Supplementary Online Content

Wallace J, Goldsmith-Pinkham P, Schwartz JL. Excess death rates for Republican and Democratic registered voters in Florida and Ohio during the COVID-19 pandemic. *JAMA Intern Med*. Published online July 24, 2023. doi:10.1001/jamainternmed.2023.1154

This supplementary material has been provided by the authors to give readers additional information about their work.

eMethods. Supplemental Description of Methods

A. Description of study data and linkages
This study made use of data from five different sources. Below, we describe each of these data sources in more detail and then discuss how these data were assembled into our analytic file.

Florida voter file:
The publicly available Florida voter file for February 2017 was accessed via the Harvard Dataverse. For additional details on the file and a link to request access to the data proceed to the following: https://dataverse.harvard.edu/dataset.xhtml?persistentId=doi:10.7910/DVN/UBIG3F. The file contains full name, date of birth, county of registration, gender, and party affiliation.

Ohio voter file:
The publicly available Ohio voter file for 2017 was accessed via the Ohio Secretary of State website at: https://www6.ohiosos.gov/ords/f?p=VOTERFTP:HOME. We accessed the link and downloaded the data on March 4, 2017. The file contains full name, date of birth, county of registration, and party affiliation.

Datavant:
Detailed mortality data for 2018 to 2021 was obtained from Datavant, an organization that augments the Social Security Administration Death Master File with information from newspapers, funeral homes, and other sources to construct an individual-level database of more than 80% of annual US deaths. For each record, the identifiable data indicate the week of death and age of the deceased individual in months, as well as individual-level identifiers that could be used to link to the voter files but that we, as researchers, did not have access to in the linked data. The Datavant mortality data was then linked, at the individual level, to the Florida and Ohio voter files on first name, last name, and date of birth using a proprietary algorithm. The Datavant algorithm uses a machine learning model to make pairwise determinations as to whether two records represent the same individual. These determinations are driven by a precision target set for a pool of records. In this study, a precision target of 95% was used, which correlates to a probability threshold of 57%. In other words, two records were considered a match if the probability they were a match was 57% as determined by the Match model. A match indicated that a registered voter in Florida and Ohio had been identified in the Datavant mortality data and provided detailed information on date of death and the age of the deceased individual in months, top-coded at 89 years old. A deidentified version of the linked data was then provided to our research team. That dataset contained date of death, year-month of birth, county of residence based on the voter registration file in 2017, party affiliation, and either gender (if from Florida) or a probabilistic guess at gender (if from Ohio) based on first name.

National Center for Health Statistics (NCHS):
To assess whether patterns of excess deaths were qualitatively similar between our linked voter and mortality data and other reliable sources, we obtained death counts from the National Center for Health Statistics (NCHS; US Centers for Disease Control and Prevention).

Centers for Disease Control and Prevention (CDC):
From the CDC, we accessed data on county-level vaccination rates as of June 23, 2022. That data can be accessed here: https://data.cdc.gov/Vaccinations/COVID-19-Vaccinations-in-the-United-States-County/8xkx-amqh. From that data, we obtained information on what share of a county's population had received at least one dose of a Covid-19 vaccine as of March 1, 2021, May 1, 2021, and October 1, 2021. We selected May 1, 2021 as the primary date for the county-level vaccination rate—one month after eligibility for vaccines opened to all adults in our study states—because it represented the approximate date when all adults would have had the opportunity to receive at least one dose of a COVID-19 vaccine if they so desired, taking into account the time that states required during April 2021 to schedule and administer vaccines to newly eligible adults seeking them. However, as a robustness check, we assessed the sensitivity of our findings to using county-level vaccination rates on alternative dates before (March 1, 2021) and after (October 1, 2021) our chosen date.

B. Excess death rate methodology

To estimate the baseline number of deaths we would expect in the absence of the COVID-19 pandemic we fit a Poisson regression model to the weekly observed death counts at the county-by-party-by-age-level for the period of January 1, 2018, through December 31, 2019.

Once we fit the model, we predicted expected deaths for the period January 1, 2018 through December 31, 2019 (in-sample) and then projected forward until December 31, 2021. Excess deaths equaled the difference between observed and expected deaths.

We estimate the model separately by state and age-bin. We adjusted for baseline differences in death counts at the party-by-age-bin-by-county level, and account for seasonality at the state by age-bin level. Our Poisson regression model took the form:

$$\log(\lambda_{\{tpac\}}) = \alpha_{\{pac\}} + \gamma_{\{as(c)\}}\omega_{\{t\}}$$
$$Y_{\{tpac\}} \sim Poisson(\lambda_{\{tpac\}})$$

Where $Y_{\{tpac\}}$ are the total deaths in period t, for party p, in age-bin a, and in county c, and $\alpha_{\{pac\}}$ are party-by-age-bin-by-county fixed effects, where $\omega_{\{t\}} = \sin(\theta_t) + \sin\left(\frac{\theta_t}{2}\right) + \cos\left(\theta_{\{t\}}\right) + \cos\left(\frac{\theta_{\{t\}}}{2}\right), \theta_{\{t\}} = \frac{2\pi t}{52.1775}$. The presence of the $\alpha_{\{pac\}}$ fixed effects makes it unnecessary to include the number of voters for each party-by-age-bin-by-county as an offset (such as in some excess death estimation approach, such as *excessmort (https://cran.r-project.org/web/packages/excessmort/index.html))*. These fixed effects directly estimate the baseline number of deaths in a cell in the estimation window, and then the model proportionately shifts the estimated deaths accordingly to this baseline level of deaths.

