Supplement to Polarization and Public Health: Partisan Differences in Social Distancing during the Coronavirus Pandemic

Hunt Allcott, *New York University, Microsoft Research, and NBER*[*]
Levi Boxell, *Stanford University*
Jacob Conway, *Stanford University*
Matthew Gentzkow, *Stanford University and NBER*
Michael Thaler, *Harvard University*
David Yang, *Harvard University and NBER*

August 2020

Appendix

Contents

[*]E-mail: hunt.allcott@nyu.edu, lboxell@stanford.edu, jcconway@stanford.edu, gentzkow@stanford.edu, michaelthaler@g.harvard.edu, davidyang@fas.harvard.edu.

List of Figures

A Appendix

A.1 Data Details

See replication code for exact details on implementation.

A.1.1 County-Level Data Build

To construct the county-level dataset used in the analysis, we proceed as follows:

1. We begin by matching SafeGraph POIs to the counties in which they are located. We use latitude and longitude from SafeGraph's July 2020 Core POI dataset, along with the 2010 TIGER county shapefile.[1] We successfully assign 99.9 percent of the POIs to a county.

2. We then merge the POI-county mapping from (1) onto SafeGraph's Patterns data using the safegraph-place-id variable. We sum visits by county for a given day, aggregating across POIs.

3. We then merge alternative county-day measures of social distancing onto the output from (2). These measures are constructed as follows from the Daily Social Distancing SafeGraph

[1] Downloaded from https://www.census.gov/geo/maps-data/data/cbf/cbf_counties.html on July 24, 2018.

data with observations at the census block group-day level for January 27 through July 12. We exclude Alaska. We restrict our sample to census block groups with active devices throughout the entire time period. We also drop one census block group with anomalous behavior as notified by SafeGraph (FIPS: 190570010001). We aggregate this series to the county level. Countable variables (e.g., 'device count') are summed, while we take a 'device count' weighted average of other variables (e.g., 'median home dwell time').

4. We then merge gridMET weather data onto the output from (3). Precipitation and temperature means for a given county day are taken as a mean across grid cell points that lie within a county boundary. Weather data was not available for Hawaii, so this particular state is dropped in regressions including weather controls.

5. We then merge The New York Times COVID-19 tracking data onto our output from (4). We assume zero cases and deaths for the observations not observed in The New York Times data. We drop the five counties associated with New York City and the four counties which overlap with Kansas City (MO), because The New York Times lists these as geographic exceptions where it either does not assign cases to these counties or excludes cases occurring within the city.

6. We then merge a dataset of county-level shelter-in-place order start dates onto the output from (5) and construct an indicator for whether a county had been subject to a shelter-in-place order by a given date. This dataset of shelter-in-place orders is the same as in Allcott et al. (2020), where its construction is described in detail. It is ultimately sourced from Keystone Strategy, a crowdsourcing effort from Stanford University and the University of Virginia, Hikma Health, and The New York Times.

7. We then aggregate the output from (6) to the county-week level using sums, averages, or start- or end-of-week observations as appropriate.

8. We then merge onto the output from (7) a dataset of county-level demographic information constructed as follows. We use the Open Census data from SafeGraph, aggregating up the data given at the census block group level to the county level. We combine this with data on county 2016 Presidential votes shares (MIT Election Data and Science Lab 2018). We define the Republican vote share to be the share of votes received by the Republican candidate over

the sum of votes across all candidates. We exclude counties without valid vote data, which drops Alaska and two additional counties (FIPS: 15005, 51515).

A.1.2 Precinct-Level Data Build

1. We begin by matching SafeGraph POIs to the precincts in which they are located. We start with the POIs successfully matched in the POI-county mapping from (1) in Section A.1.1. We use POI latitude and longitude along with 2016 precinct-level shapefiles (Voting and Election Science Team 2018). Of these POIs, we successfully match 99.6 percent to a unique precinct in the states covered by these precinct shapefiles and drop the 0.001 percent matched to two precincts. The precinct shapefiles cover the following 42 states: AK, AR, AZ, CA, CO, DC, DE, FL, GA, HI, IA, ID, IL, KS, KY, LA, MA, MD, ME, MI, MN, MO, MT, NC, ND, NE, NH, NM, NV, OK, OR, RI, SC, SD, TN, TX, UT, VA, VT, WA, WI, and WY.

2. We merge the output from (1) with the Patterns dataset from SafeGraph using the safegraph-place-id variable. We sum visits by precinct in a given week, aggregating across POIs.

3. We then merge alternative precinct-day measures of social distancing onto the output from (2). We start with the cleaned census block group level measures produced from (3) in Section A.1.1. We then aggregate our (countable) alternative social distancing variables to the precinct level as follows. We do this by first constructing the geographic intersections formed by our precinct shapefiles and 2019 Tiger census block group shapefiles.[2] Let a_p, a_b, and a_{bp} denote the area of precinct p, census block group b, and of their intersection respectively. For a given count variable x_b given at the block group level, we construct a precinct-level estimate as: $\hat{x}_p := \sum_b \frac{a_{bp}}{a_b} x_b$.[3] We then form ratios (e.g., 'share leaving home') using these summed precinct-level estimates as needed.

4. We then merge gridMET weather data onto the output from (3). Weather data by precinct day was constructed as a land area weighted average of weather data for overlapping census block groups. Census block group centroids were first associated with the nearest grid cell centroid in gridMET. Weather data was not available for Hawaii, so this particular state is dropped in regressions including weather controls.

[2]Downloaded from ftp://ftp2.census.gov/geo/tiger/TIGER2019/BG/ on April 1, 2020.
[3]This estimate is exactly correct if a given demographic x_b is evenly distributed across a census block group's area.

5. We then merge The New York Times COVID-19 county-level tracking data onto our output from (4). Precincts were first associated with the county corresponding to the census block group of their largest intersection. We drop the five counties associated with New York City and the four counties which overlap with Kansas City (MO), because The New York Times lists these as geographic exceptions where it either does not assign cases to these counties or excludes cases occurring within the city.

6. We then merge the dataset of county-level shelter-in-place order start dates from (6) in Section A.1.1 onto the output from (5). We construct an indicator for whether a county had been subject to a shelter-in-place order by a given date.

7. We then aggregate the output from (6) to the precinct-week level using sums, averages, or start- or end-of-week observations as appropriate.

8. We then merge onto the output from (7) a dataset of precinct-level demographic information constructed as follows. We start with the Open Census census block group level demographics from (8) in Section A.1.1. We then aggregate countable variables to the precinct-level using the intersection share procedure described in (3). We then form ratios (e.g., 'population density' or 'share hispanic') using these summed precinct-level estimates. We then merge precinct-level 2016 Presidential votes shares (Voting and Election Science Team 2018) onto the resulting output, constructed as in A.1.1 step (8). We drop Alaska in order to be consistent with our county construction.

A.2 Survey Details

A.2.1 Data

We clean the survey data from Qualtrics as follows:

1. We match participant IDs from Qualtrics with a list of emailed IDs from CloudResearch and drop observations that do not match to remove test subjects. There is one exception, where the ID on Qualtrics did not correctly generate. We find exactly one remaining participant with the same demographics in the CloudResearch, so we keep this participant.

2. We change one miscoded age from .23 to 23 and one miscoded ZIP code from ,43011 to 43011.

3. We merge ZIP code data with 2010 US Census data and match ZIP codes to states to get population density.

4. We match ZIP codes to counties and use the week of March 29-April 4 and get county-level COVID cases and deaths via The New York Times. All ZIP codes in New York City are matched to the city-level cases and deaths since county-level data is unavailable from The New York Times. For analyses, we control for log(county cases + 1) and log(county deaths + 1).

5. We weight observations across age category, gender, race/ethnicity, and party affiliation using the Stata ebalance command. Weights are prespecified in the pre-analysis plan.

6. News sources are numbered in the data in the following order: (1) Network news; (2) Breitbart; (3) CNN; (4) Facebook; (5) Fox News; (6) MSNBC; (7) New York Times; (8) Wall Street Journal; (9) Twitter; (10) Wikipedia; (11) CDC; (12) WHO. News sources (1)-(3) and (5)-(8) are ranked by partisanship as specified in our pre-analysis plan. For our news consumption specification, we use the question on consumption of news about the coronavirus. Often is coded as 1, Sometimes as 2/3, Rarely as 1/3, and Never/Not Familiar as 0. Answers are then multiplied by the pre-specified partisanship of each source (-1 for NYT; -2/3 for MSNBC; -1/3 for CNN; 0 for Network; 1/3 for WSJ; 2/3 for Fox News; 1 for Breitbart). The weighted average of news partisanship for each participant equals the sum of (Answer * Source Partisanship) divided by the sum of Answer.

