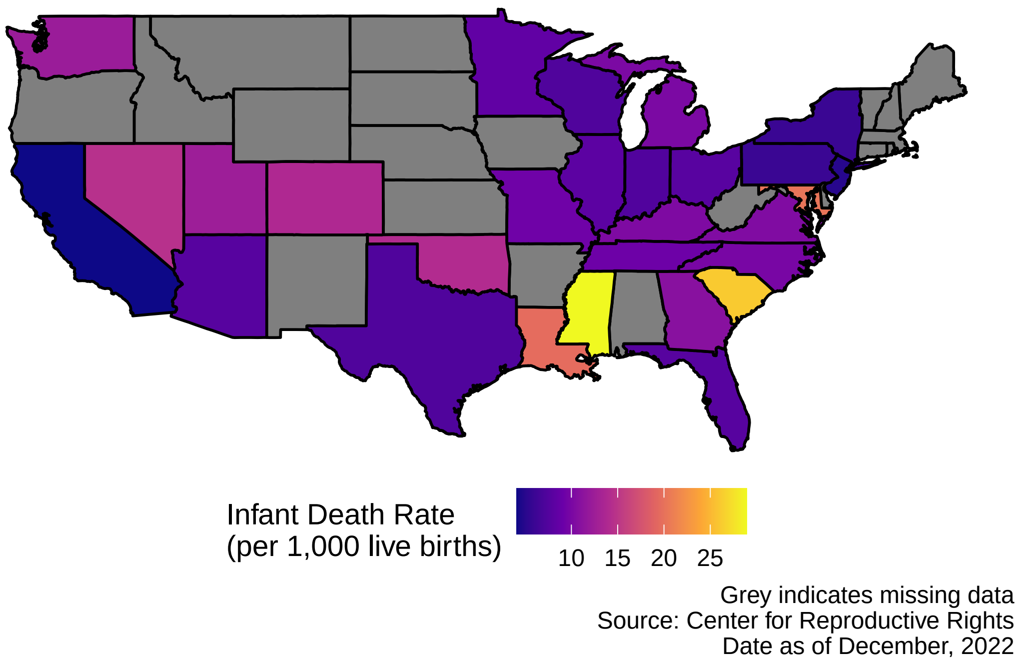


Figure 4: Infant Death Rates Across U.S. States (per 1,000 Live Births), December 2022. Infant Mortality Rate is much higher iin Southeastern. Alabama and Mississippi show the highest rates, while California and New York exhibit lower rates.

darker shades (purple). States with missing data are indicated in grey. The map highlights disparities in infant mortality rates, with some states, particularly in the southeastern U.S., exhibiting notably higher rates compared to others. States like Alabama and Mississippi show the highest rates, while states such as California and New York exhibit lower rates.

2.5.3 Age of Mother

The Age of Mother (`age_of_mother`) refers to the age group of the mother at the time of childbirth.In our observational data, age is categorized into the brackets 15–19 years, 20–24 years, 25–29 years, and 30–34 years in the raw data, in compliance with privacy requirements.

Table 2 displays the average infant death rate (per 1,000 live births) by the age group of mothers. It shows a clear trend: younger mothers, particularly those aged 15–19 and 20–24, have higher infant death rates (9.0 and 10.0, respectively), while older mothers in the 25–29 and 30–34 age groups have lower rates (7.4 and 6.6, respectively). This suggests that maternal age may be a factor influencing infant mortality, with younger mothers experiencing higher risks.

Table 2: Average Infant Death Rate (per 1,000 live births) of different age group of Mother. It shows that infant death rates decrease with increasing maternal age, with the highest rate (10.0) among mothers aged 20–24 and the lowest (6.6) among those aged 30–34.

Age of Mother	Average Infant Death Rate (per 1,000 live births)
15-19	9.0
20-24	10.0
25-29	7.4
30-34	6.6

Note:

Data Source: CDC Wonder

2.5.4 Mother's Single Race

The self-identified race of the mother, recorded as a single race category. In this dataset, the available categories are: Asian, Black or African American, and White. This classification facilitates the analysis of maternal and infant health outcomes across these racial groups.