It is easy to accommodate overall linear time trends or state-specific linear time trends:

$$\log(\lambda_{\{tpac\}}) = \alpha_{\{pac\}} + \gamma_{\{as(c)\}}\omega_{\{t\}} + \beta_{\{s(c)\}}t$$
$$Y_{\{tpac\}} \sim Poisson(\lambda_{\{tpac\}})$$

Where t is a linear time trend for the number of weeks since the first week of 2017, and $\beta_{\{s(c)\}}$ can either vary by state or be fixed across both.

The parameters for the model are estimated using Poisson regression with the *fixest* package version 0.11.1, with the $\alpha_{\{pac\}}$ treated as nuisance parameters, and standard errors clustered at the county-level.

To compute prediction intervals for excess death rates in Figure 1, Table 1, and eTable 2 we used the following procedure.

1. We extracted the estimated asymptotic covariance matrix for the model after estimation and sampled from a multivariate normal distribution centered at the estimated values. We drew 100 samples from this parameter distribution and for each sample computed $\hat{\lambda}_{\{tpac\}}$.
2. Because independent Poissons are additive, we sum $\hat{\lambda}_{\{tpac\}}$ up to whatever level of aggregation we are reporting. For example, when predicting the level of deaths in aggregate for a time period *t*, we calculated the total $\hat{\lambda}_t = \sum \hat{\lambda}_{\{tpac\}}$.
3. We then simulate 100 times from a Poisson distribution for each of the 100 simulated $\hat{\lambda}_t$, giving us 10,000 samples from the empirical predictive distribution. The reported 95% prediction intervals reflect the 2.5 and 97.5th percentiles of this distribution.

When taking the difference in excess death rates between Republican and Democratic voters, we adjust for the relative predicted death rates across different groups. Our parameter of interest is

$$\hat{\tau}_{\{t\}} = \sum_{a,s} w_{tas}\left[\left(\left(Y_{\{tra,s(c)\}} - \hat{Y}_{\{tra,s(c)\}}\right)/\hat{Y}_{\{tra,s(c)\}}\right) - \left(\left(Y_{\{tda,s(c)\}} - \hat{Y}_{\{tda,s(c)\}}\right)/\hat{Y}_{\{tda,s(c)\}}\right)\right]$$

$$w_{\{tas\}} = \frac{\hat{Y}_{\{tra,s(c)\}} + \hat{Y}_{\{tda,s(c)\}}}{\sum_{a,s} \hat{Y}_{\{tra,s(c)\}} + \hat{Y}_{\{tda,s(c)\}}}$$

where $\hat{Y}_{\{tra,s(c)\}}$ is the predicted number of deaths in period *t* for Republicans in age-bin for state s, and $Y_{\{tra,s(c)\}}$ is the realized number of deaths. Hence, $w_{\{tas\}}$ is the period *t* share of predicted deaths in age-bin *a* and state *s*, which readjusts the relative weighting of death rate comparisons to account for differences in the age groups and states. For the predicted value for $\hat{\tau}_{\{t\}}$, we use $\widehat{Y_{tra,s(c)}} = \hat{\lambda}_{\{tra,s(c)\}}$. To construct prediction intervals, we perform the same process as above, simulating 100 values of $\hat{\lambda}_{\{tra,s(c)\}}$ and then sampling 100 values from the Poisson distribution before constructing $\hat{\tau}_{\{t\}}$. We then construct the 95% prediction intervals in the same way.

We would like to further adjust for age-bin-by-county level differences, but the sampled probabilities for $\hat{Y}_{\{trac\}}$ can include zeros because the cell bins are so small. As a result, we cannot calculate an excess death rate for group for these cells (since we would be dividing by zero). We can, however, estimate the mean differences using $\hat{\lambda}_{\{trac\}}$ because these estimated means are all greater than zero in expectation). We report the following in supplemental results to show robustness:

$$\tilde{\tau}_{\{t\}} = \sum_{a,c} w_{tac}\left[\left(\left(Y_{\{tra,s(c)\}} - \hat{\lambda}_{\{trac\}}\right)/\hat{\lambda}_{\{trac\}}\right) - \left(\left(Y_{\{tda,s(c)\}} - \hat{\lambda}_{\{tdac\}}\right)/\hat{\lambda}_{\{tdac\}}\right)\right]$$