We have the following demographic groups prior to weighting:

- Age: 45.7% 18-39, 33.8% 40-59, 20.5% 60+
- Gender: 51.9% Female, 47.75% Male, 0.35% Other / Non-binary
- Race: 66.6% White (Not Hispanic or Latinx), 15.25% Hispanic or Latinx, 11.2% Black or African American (Not Hispanic or Latinx), 4.95% Asian or Pacific Islander, 2.0% Other.
- Party: 34.65% Democratic, 31.25% Republican, 32.8% Independent, 1.3% Other

A.2.2 Survey Questions

Screening

- What is your gender? [Male; Female; Other / Non-binary]

- What race/ethnicity best describes you? [American Indian or Alaska Native; Asian or Pacific Islander; Black or African American (Not Hispanic or Latinx); Hispanic or Latinx; White (Not Hispanic or Latinx); Other]
- Do you consider yourself a Republican, a Democrat, or an Independent? [Democrat (Strongly Democratic); Democrat (Weakly Democratic); Independent (Lean toward the Democratic Party); Independent (Do not lean towards either party); Independent (Lean toward the Republican Party); Republican (Weakly Republican); Republican (Strongly Republican); Other / prefer not to say]
- What is your age?
- Do you currently live in the United States? [Yes; No]

Consent

[Page seen if age $>$ 18, United States = Yes, and not screened out due to demographic quotas.]

Congratulations! You are eligible to participate. Please read the consent form below:

DESCRIPTION: You are invited to participate in an online research study on your views about the news and predictions of what will happen in the future. This is a research project being conducted by researchers at Harvard University and New York University.

TIME INVOLVEMENT: Your participation will take approximately 20 minutes, and the entire study will take place online.

RISKS AND BENEFITS: We will ensure that your individual responses are strictly confidential, and research results will only be presented in the aggregate. Your responses will not be shared with government officials or any 3rd party. We hope that the knowledge gained from this study will benefit society in general. **We cannot and do not guarantee or promise that you will receive any direct benefits from this study.**

PAYMENTS: If you are eligible for the study, and once you complete the study, you will receive a participation fee. You may also earn a bonus payment of up to $100 via an Amazon gift card. All payments will be through your research provider.

PARTICIPANT'S RIGHTS: If you have read this form and have decided to participate in this project, please understand your **participation is voluntary** and you have the **right to withdraw your consent or discontinue participation at any time without penalty or loss of benefits to**

which you are otherwise entitled. The alternative is not to participate. You have the right to refuse to answer particular questions. The results of this research study may be presented at scientific or professional meetings or published in scientific journals.

CONTACT INFORMATION:

Questions: If you have any questions, concerns or complaints about this research, its procedures, risks and benefits, contact the researchers at rb4337@nyu.edu.

Independent Contact: If you are not satisfied with how this study is being conducted, or if you have any concerns, complaints, or general questions about the research or your rights as a participant, please contact the Harvard University Area Institutional Review Board (IRB) to speak to someone independent of the research team at cuhs@harvard.edu, (617)-496-2847. You can also write to the Committee on the Use of Human Subjects, Harvard University, 44-R Brattle Street, Suite 200, Cambridge, MA 02138.

Please retain a copy of this form for your records.

If you wish to participate in this study, please click "I consent" to proceed. This serves as an electronic signature indicating your consent to participate in the study.

[I consent; I do not consent]

[Only consenting subjects proceed]

Demographics

- How many children under the age of 18 do you have? [0; 1; 2; 3; 4; 5 or more]
- What is the highest degree or level of schooling that you have completed? [Less than a high school diploma; High school diploma or equivalent (for example: GED); Some college but no degree; Associate's degree; Bachelor's degree; Graduate degree (for example: MA, MBA, JD, PhD)]
- What was your total income in 2019? Please include only employment income (wages, salary, bonuses, tips, and any income from your own businesses). [I did not earn income in 2019; $1 to $9,999; ...; $50,000 to $59,999; $60,000 to $74,999; $75,000 to $99,999; $100,000 to $124,999; $125,000 to $149,999; $150,000 or more] [Coded as midpoint of range in thousands of dollars except for top bracket, who is coded at 200. Log(income + 1) is used as the control.]
- In what ZIP Code do you currently live? Please enter your 5-digit ZIP Code.

- In general, how would you rate your OVERALL health? [Excellent / Very good / Good / Fair / Poor]
- Has a doctor ever told you that you had the following conditions? [Yes / No]
  - Diabetes or high blood sugar
  - Lung disease such as chronic bronchitis or emphysema
  - A heart attack, coronary heart disease, angina, congestive heart failure, or other heart problems
- Please answer the following yes/no questions:
  - In the past week, have you had to go to a work environment in which you were within six feet of others?
  - Have you smoked at least 100 cigarettes in your entire life?
  - Have you smoked at least 10 cigarettes in the past week?

Information sources

- All of the following questions were asked about the following 12 news sources: Network news (ABC, CBS, NBC); Breitbart; CNN; Facebook; Fox News; MSNBC; The New York Times; The Wall Street Journal, Twitter, Wikipedia, The Centers for Disease Control (CDC); The World Health Organization (WHO).
  - **Last year**, how much **trust and confidence** did you have in each of the following sources when it comes to **reporting about politics and current events fully, accurately, and fairly**? [A great deal / A fair amount / Not very much / None at all / Not familiar with this outlet]
  - **Last year**, how **frequently did you get news and information** from each of the following sources **about politics and current events** through any medium (including reading online, watching on TV, etc.)? [Often / Sometimes / Rarely / Never / Not familiar with this outlet]
  - How much **trust and confidence** do you have in each of the following sources when it comes to **reporting about the coronavirus fully, accurately, and fairly**? [A great deal / A fair amount / Not very much / None at all / Not familiar with this outlet]

  - How **frequently are you getting news and information** from each of the following sources **about the coronavirus** through any medium (including reading online, watching on TV, etc.)? [Often / Sometimes / Rarely / Never / Not familiar with this outlet]

Changes in behavior and effects of social distancing

- Think about the ways you may have changed your daily routine in the past two weeks specifically because of the coronavirus. For example, you may be washing your hands more, avoiding restaurants and other public places, and/or reducing interactions with friends and family.

- By what percent have you reduced your overall contact with other people as a result of the coronavirus outbreak? Please enter a percentage from 0 to 100.

- Think back to two weeks ago.

- As of two weeks ago, by what percent had you reduced your overall contact with other people as a result of the coronavirus outbreak? Please enter a percentage from 0 to 100.

- Imagine that starting today and for the rest of the month, you went back to your **normal daily routine from before the coronavirus**. What do you think is the probability that you would catch the coronavirus in the next month? Please enter a percentage from 0 to 100. [Subjects who answer 0 for the percent reduction question see "continued with" instead of "went back to."]

- Imagine that starting today and for the next month, **you cut off all in-person contact with people outside your household.** What do you think is the probability that you would catch the coronavirus in the next month? Please enter a percentage from 0 to 100.

- We'd like to quantify the overall costs (in terms of time, money, and inconvenience) that social distancing imposes on you. Consider a hypothetical situation in **a normal month in the future, after the coronavirus outbreak is completely over.**

  Imagine you had a choice between:

  (A) following your normal routine for one month,

  OR

(B) cutting off all in-person contact with people outside your household for one month, AND receiving $X cash.

Presumably if you were offered a large amount of cash ($X is large), you'd be willing to cut off all social contact. If you weren't offered any cash ($X is 0), you'd prefer to stick with your normal routine. What value of X would make you equally happy with these two options? Please answer in dollars.

Economic trade-offs

- When there was no "stay-at-home" order for your area, what did you think was the best way to help the country in this time of crisis? [7-point scale from "Go out more to help the economy" to "Go out less to avoid spreading the coronavirus"]

Predictions

[If unincentivized:]

- You will now be asked to make a few predictions.