Table 3: Average Infant Death Rate (per 1,000 live births) of different age group of Mother. It shows that infant death rates decrease with increasing maternal age, with the highest rate (10.0) among mothers aged 20–24 and the lowest (6.6) among those aged 30–34.

Mother's Single Race	Average Infant Death Rate (per 1,000 live births)

Asian	5.4
Black or African American	14.5
White	5.8

Note:
Data Source: CDC Wonder

3 Model

For our analysis, we use a Bayesian Difference-in-Differences (DID) model to quantify the impact of abortion injunctions on the infant death rate across states in the USA. This method compares changes in infant death rates over time between a treatment group (states where abortion injunctions were implemented) and a control group (states where abortion remains legal), providing a robust framework for estimating causal effects.

The model focuses on the infant death rate per 1,000 live births as the dependent variable and includes the following independent variables:
- Treatment Group Indicator `abortion_illegal`: A binary variable indicating the group: 1 if abortion is illegal in the state. 0 if abortion remains legal.
- Time Indicator `after_injunction`: A binary variable indicating the time period: 0 for the period before the injunction implemented. 1 for the period after the injunction implemented.
- Interaction Term`after_injunction` $\times$ `abortion_illegal`: The interaction term captures the additional change in the outcome (e.g., percentage of support, behavior, or health-related metrics) for regions where abortion is illegal after the injunction implemented, beyond any changes observed in regions where abortion remains legal.

To account for potential regional variations in infant death rates, we include state as a random effect in the model. This allows us to control for unobserved heterogeneity across states that might introduce bias. Additionally, we include age_of_mother, mothers_single_race, and year_of_death as control variables to adjust for demographic and temporal factors that could influence infant mortality rates, ensuring that the estimated effects are not confounded by these characteristics.

By isolating the impact of the abortion injunction through this interaction term, the model identifies changes in infant death rates that can be attributed to the policy intervention. Additional details and diagnostics supporting the model's implementation are provided in Appendix- B.

3.1 Difference-In-Difference Approach (DID)

DID is a causal inference method that estimates the impact of an intervention by comparing changes in outcomes over time between a treatment group and a control group. By assuming parallel trends in the absence of the intervention, the DID framework isolates the intervention's effect through the interaction between time periods (pre- and post-intervention) and group membership. This makes it a robust method for evaluating policy changes and other significant events.

3.2 Model set-up

In our analysis, we employ two models. The first model includes only the predictor variables and the random effect for state. The second model builds on the first by incorporating additional control variables to account for potential confounding factors.

Define y_i as as the infant death rate per 1,000 live births.

(1)

(2)

$$\mu_i = \beta_0 + \beta_1 \cdot \text{After Injunction}_i + \beta_2 \cdot \text{Abortion Illegal}_i + \beta_3 \cdot \text{After Injunction}_i \cdot \text{Abortion Illegal}_i + \beta_4 \cdot \text{Year of}$$

(3)

(4)

(5)

(6)

(7)

(8)

(9)

(10)

(11)

(12)

The models are run in R (R Core Team 2023) using the rstanarm package of Goodrich et al. (2022). The default priors from rstanarm are used for both GLM Bayesian models. The intercept normal prior with $= 50$ and $= 10$ reflects the central tendencies of Harris and Trump's polling percentages, influenced by prior knowledge and the predictors use a normal prior with $= 0$ and $= 5$ (Goodrich et al. 2022). We run the model in R (R Core Team 2023) using the `rstanarm` package of Goodrich et al. (2022). We use the default priors from `rstanarm`.

3.2.1 Model justification

The Bayesian framework enhances the DID approach by incorporating prior information and quantifying uncertainty in a probabilistic manner. Using Markov Chain Monte Carlo (MCMC) sampling, the model estimates the posterior distributions of the parameters, offering flexibility in handling complex relationships and uncertainties in the data.

4 Results

Our results are summarized in Table 4.

5 Discussion

5.1 First discussion point

If my paper were 10 pages, then should be be at least 2.5 pages. The discussion is a chance to show off what you know and what you learnt from all this.

5.2 Second discussion point

Please don't use these as sub-heading labels - change them to be what your point actually is.