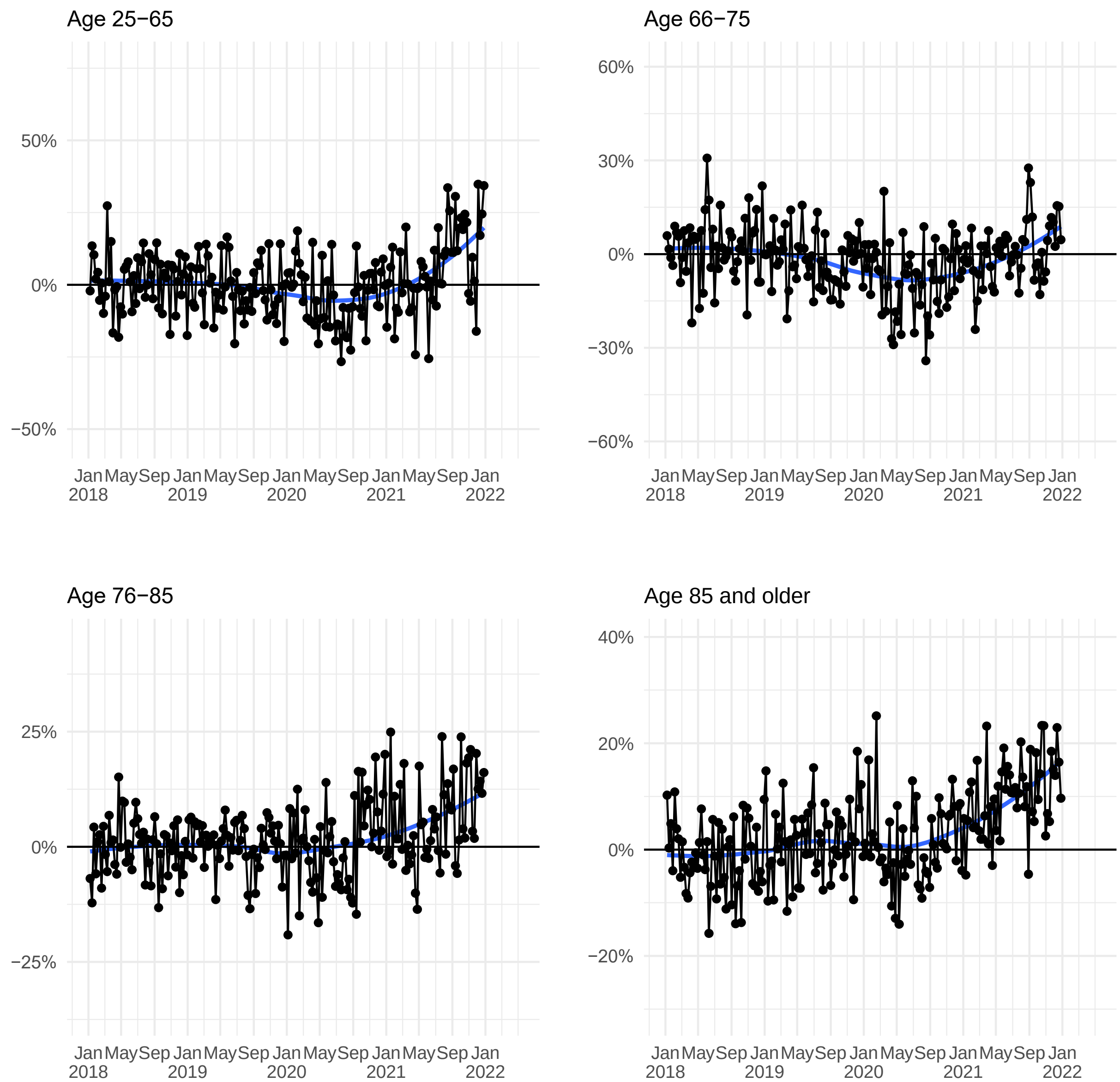
$$w_{\{tac\}} = \frac{\hat{\lambda}_{\{trac\}} + \hat{\lambda}_{\{tdac\}}}{\sum_{a,c} \hat{\lambda}_{\{trac\}} + \hat{\lambda}_{\{tdac\}}}.$$

Intuitively, this reweighting balances the distribution of age-bin-by-county across the registered Republican and Democratic voters. This adjusts for potential differences in excess death rates at the age-bin-by-county level. When presenting these results, we present only point estimates without prediction intervals because we are unable to simulate within the cells due to the zero predicted deaths issue described above.