[If incentivized:]

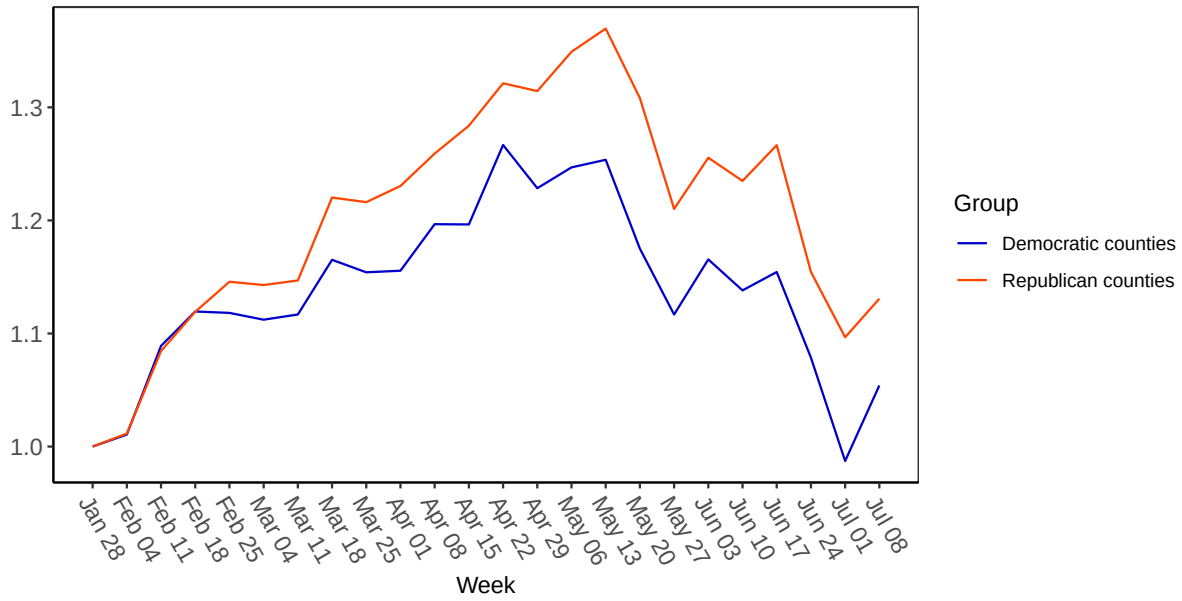
- You will now be asked to make a few predictions. Think carefully! We'll randomly select 10 participants for an accuracy reward. If you're selected, we'll pay you up to $100 depending on how accurate your prediction was. For example:
  - If your answer is exactly right, we'll give you $100
  - If your answer is 1% off, we'll give you $99
  - If your answer is 2% off, we'll give you $98
  - ...
  - If your answer is 50% off, we'll give you $50
  - etc.

All subjects see:

- We want to know how well you think the U.S. will limit the spread of the coronavirus in the next month. There had been 177,226 known cases of coronavirus in the U.S. by March 31. How many additional known cases will there be in the U.S. in the month of April?

- RealClearPolitics reports polling data on public approval of President Trump's handling of the coronavirus outbreak. What percent of people will say they approve of Trump's handling of the coronavirus outbreak on the latest poll that ends before April 30?

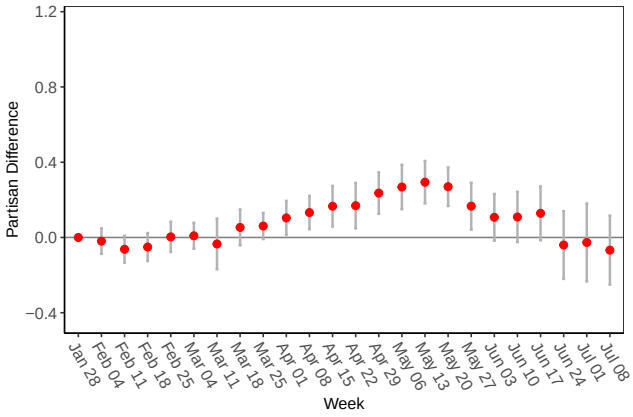
Appendix Figure A1: POI Visits in 2019



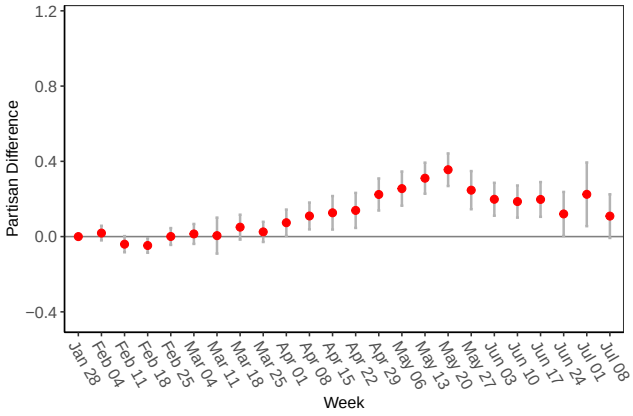
Note: Figure shows the aggregate number of POI visits (normalized to one) for 24 weeks starting on January 28, 2019 for Republican counties and Democratic counties. Republican counties are defined to be those whose 2016 Republican vote share is greater than the median vote share (66.4%) across the counties in our sample. Counties covering New York City, Kansas City, and Alaska are excluded from these counts, as in Figure 3 and as noted in Section A.1.1.

Appendix Figure A2: Partisan Differences in Social Distancing, 2019

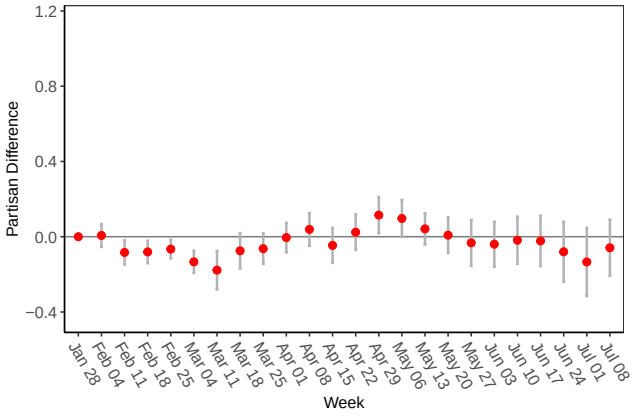
Panel A: Only County & Time FE



Panel B: Adds State-Time FE



Panel C: Adds Health + Econ + Weather Controls

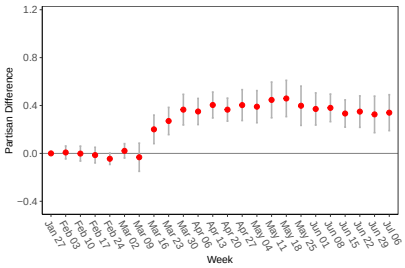


Note: Figure shows the estimated coefficients for county Republican vote share ρ_i on the log number of POI visits in the county as in Figure 4, except that 24 weeks of data from January 28, 2019 are used instead of January 27, 2020. For Panel A, only county and time fixed effects are included as controls. Panel B is the same as Panel A except state-time fixed effects replace the time fixed effects. Panel C is the same as Panel B except that health, economic, and weather covariates are included (flexibly), as described in the main text. The grey error bars indicate 95 percent confidence intervals constructed using standard errors clustered at the state-level.

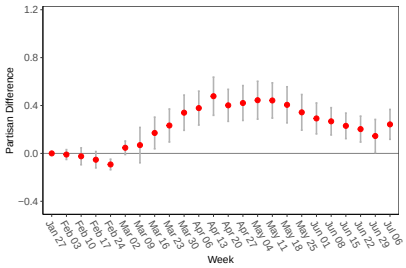
Appendix Figure A3: Partisan Differences in Social Distancing, Alternative Specifications