5.3 Third discussion point

5.4 Weaknesses and next steps

Weaknesses and next steps should also be included.

Table 4: Explanatory models of flight time based on wing width and wing length

	(1)	(2)
(Intercept)	9.681	−649.462
after_injunction1	−0.540	−0.951
abortion_illegal1	3.572	1.371
after_injunction1 × abortion_illegal1	0.128	0.268
year_of_death		0.329
age_of_mother20-24		−3.573
age_of_mother25-29		−5.736
age_of_mother30-34		−6.164
mothers_single_raceBlack or African American		6.555
mothers_single_raceWhite		−2.613
Num.Obs.	659	659
R2	0.256	0.793
R2 Adj.	0.209	0.778
R2 Marg.	0.090	0.632
ICC	0.6	0.7
Log.Lik.	−1885.283	−1460.386
ELPD	−1911.7	−1496.6
ELPD s.e.	33.9	45.9
LOOIC	3823.4	2993.3
LOOIC s.e.	67.8	91.8
WAIC	3820.7	2989.5
RMSE	4.21	2.26
r2.adjusted.marginal	0.209125275079672	0.778197133826244

Note. All models include a control for state and pollsters. The reference level of the state is Alaska, the reference level of population is likely voter, and the reference level of pollsters is Angus Reid.

Appendix

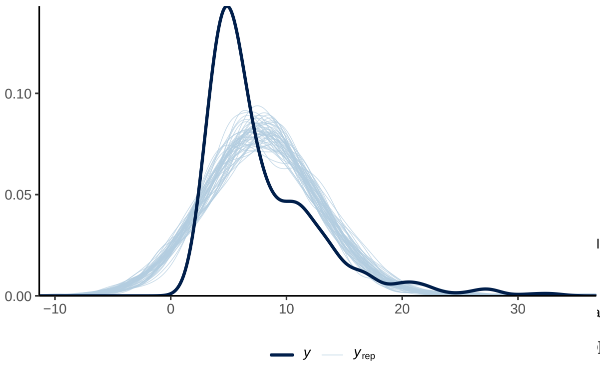
A Additional data details

B Model details

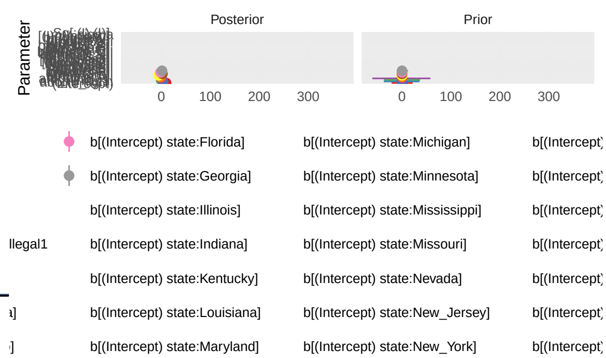
B.1 Posterior predictive check

In Figure 5a we implement a posterior predictive check. This shows...

In Figure 5b we compare the posterior with the prior. This shows...