C. Measuring the Date of Open Eligibility for Vaccines Among Adults

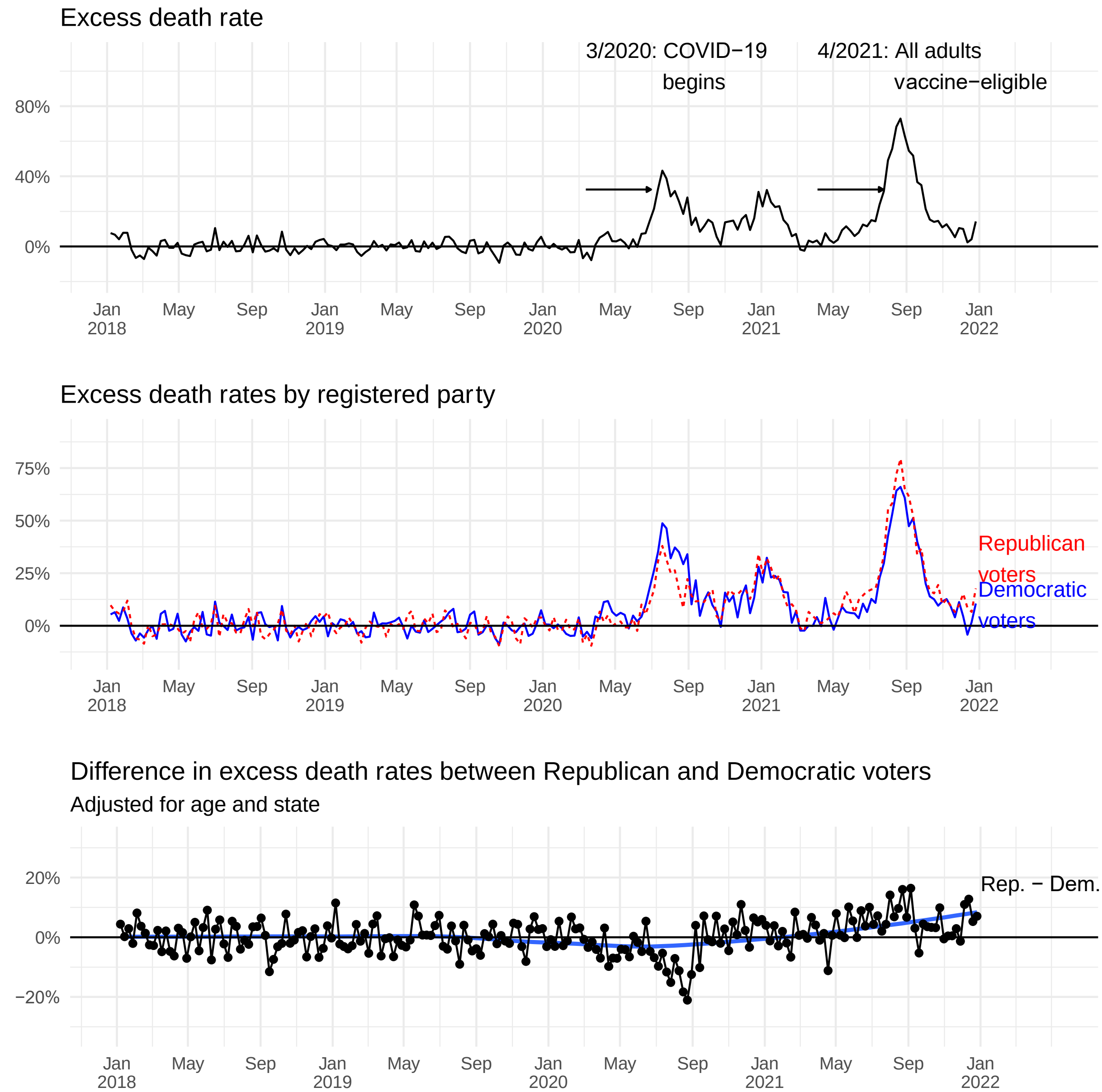
We selected April 5, 2021 as a cutoff for when vaccines were available to all adults in Florida and Ohio. The federal deadline for when states were required to open eligibility for vaccines for all adults was April 19, 2021, but Florida and Ohio did so earlier. Florida made vaccines available to all adults on April 5, 2021. For more details on Florida's policy see: https://floridahealthcovid19.gov/news/. Ohio made vaccines available to all adults earlier, on March 29, 2021. For more details on Ohio's policy see: https://governor.ohio.gov/media/news-and-media/expanded-vaccine-eligibility-+cleveland-mass-vaccination-clinic-opens-to-public-03162021. We selected the latter of the two dates, April 5, 2021, as the date by which all adults in both states were eligible for vaccines. Because our data is weekly, we use April 1, 2021 data and onwards to reflect this period, which corresponds closely to when vaccines were available to all adults in our study states.