Panel A: Dropping Controls

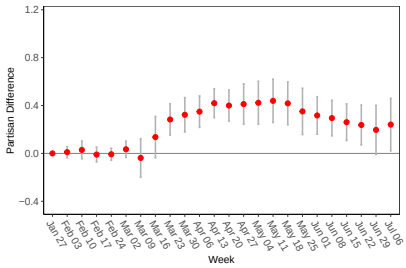
Drops Health Controls



Drops Economic Controls

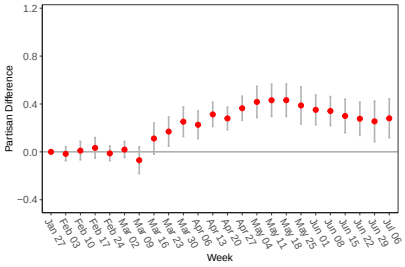


Drops Weather Controls

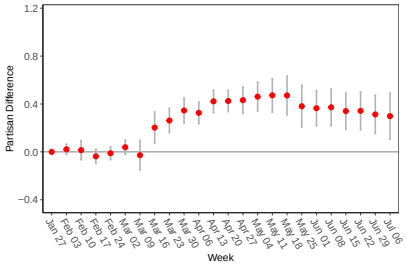


Panel B: Additional Specifications

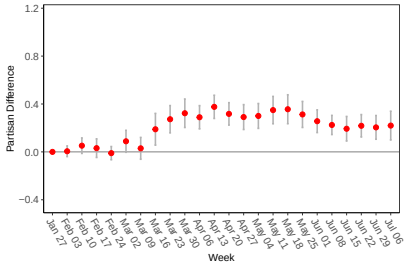
Drops State-Time FE



Linear Controls × Time FE

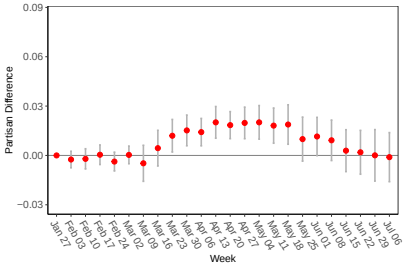


Adds Additional Controls

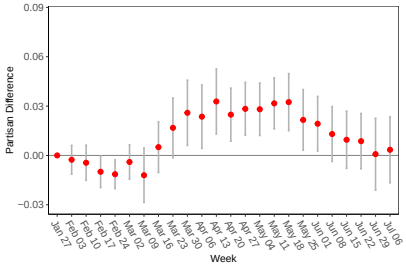


Panel C: Partisanship Indicators

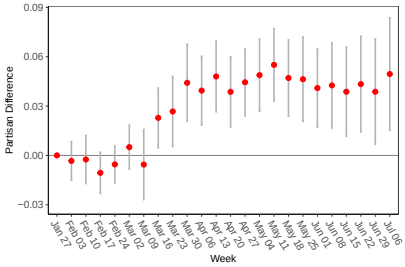
Above or Below Median



Top or Bottom Quartile



Top or Bottom Decile



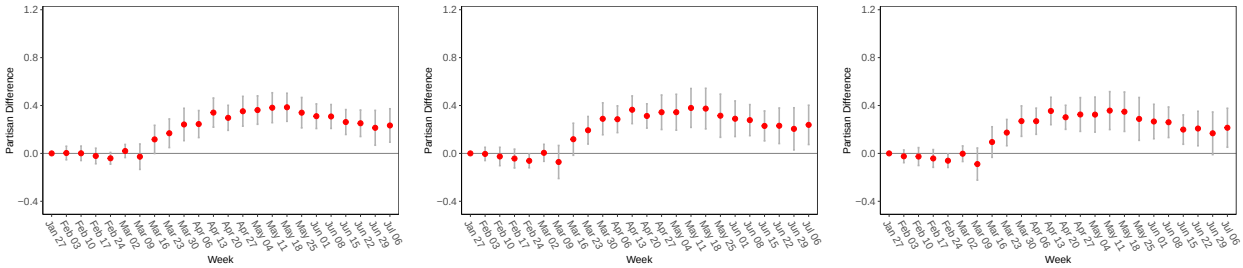
Note: Figure shows the estimated coefficients for county Republican vote share ρ_i on the log number of POI visits in the county. The specifications are analogous to our baseline in Panel C of Figure 4 except for the following deviations.

- Panel A: The first plot drops the health controls; the second plot drops the economic controls; and the third plot drops the weather controls.
- Panel B: The first plot drops state-time fixed effects (keeping county fixed effects and time fixed effects); the second plot interacts time fixed effects with linear versions of all controls, rather than with decile indicators; the third plot adds the following controls, each with time-varying decile indicators: share under age 18; shares with degrees in science+engineering, business, or arts+humanities; share of households with at least one vehicle; share of homes rented; share speaking only English at home; share with health insurance; shares commuting by auto, taxi, cycle, walking, or without commute; share enrolled in grad+professional school; share citizens; share married; shared of households with an age 60+ occupant; shares of households of size 1, 2, or 3-5; share of households which are a family. This third plot also adds as a control the log of POI visits (plus one) in the same week and county but during the previous year, included as a linear control interacted with time fixed effects.
- Panel C: The first plot defines partisanship ρ_i to be 1 if the Republican vote share is greater than the median and -1 otherwise; the second plot defines partisanship ρ_i to be 1 if the Republican vote share is in the top quartile, -1 if in the bottom quartile, and 0 otherwise; and the third plot defines partisanship ρ_i to be 1 if the Republican vote share is in the top decile, -1 if in the bottom decile, and 0 otherwise.

Appendix Figure A3: Partisan Differences in Social Distancing, Alternative Specifications cont.

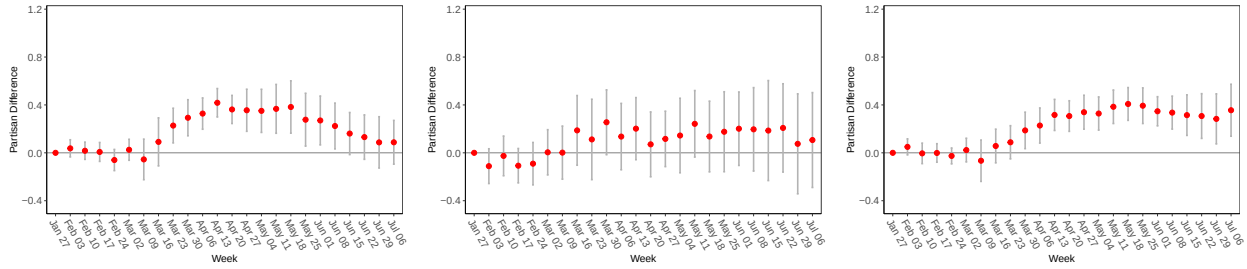
Panel D: Sample Restrictions by Population and State

Population above 3,000 | Population below 500,000 | Drops CA, WA, and NY



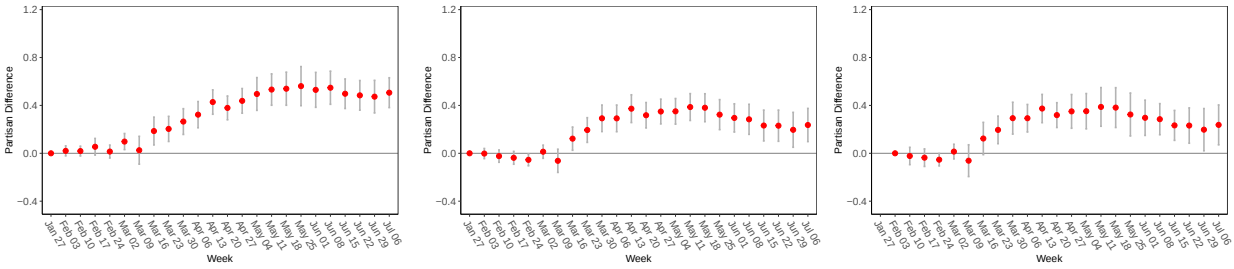
Panel E: Sample Restrictions by Republican Vote Share

Drops Bottom and Top Deciles | Above Median Vote Share | Below Median Vote Share



Panel F: Weighting, County Clustering, and Alternative Start Date

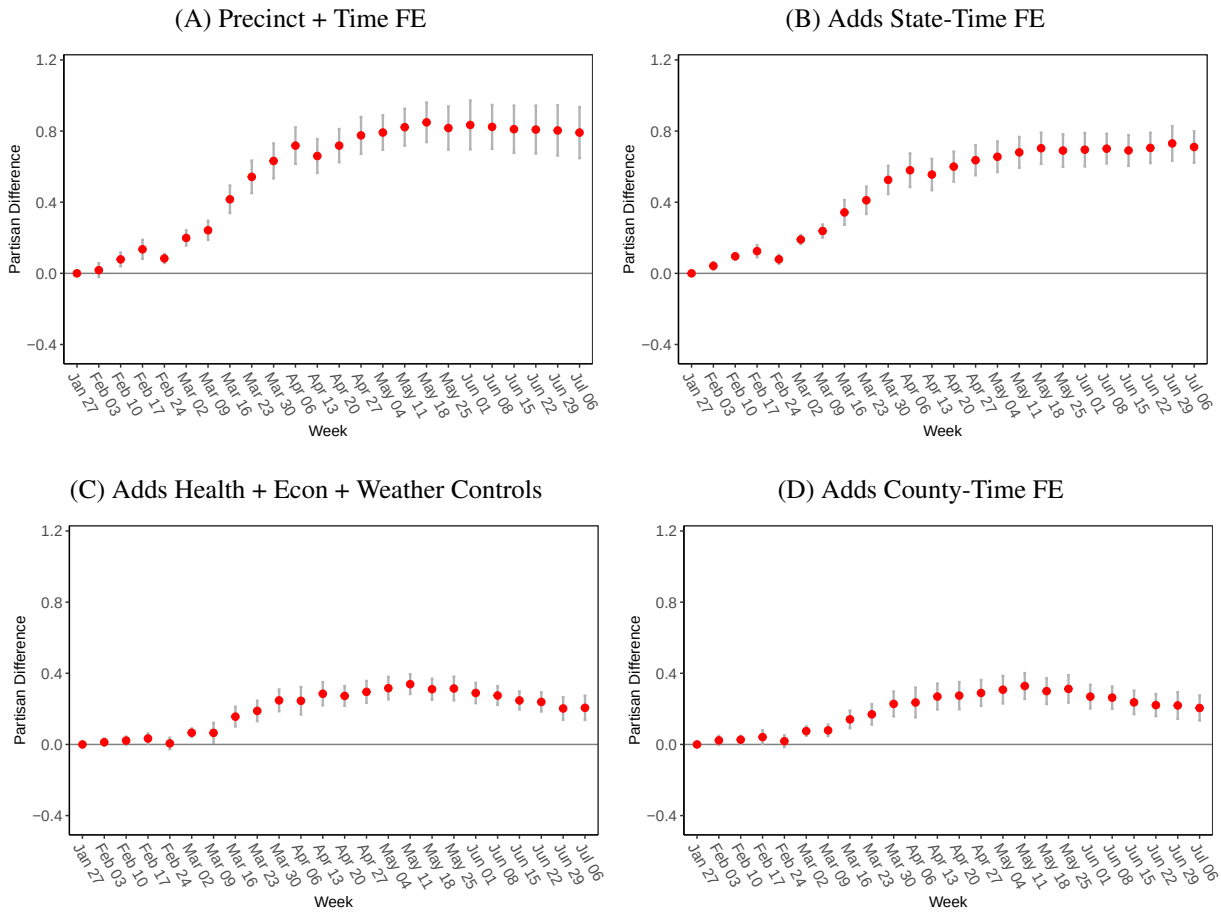
Weights by Population | Clusters SEs by County | Drops Week of January 27