(a) Posterior prediction check

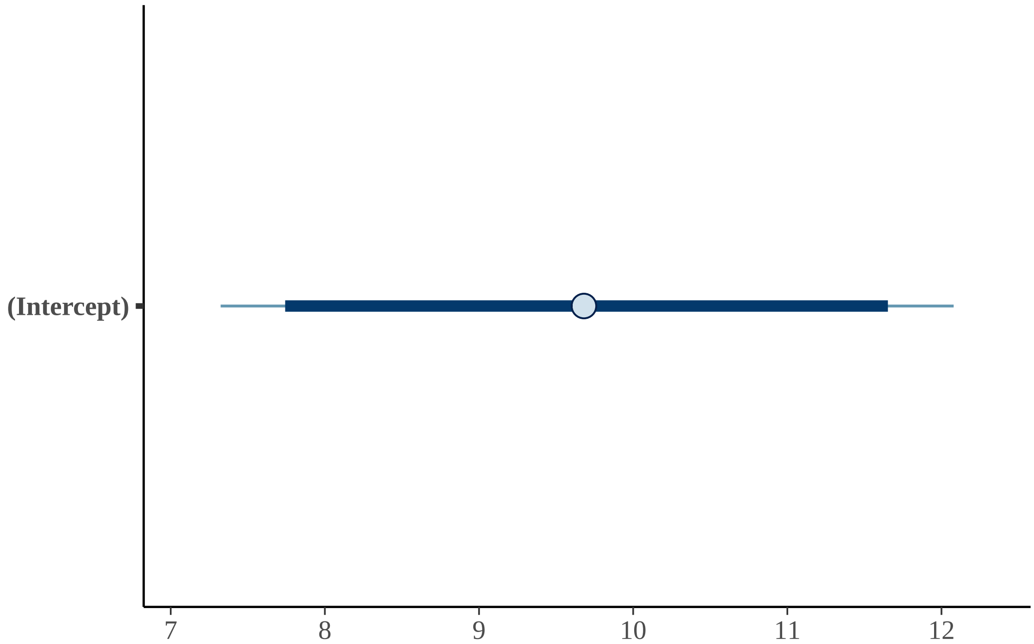


(b) Comparing the posterior with the prior

Figure 5: Examining how the model fits, and is affected by, the data

```
plot(first_model, pars = "(Intercept)", prob = 0.95)
```

```
Warning: `prob_outer` (0.9) is less than `prob` (0.95)
... Swapping the values of `prob_outer` and `prob`
```




```r
prior_summary(second_model)
```

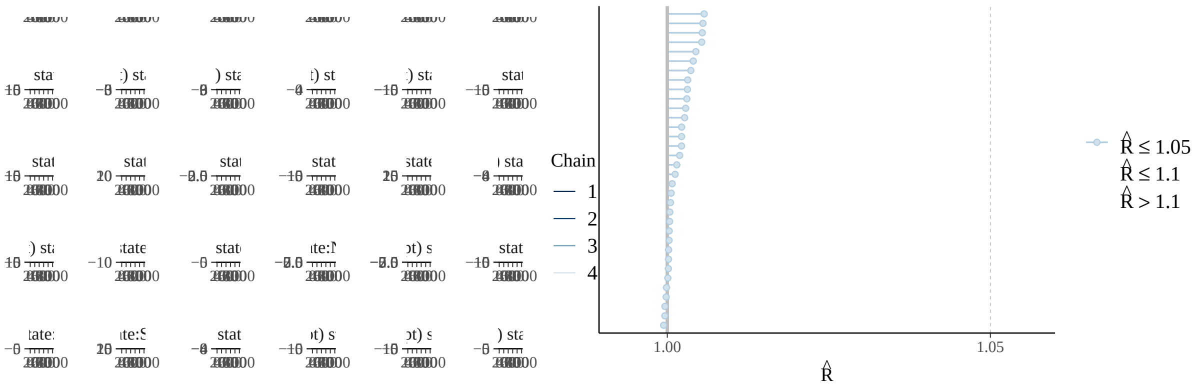
```
Priors for model 'second_model'
------
Intercept (after predictors centered)
  Specified prior:
    ~ normal(location = 0, scale = 2)
  Adjusted prior:
    ~ normal(location = 0, scale = 9.9)

Coefficients
  Specified prior:
    ~ normal(location = [0,0,0,...], scale = [2,2,2,...])
  Adjusted prior:
    ~ normal(location = [0,0,0,...], scale = [21.47,22.74,19.82,...])

Auxiliary (sigma)
  Specified prior:
    ~ exponential(rate = 1)
  Adjusted prior:
    ~ exponential(rate = 0.2)

Covariance
```

```
 ~ decov(reg. = 1, conc. = 1, shape = 1, scale = 1)
------
See help('prior_summary.stanreg') for more details
```

B.2 Diagnostics

Figure 6a is a trace plot. It shows... This suggests...

Figure 6b is a Rhat plot. It shows... This suggests...