eFigure 1. Excess Death Rates by Age in Florida and Ohio: 2018 – 2021



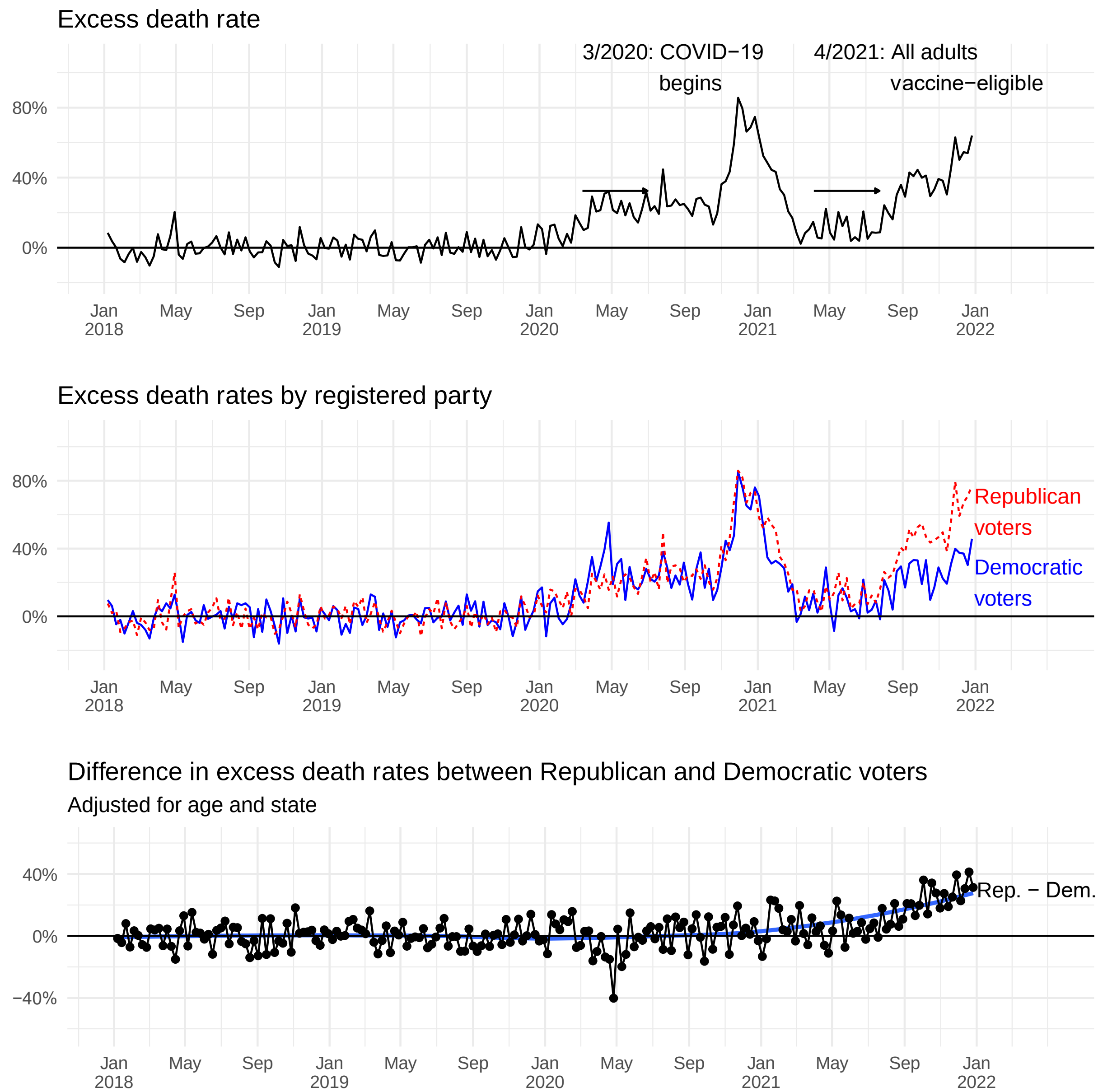
Legend: The figure plots the percentage point difference in excess death rates between Republican and Democratic voters in Florida and Ohio separately by age bin after adjusting for state-level differences. For additional details on the excess death methodology and statistical analyses refer to the **Methods** section and **eMethods** in the **Supplement**.