Note: Figure shows the estimated coefficients for county Republican vote share ρ_i on the log number of POI visits in the county. The specifications are analogous to our baseline in Panel C of Figure 4 except with the following deviations.

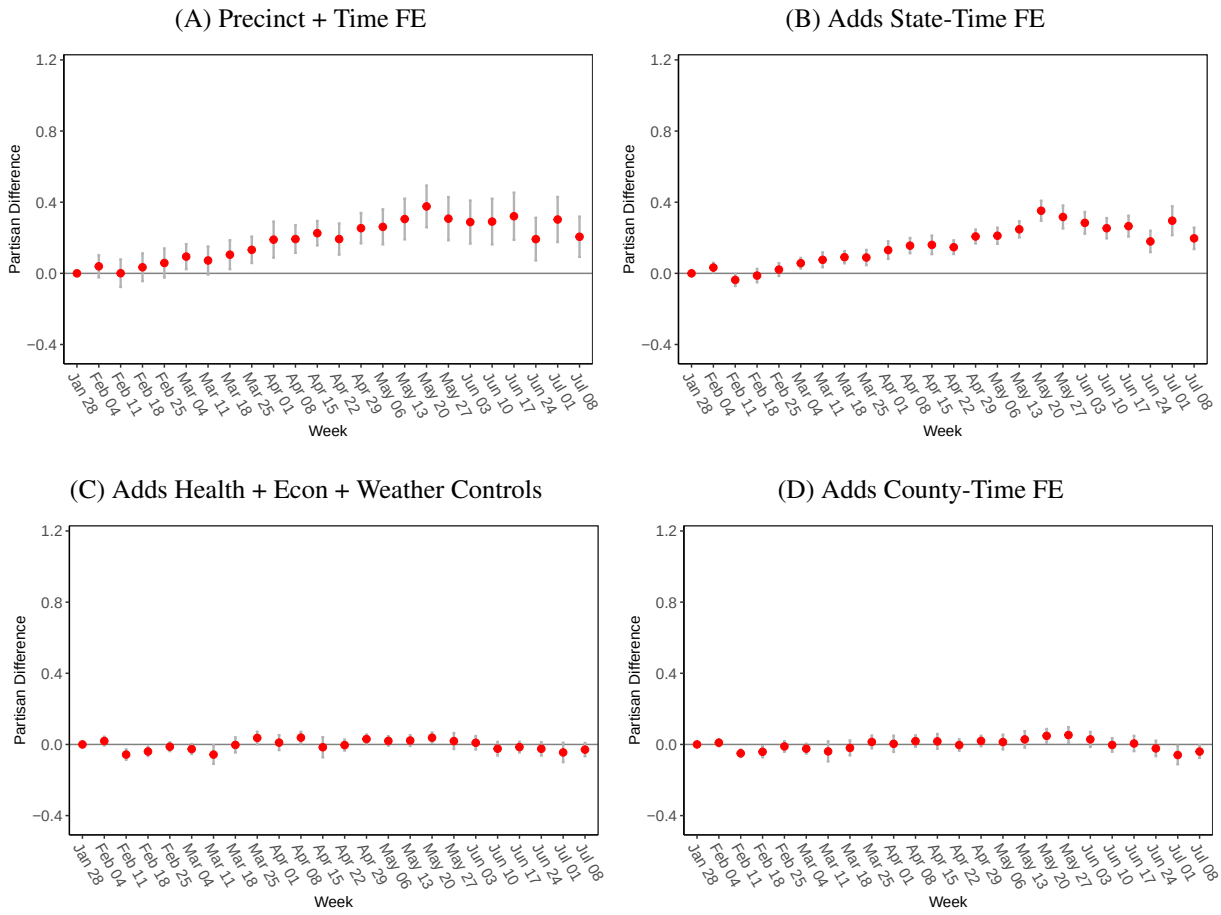
- Panel D: The first plot only keeps counties with a population above 3,000 (dropping 5.1 percent of counties in our sample); the second plot only keeps counties with a population below 500,000 (dropping 4.1 percent of counties in our sample); the third plot drops California, Washington, and New York.
- Panel E: The first plot drops counties for which the Republican vote share was in the bottom or top decile; the second plot keeps counties for which the Republican vote share is greater than the median; and the third plot keeps counties for which the Republican vote share is less than or equal to the median.
- Panel F: The first plot weights observations by the county's population. The second plot clusters standard errors at the county level. The third plot drops the week of January 27 and normalizes the estimates relative to the week of February 3.

Appendix Figure A4: Partisan Differences in Social Distancing, Precinct



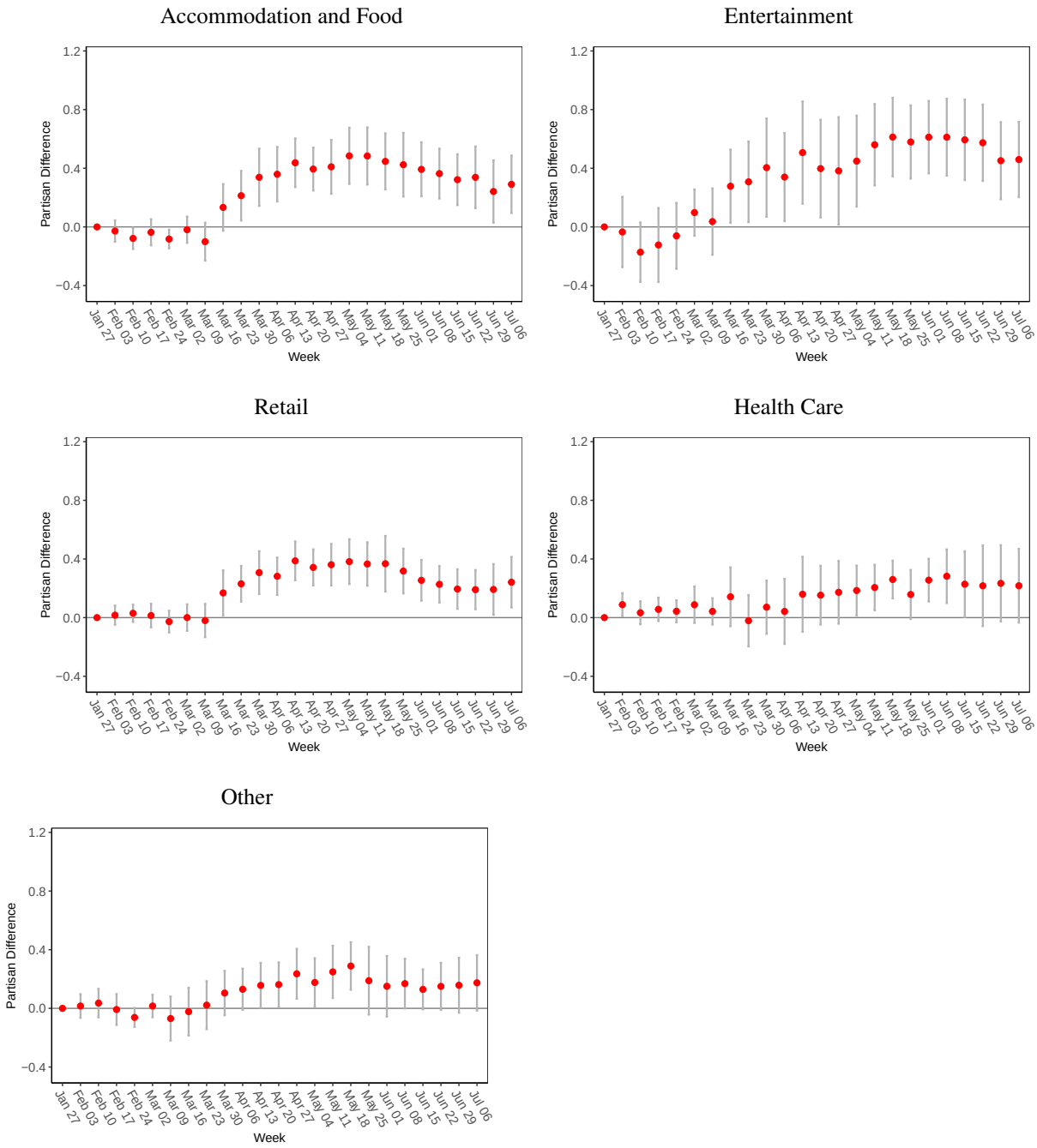
Note: Figure shows the estimated coefficients for precinct partisanship ρ_i on the log number of POI visits in the precinct using the specification outlined in the main text. For Panel A, only precinct and time fixed effects are included as controls. Panel B is the same as Panel A except state-time fixed effects replace the time fixed effects. Panel C is the same as Panel B except that health, economic, and weather covariates are included (flexibly), as described in the main text. Panel D is the same as Panel C except that county-time fixed effects replace the state-time fixed effects. The county-level COVID-19 controls are also subsumed in this specification. The grey error bars indicate 95 percent confidence intervals constructed using standard errors clustered at the state-level. See footnote 14 for limitations regarding this precinct-level analysis.