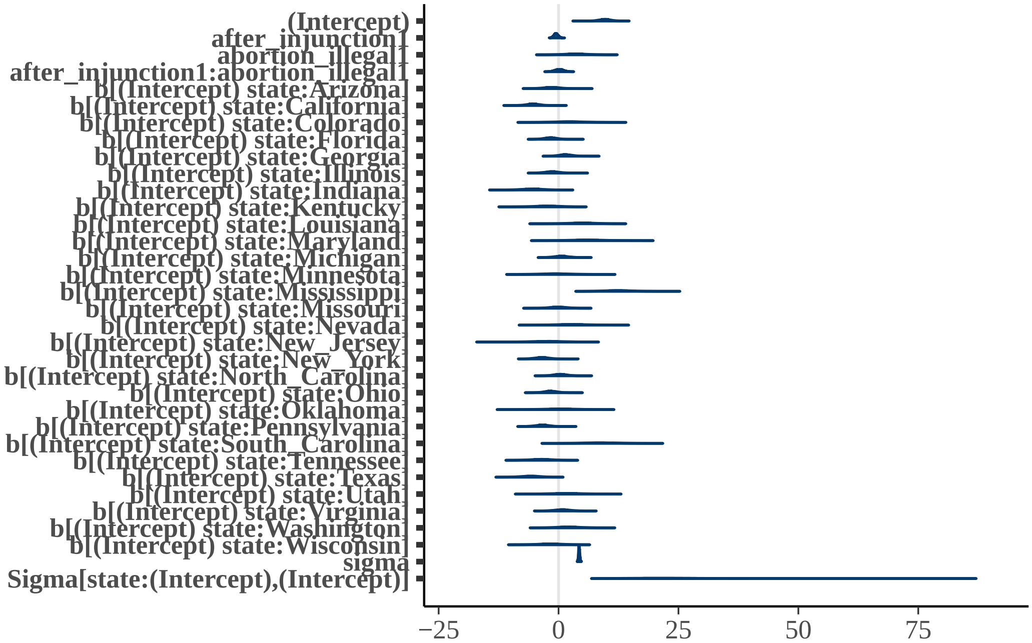


(a) Trace plot

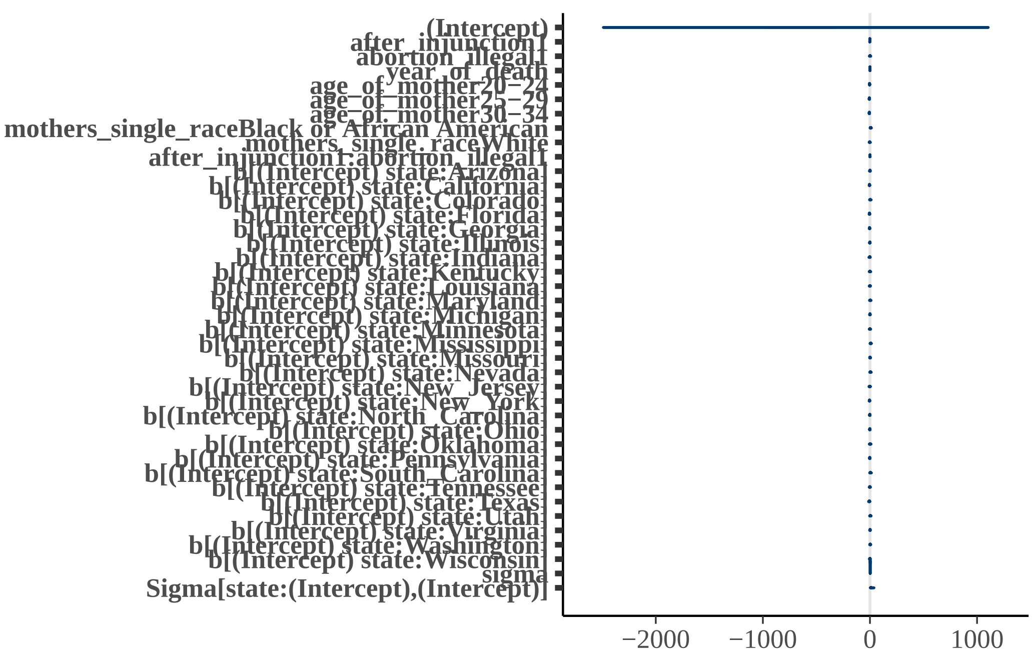
(b) Rhat plot

Figure 6: Checking the convergence of the MCMC algorithm

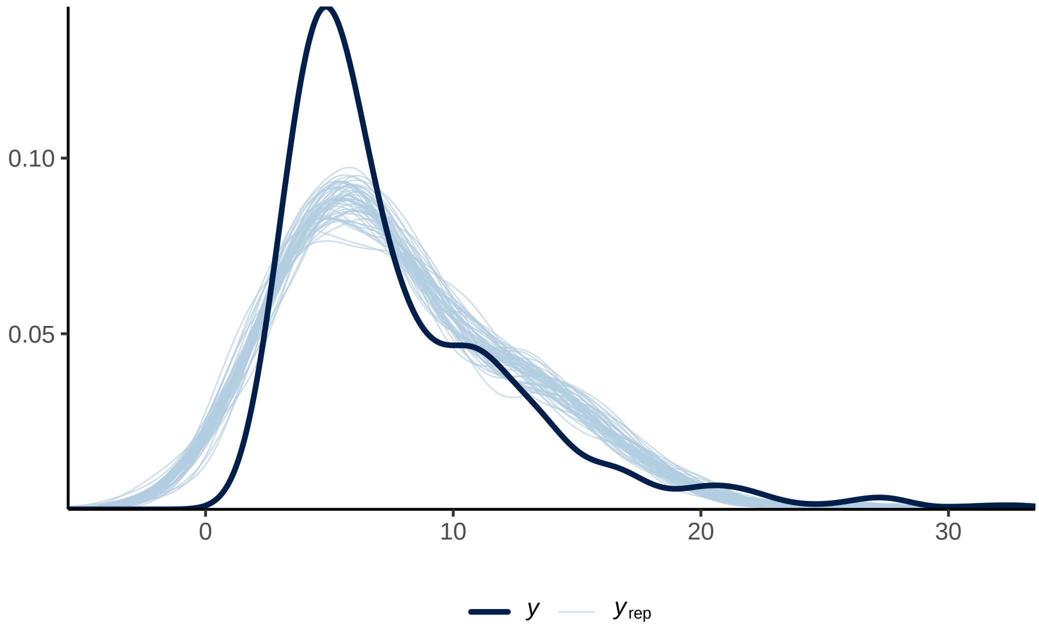
```
plot(
  first_model,
  "areas"
)
```




```r
plot(
  second_model,
  "areas"
)
```



```r
pp_check(second_model) +
  theme_classic() +
  theme(legend.position = "bottom")
```



References

Alexander, Rohan. 2023. *Telling Stories with Data*. Chapman; Hall/CRC. https://tellingstorieswithdata.com/.

Burdick, Kendall J., Catherine G. Coughlin, Gabrielle R. D'Ambrosi, Michael C. Monuteaux, Katherine E. Economy, Rebekah C. Mannix, and Lois K. Lee. 2024. "Abortion Restrictiveness and Infant Mortality: An Ecologic Study, 2014-2018." *American Journal of Preventive Medicine* 66 (3): 418–26. https://doi.org/10.1016/j.amepre.2023.10.010.

Chuck, Elizabeth. 2024. "Texas Abortion Ban May Have Contributed to Deaths of at Least 30 Pregnant Women, New Report Finds." 2024. https://www.nbcnews.com/health/womens-health/texas-abortion-ban-deaths-pregnant-women-sb8-analysis-rcna171631.

Disease Control, Centers for, and National Center for Health Statistics Prevention. 2024. "National Vital Statistics System, Linked Birth / Infant Deaths on CDC WONDER Online Database." http://wonder.cdc.gov/lbd-current-expanded.html.

Goodrich, Ben, Jonah Gabry, Imad Ali, and Sam Brilleman. 2022. "rstanarm: Bayesian applied regression modeling via Stan." https://mc-stan.org/rstanarm/.

R Core Team. 2023. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing. https://www.R-project.org/.

Reproductive Rights, Center for. 2024. "Abortion Laws by State." 2024. https://reproductiverights.org/maps/abortion-laws-by-state/.