eFigure 2. Excess Death Rates in Florida: 2018 – 2021



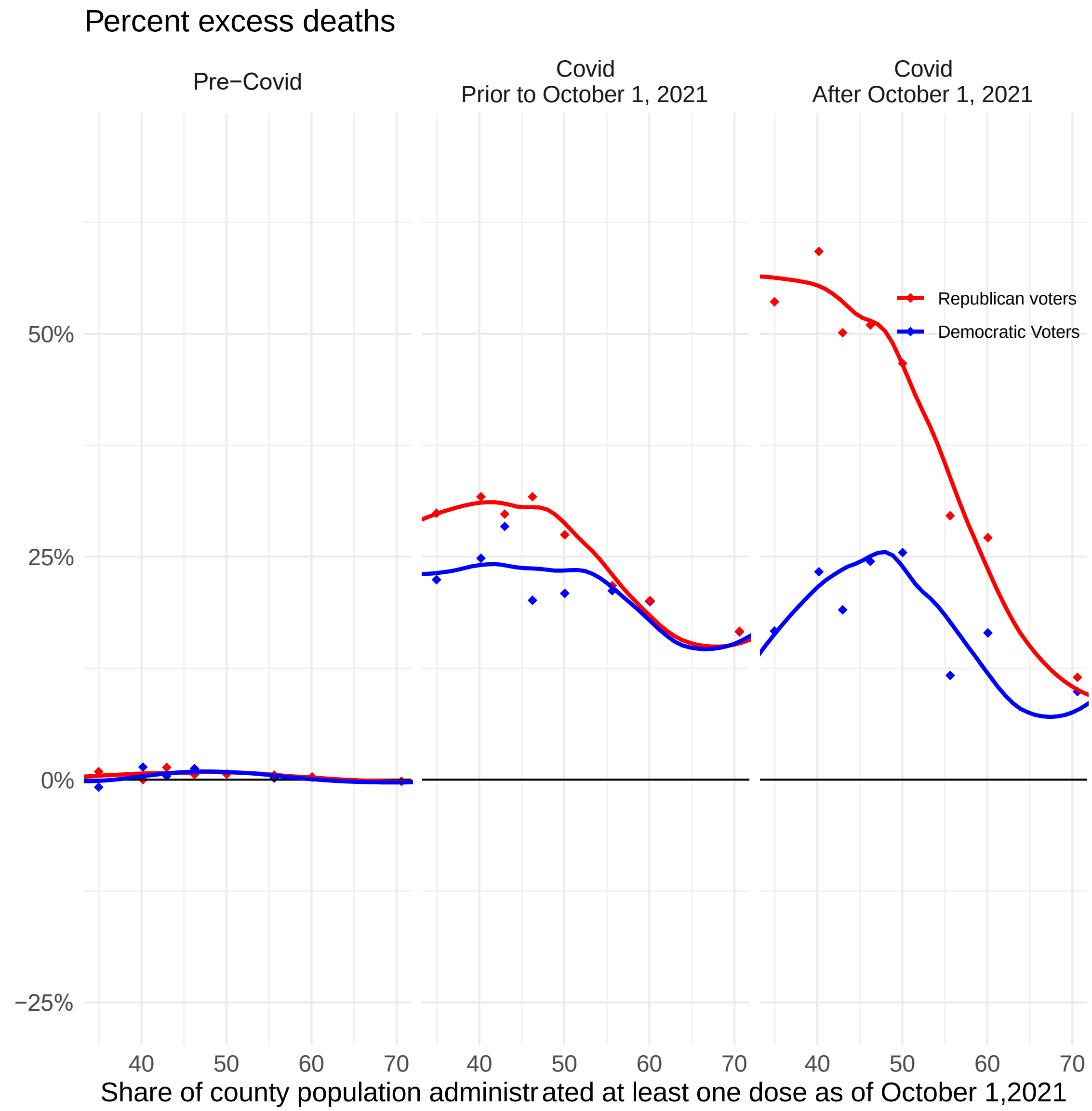
Legend: The figure plots weekly excess deaths for Florida based on mortality records linked to voter registration files. Excess death rates are calculated for each week by comparing the observed deaths in that week to expected deaths based on a Poisson model. We include 95% prediction intervals using simulations from the Poisson coefficient and outcome distribution, with standard errors clustered at the county level. The top panel plots overall excess death rates. The middle panel plots excess death rates separately for registered Republican and Democratic voters. The bottom panel plots the percentage point difference in excess death rates between Republican and Democratic voters after adjusting for age and state-level differences. The blue curve in the bottom panel was fit with a LOESS smoother. For additional details on the excess death methodology and statistical analyses refer to the **Methods** section and the **eMethods** section in the **Supplement**.