Appendix Figure A5: Partisan Differences in Social Distancing, Precinct 2019



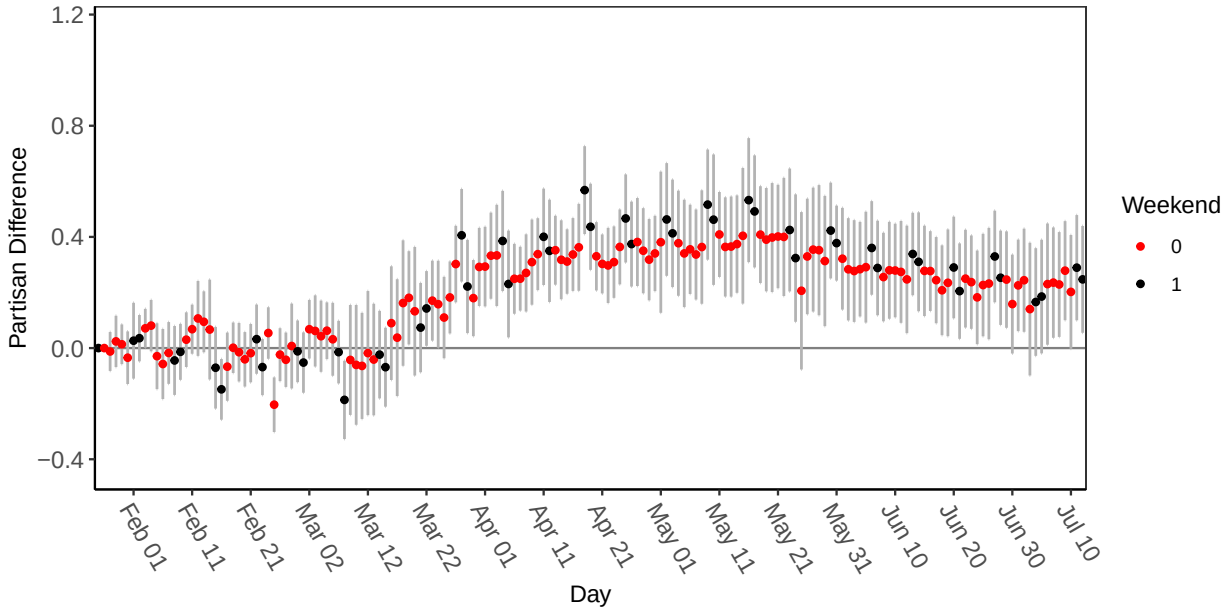
Note: Figure shows the estimated coefficients for precinct partisanship ρ_i on the log number of POI visits in the precinct. The figure mirrors Appendix Figure A4, except that 24 weeks of data from January 28, 2019 are used instead of January 27, 2020. For Panel A, only precinct and time fixed effects are included as controls. Panel B is the same as Panel A except state-time fixed effects replace the time fixed effects. Panel C is the same as Panel B except that health, economic, and weather covariates are included (flexibly), as described in the main text. Panel D is the same as Panel C except that county-time fixed effects replace the state-time fixed effects, the county-level COVID-19 controls are also dropped in this specification. The grey error bars indicate 95 percent confidence intervals constructed using standard errors clustered at the state-level. See footnote 14 for limitations regarding this precinct-level analysis.

Appendix Figure A6: Partisan Differences in Social Distancing by 2-Digit NAICS Code Industry



Note: Figure shows the estimated coefficients for county Republican vote share ρ_i on the log number of POI visits in the county after restricting POI visits to various 2-digit NAICS codes. The NAICS code groups are: Accommodation and Food (NAICS 72), Entertainment (NAICS 71), Retail Trade (NAICS 44 and 45), Health Care (NAICS 62), and Other Industries (All NAICS codes not previously used). The same controls are used as in Panel C of Figure 4. The grey error bars indicate 95 percent confidence intervals constructed using standard errors clustered at the state-level.

Appendix Figure A7: Partisan Differences in Social Distancing, Daily

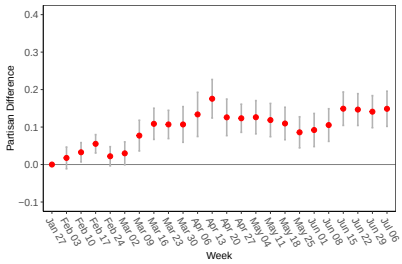


Note: Figure shows the estimated coefficients for county Republican vote share ρ_i on the log number of POI visits in the county. The same controls as in Panel C of Figure 4 are used except that state-time fixed effects occur at the day level, and we add separate county fixed effects for weekdays vs. weekends so that weekday and weekend series are normalized separately. The grey error bars indicate 95 percent confidence intervals constructed using standard errors clustered at the state-level.

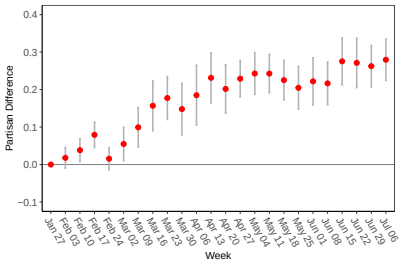
Appendix Figure A8: Partisan Differences in Social Distancing, Alternative Measures by Home Geography