eFigure 3. Excess Death Rates in Ohio: 2018 – 2021



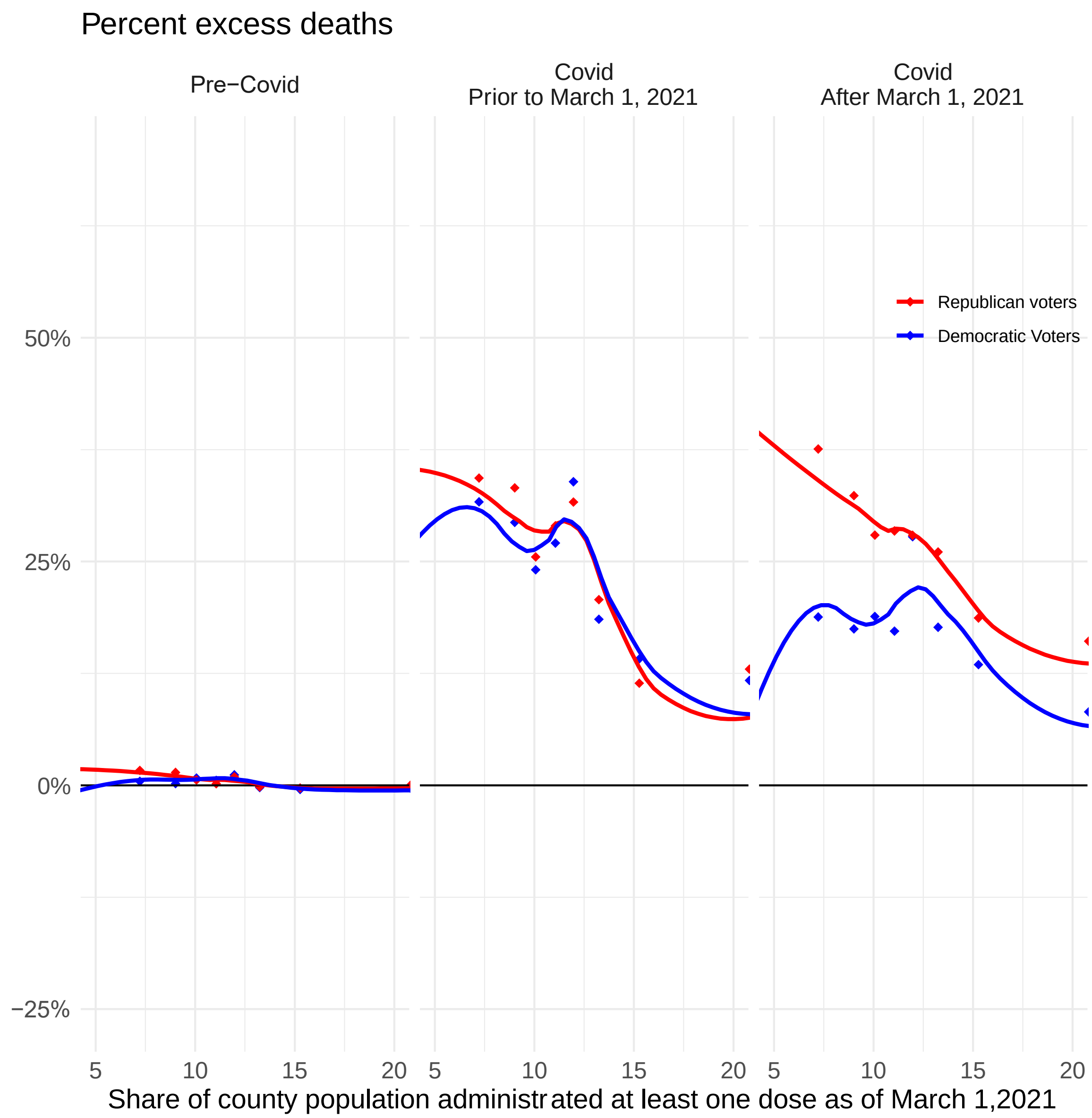
Legend: The figure plots weekly excess deaths for Ohio based on mortality records linked to voter registration files. Excess death rates are calculated for each week by comparing the observed deaths in that week to expected deaths based on a Poisson model. We include 95% prediction intervals using simulations from the Poisson coefficient and outcome distribution, with standard errors clustered at the county level. The top panel plots overall excess death rates. The middle panel plots excess death rates separately for registered Republican and Democratic voters. The bottom panel plots the percentage point difference in excess death rates between Republican and Democratic voters after adjusting for age and state-level differences. The blue curve in the bottom panel was fit with a LOESS smoother. For additional details on the excess death methodology and statistical analyses refer to the **Methods** section and the **eMethods** section in the **Supplement**.

eFigure 4. Excess Death Rates and Vaccination Rates in Florida and Ohio During the COVID-19 Pandemic using October 1, 2021, Vaccination Rates



Legend: This figure presents estimates of excess death rates using linked mortality and voter data that has been aggregated to the county level. These data are presented separately for Republican and Democratic voters. The y-axis contains the party-specific estimates of excess death rates and the x-axis contains the share of the county population administered at least one dose of the vaccine as of October 1, 2021. The blue and red curves were fit using a loess smoother. In the pre-Covid period (prior to April 2020), excess death rates for both Republican and Democratic voters hover around zero. During the beginning of the pandemic but prior to October 1, 2021, the association between excess death rates and county-level vaccination rates are negative and somewhat similar for Republican and Democratic voters. In the period after October 1, 2021 there is a clear visual difference between Republican and Democratic voters, with higher excess death rates for Republicans concentrated in counties with lower overall vaccination rates and minimal differences in counties with the highest vaccination rates.

eFigure 5. Excess Death Rates and Vaccination Rates in Florida and Ohio During the COVID-19 Pandemic using March 1, 2021, Vaccination Rates



Legend: This figure presents estimates of excess death rates using linked mortality and voter data that has been aggregated to the county level. These data are presented separately for Republican and Democratic voters. The y-axis contains the party-specific estimates of excess death rates and the x-axis contains the share of the county population administered at least one dose of the vaccine as of March 1, 2021. The blue and red curves were fit using a loess smoother. In the pre-Covid period (prior to April 2020), excess death rates for both Republican and Democratic voters hover around zero. During the beginning of the pandemic but prior to March 1, 2021, the association between excess death rates and county-level vaccination rates are negative and similar for Republican and Democratic voters. In the period after March 1, 2021 there is a clear visual difference between Republican and Democratic voters, with higher excess death rates for Republicans concentrated in counties with lower overall vaccination rates and smaller differences in counties with the highest vaccination rates.