Panel A: Alternative Measures by County

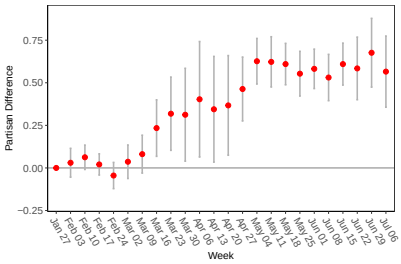
Log Devices Leaving Home



Log Stops in Non-Home CBGs

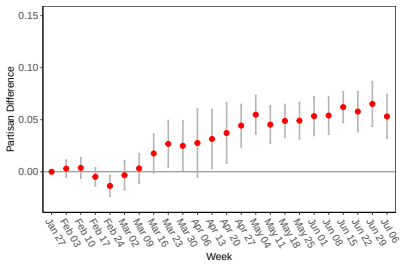


Log Median Time Away from Home

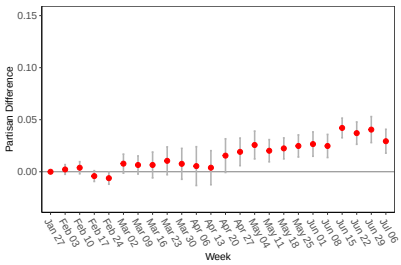


Panel B: Share of Devices Leaving Home by County

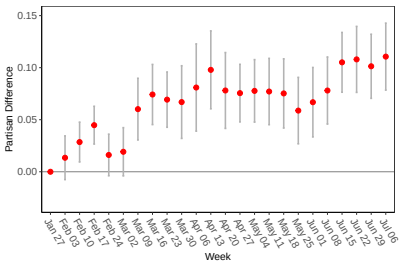
Share of Devices



Share of Candidate Devices

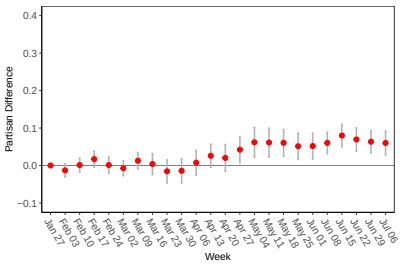


Share Adjusted for Attrition

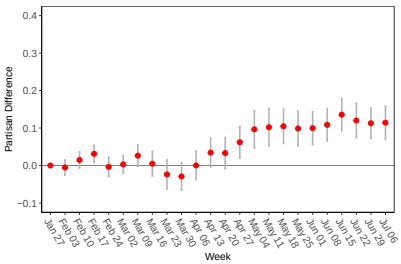


Panel C: Alternative Measures by Precinct

Log Devices Leaving Home



Log Stops in Non-Home CBGs



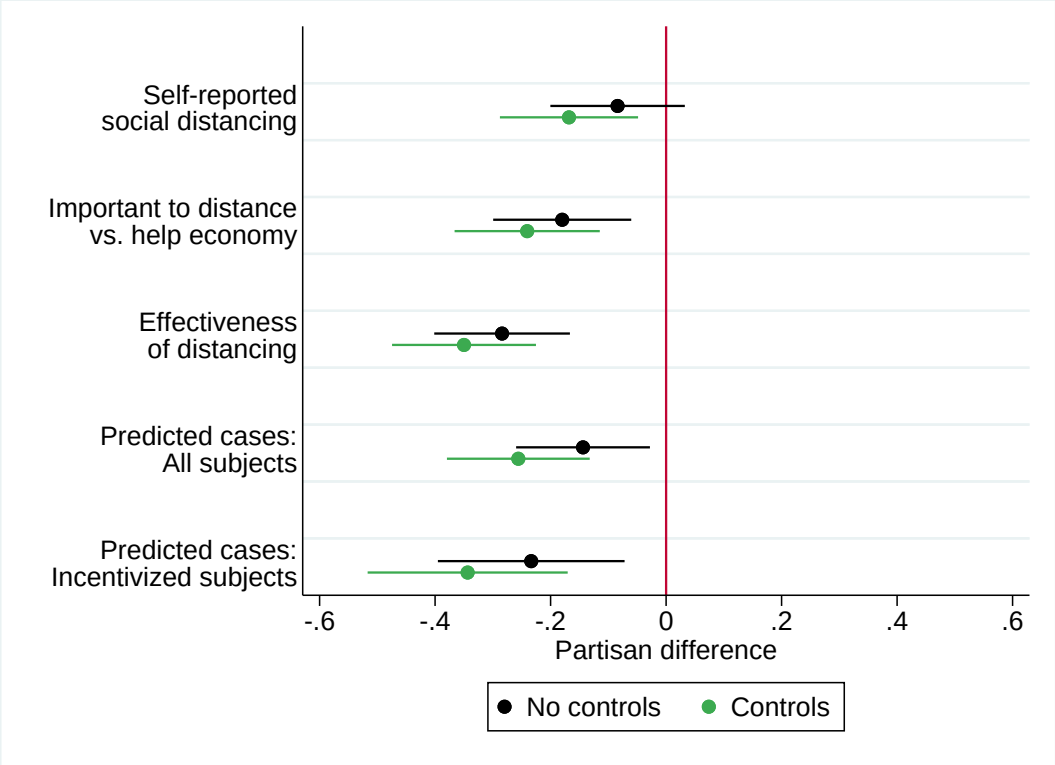
Note: Figure shows the estimated coefficients for Republican vote share ρ_i on alternative social distancing measures. The specifications are analogous to our baseline in Panel C of Figure 4 or Panel D of Figure A4 for county- and precinct-level regressions respectively, except that we replace log visits with the following alternative outcomes:

- Panel A: 'Log Devices Leaving Home' is the log of one plus the number of active devices in the panel minus the active devices never observed leaving their geohash-7 home. 'Log Stops in Non-Home CBGs' counts, by home county, the number of devices which stop in a given non-home census block group. We then sum across non-home census block groups and days to form our county-week outcome. 'Log Median Time Away from Home' calculates the median time a device is observed outside its geohash-7 home, by home census block group. We then take a device-weighted average across census block groups and days to form our county-week measure, and then take the log of this value.

- Panel B: This panel shows the share of devices which are observed outside their geohash-7 home, making different assumptions about attrition observed in the data. 'Share of Devices' is defined to be $1 - \frac{\text{home devices}}{\text{current device count}}$, where 'home devices' are active devices never observed leaving their geohash-7 home and 'current device count' is the number of active devices for the current week. 'Share of Candidate Devices' is similarly defined as $1 - \frac{\text{home devices}}{\text{candidate device count}}$, where 'candidate device count' is the number of devices regardless of activity. 'Share Adjusted for Attrition' is defined to be $1 - \frac{\max\{0, \text{home devices} + (\text{inititial device count} - \text{current device count})\}}{\text{initial device count}}$, where 'initial device count' is the number of active devices for the week of February 1.

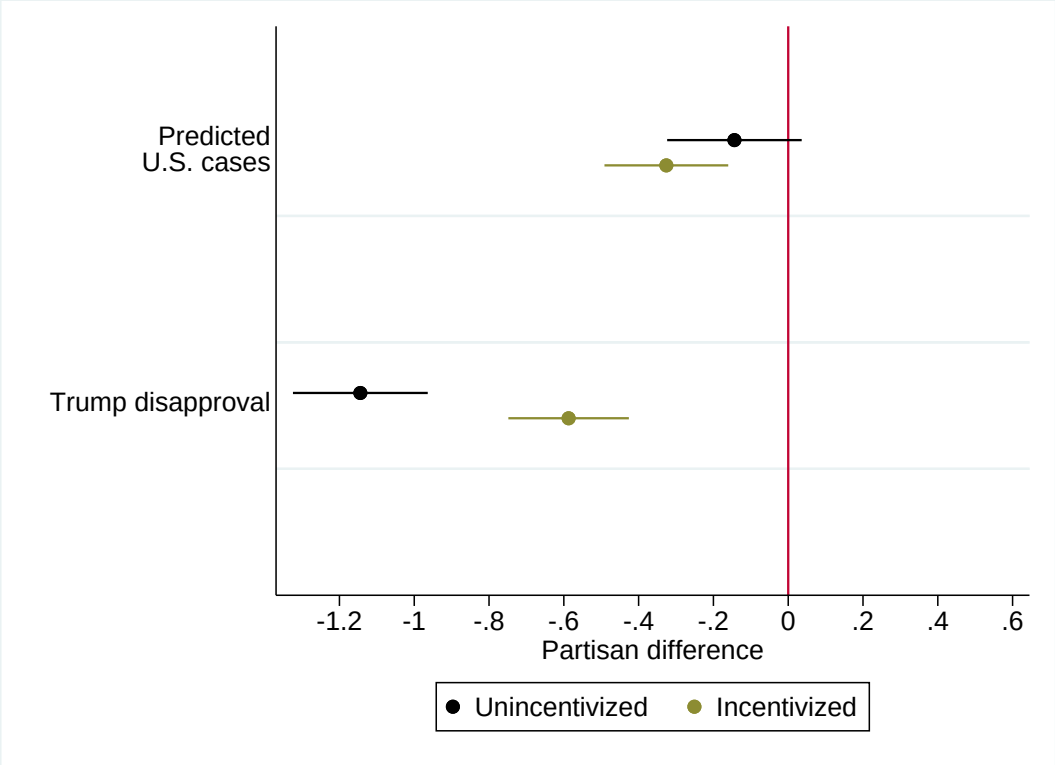
- Panel C: This panel produces precinct-level analogues of the first two plots in Panel A. Our specification matches Panel D of Figure A4 (including county-time fixed effects), and we map our original census block group-level social distancing measures to precincts using the method described in Section A.1.2. See footnote 14 for limitations regarding this precinct-level analysis.