eTable 1. Summary Statistics

	Florida	Ohio	Total
Panel A. Voter registration data			
Counts of 2017 registered voters			
Republican voters	4,781,802	2,044,251	6,826,053
Democratic voters	5,184,144	1,302,223	6,486,367
Other voters	3,744,412	4,546,770	8,291,182
Total voters	13,710,358	7,893,244	21,603,602
Democratic voter age distributions			
Voters aged 25-64	3,360,007	816,532	4,176,539
Voters aged 65-74	781,810	252,676	1,034,486
Voters aged 75-84	437,754	126,080	563,834
Voters aged 85+	248,465	54,453	302,918
Republican voter age distributions			
Voters aged 25-64	2,890,858	1,301,001	4,191,859
Voters aged 65-74	832,564	399,015	1,231,579
Voters aged 75-84	533,035	211,131	744,166
Voters aged 85+	246,394	79,190	325,584
Panel B. Linked Mortality data			
Counts of Democratic voter deaths, by age			
Voters aged 25-64	32,029	10,367	42,396
Voters aged 65-74	34,754	15,124	49,878
Voters aged 75-84	48,592	19,543	68,135
Voters aged 85+	66,890	21,576	88,466
Counts of Republican voter deaths, by age			
Voters aged 25-64	28,053	14,153	42,206
Voters aged 65-74	34,196	20,522	54,718
Voters aged 75-84	58,357	30,369	88,726
Voters aged 85+	72,034	31,801	103,835
Age at death, mean (IQR)			
Democratic voters	77 (70, 89)	77 (70, 88)	77 (70, 89)
Republican voters	78 (72, 89)	78 (71, 88)	78 (72, 89)
Democratic and Republican voters	78 (71, 89)	78 (71, 88)	78 (71, 89)

Note: Table presents summary statistics for the study cohort of 2017 registered voters in Florida and Ohio along with observed death counts by year based on mortality records linked to voter registration files at the individual level. For additional details on the data sources refer to the **Methods** section and **eMethods** in the **Supplement**.

eTable 2. Sensitivity of Estimated Difference in Excess Death Rates between Republican and Democratic Voters to Alterations in Excess Death Methodology and Statistical Model

		Sensitivity to changes in Poisson model used to predict baseline death counts			Sensitivity to adjustments used to compare excess death rates between Rep. and Dem. voters	
	Primary model (1)	**Add linear trend (2)**	**Add state-varying linear trend (3)**	**Add 6-month periodicity term for seasonality (4)**	**Remove all adjusters from primary model (5)**	**Add county-by-age adjusters to primary model (6)**
Full period: 4/1/2020-12/31/2021, all groups	2.78 [1.56,3.7]	3.56 [2.48,4.59]	2.82 [1.6,3.77]	2.82 [1.61,3.81]	2.80 [1.59,3.77]	3.09
Panel A. By Period						
Prior to open vaccine eligibility: 4/1/2020-3/31/2021	-0.882 [-2.5,0.33]	0.519 [-0.82,1.84]	-0.895 [-2.5,0.32]	-0.895 [-2.51,0.33]	-0.888 [-2.47,0.37]	-0.510
After open vaccine eligibility: 4/1/2021-12/31/2021	7.71 [5.96,9.26]	7.65 [6.06,9.26]	7.83 [5.95,9.53]	7.82 [5.96,9.54]	7.76 [5.89,9.46]	7.94
Panel B. By Age						
Voters aged 25-64	0.875 [-1.84,3.48]	0.912 [-1.76,3.59]	0.888 [-1.89,3.62]	0.888 [-1.88,3.66]	0.880 [-1.97,3.67]	1.58
Voters aged 65-74	-3.60 [-6.32, -1.3]	-3.09 [-5.61, -0.7]	-3.65 [-6.51,-1.34]	-3.63 [-6.4,-1.29]	-3.60 [-6.35,-1.23]	-2.34
Voters aged 75-84	3.98 [1.78,5.85]	4.59 [2.53,6.47]	4.04 [1.83,5.9]	4.03 [1.79,5.92]	3.99 [1.82,5.96]	4.28
Voters aged 85+	6.09 [4.29,7.7]	7.28 [5.61,8.94]	6.18 [4.3,7.91]	6.20 [4.32,7.93]	6.16 [4.33,7.92]	5.73
Panel C. By State						
Florida	0.863 [-0.35,2.08]	0.834 [-0.35,2.03]	0.876 [-0.36,2.1]	0.906 [-0.44,2.17]	0.902 [-0.41,2.16]	1.52
Ohio	7.43 [4.82,9.18]	7.61 [5.39,9.54]	7.54 [4.87,9.32]	6.96 [4.58,8.8]	6.87 [4.52,8.63]	6.91

Note: This table presents estimates of differences in excess death rates for Republican and Democratic voters in Florida and Ohio based on mortality records linked to 2017 voter registration files at the individual level. The word "Rep." is short for Republican and "Dem." is short for Democratic. Our primary model in column (1) is the one that is used to produce the estimates in the rightmost column of **Table 1** in the manuscript. Columns (2)-(4) present the estimated differences in excess death rates between Republican and Democratic voters from Poisson models that vary in how they model expected death counts, allowing for linear time trends and additional periodicity terms to capture seasonality. Columns (5) and (6) assess the sensitivity of our estimates to using different sets of adjusters for differences between groups in excess death rates during COVID-19. Our primary model adjusts for differences between groups at the state-by-age-bin level. Column (5) removes this adjustor. Column (6) retains state-by-age-bin level adjusters and adds an additional adjustment at the county-by-age-bin level. We present only point estimates in column (6), because we are unable to simulate within the cells due to the zero predicted deaths issue described in the **eMethods** section in the **Supplement**. For additional details on the excess death rate methodology and statistical analyses please refer to the **Methods** section and the **eMethods** section in the **Supplement**.