Appendix Figure A9: Partisan Differences in Beliefs and Actions: Unweighted



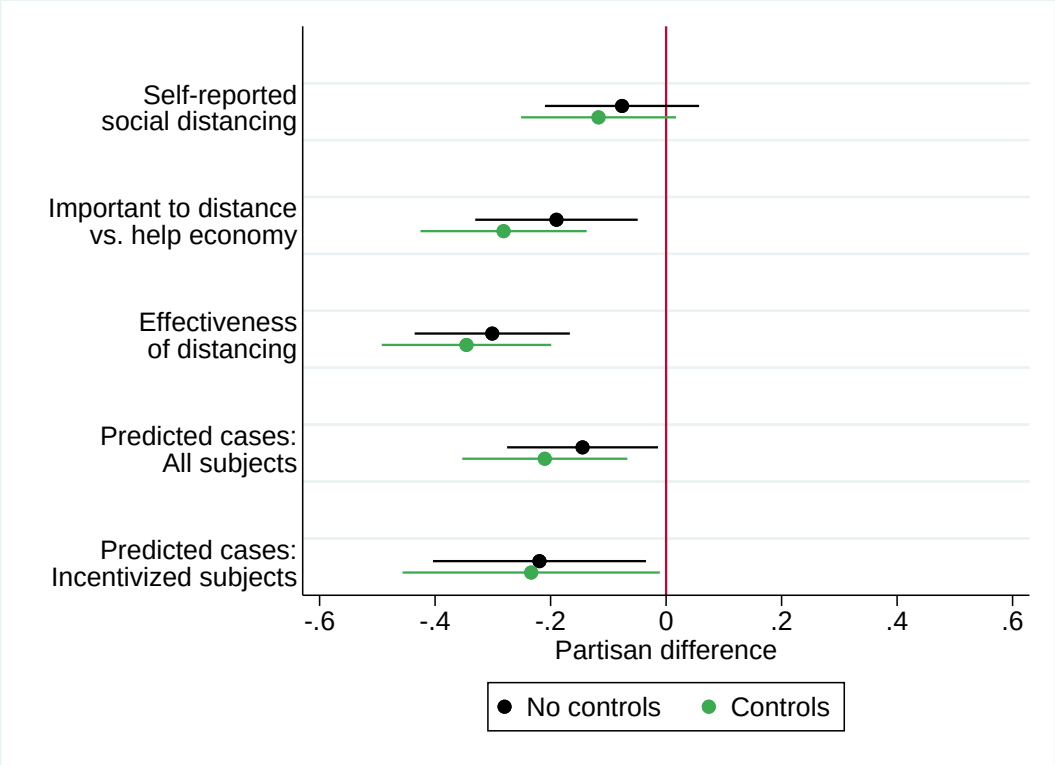
Note: Figure shows coefficient plots of regressing normalized measures of beliefs and actions on Republican party lean, without weighting observations. Negative values indicate less concern about COVID-19 or social distancing. Demographic controls are age, race, income, education, number of children, ZIP code logged population density, county-level deaths and cases, and state fixed effects. 2 percent of observations are set to the mean due to an invalid ZIP code. Self-reported social distancing is the percent reduction in contact with others over one month; effectiveness of distancing is the estimated likelihood of catching COVID-19 in one month without social distancing; importance of distancing vs. economy is subjects' perception of whether it is more important to go out and stimulate the economy versus staying in and preventing the spread of COVID-19; predicted cases are predictions about the number of new COVID-19 cases in the U.S. in April; incentivized subjects restrict to the subsample whose answers are incentivized. Error bars represent 95 percent confidence intervals.

Appendix Figure A10: Effect of Incentives on Beliefs



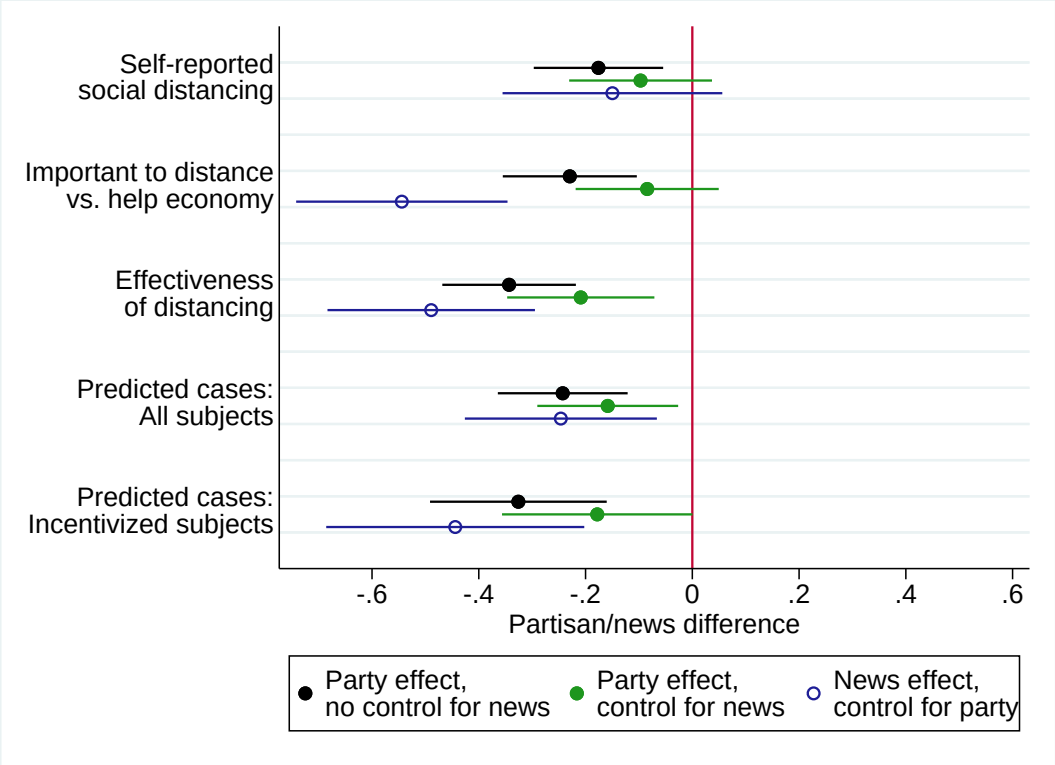
Note: Figure shows coefficient plots of regressing beliefs on Republican party lean, with and without incentives for getting close to the correct answer. Trump disapproval is a low-stakes question that is susceptible to partisan cheerleading (Bullock et al. 2015; Prior et al. 2015). These results show that predicting COVID-19 cases does not appear susceptible to the same behavior. Observations are weighted to mimic a representative sample as described in the text. Error bars represent 95 percent confidence intervals.

Appendix Figure A11: Partisan Differences in Beliefs and Actions: County Fixed Effects



Note: Figure shows coefficient plots of regressing normalized measures of beliefs and actions on our seven-point measure of partisan affiliation which ranges between 0 (Strongly Democratic) and 1 (Strongly Republican). Negative estimates indicate less concern about COVID-19 or social distancing. Demographic controls are age, race, income, education, number of children, log population at the ZIP code level, and county fixed effects. 21.5 percent of observations are dropped due to an invalid ZIP code or unique county. The remaining observations are weighted to mimic a representative sample as described in the text. Self-reported social distancing is the percent reduction in contact with others over one month; effectiveness of distancing is the estimated likelihood of catching COVID-19 in one month without social distancing; importance of distancing vs. economy is subjects' perception of whether it is more important to go out and stimulate the economy versus staying in and preventing the spread of COVID-19; predicted cases are predictions about the number of new COVID-19 cases in the US in April; incentivized subjects restrict to the subsample whose answers are incentivized. Error bars represent 95 percent confidence intervals.

Appendix Figure A12: Differences in Beliefs and Actions by Party and News Consumption



Note: Figure shows coefficient plots of regressing normalized measures of beliefs and actions on coronavirus news consumption and our our seven-point measure of partisan affiliation which ranges between 0 (Strongly Democratic) and 1 (Strongly Republican). Negative estimates indicate less concern about COVID-19 or social distancing. News consumption weights respondents' ratings to "How frequently are you getting news and information from each of the following sources about the coronavirus?" as specified in the Survey Details of the Appendix. All coefficients include demographic controls of age, race, income, education, number of children, log population density at the ZIP code level, and state fixed effects. 2 percent of observations are set to the mean due to an invalid ZIP code. For each outcome, the top plot does not control for news, the middle plot linearly controls for news, and the bottom plot shows the coefficient on news in the middle specification. Self-reported social distancing is the percent reduction in contact with others over one month; effectiveness of distancing is the estimated likelihood of catching COVID-19 in one month without social distancing; importance of distancing vs. economy is subjects' perception of whether it is more important to go out and stimulate the economy versus staying in and preventing the spread of COVID-19; predicted cases are predictions about the number of new COVID-19 cases in the U.S. in April; incentivized subjects restrict to the subsample whose answers are incentivized. Observations are weighted to mimic a representative sample as described in the text. Error bars represent 95 percent confidence intervals.