By Jing Huang, Brian T. Fisher, Vicky Tam, Zi Wang, Lihai Song, Jiasheng Shi, Caroline La Rochelle, Xi Wang, Jeffrey S. Morris, Susan E. Coffin, and David M. Rubin

DOI: 10.1377/hlthaff.2021.01072
HEALTH AFFAIRS 41, NO. 3 (2022): 445–453

The Effectiveness Of Government Masking Mandates On COVID-19 County-Level Case Incidence Across The United States, 2020

ABSTRACT Evidence for the effectiveness of masking on SARS-CoV-2 transmission at the individual level has accumulated, but the additional benefit of community-level mandates is less certain. In this observational study of matched cohorts from 412 US counties between March 21 and October 20, 2020, we estimated the association between county-level public masking mandates and daily COVID-19 case incidence. On average, the daily case incidence per 100,000 people in masked counties compared with unmasked counties declined by 25 percent at four weeks, 35 percent at six weeks, and 18 percent across six weeks postintervention. The beneficial effect varied across regions of different population densities and political leanings. The most concentrated effects of masking mandates were seen in urban counties; the benefit of the mandates was potentially stronger within Republican-leaning counties. Although benefits were not equally distributed in all regions, masking mandates conferred benefit in reducing community case incidence during an early period of the COVID-19 pandemic.

Jing Huang (jing14@pennmedicine.upenn.edu), University of Pennsylvania and Children's Hospital of Philadelphia, Philadelphia, Pennsylvania.

Brian T. Fisher, University of Pennsylvania and Children's Hospital of Philadelphia.

Vicky Tam, Children's Hospital of Philadelphia.

Zi Wang, Children's Hospital of Philadelphia.

Lihai Song, Children's Hospital of Philadelphia.

Jiasheng Shi, University of Pennsylvania and Children's Hospital of Philadelphia.

Caroline La Rochelle, Children's Hospital of Philadelphia.

Xi Wang, Children's Hospital of Philadelphia.

Jeffrey S. Morris, University of Pennsylvania.

Susan E. Coffin, University of Pennsylvania and Children's Hospital of Philadelphia.

David M. Rubin, University of Pennsylvania and Children's Hospital of Philadelphia.

As COVID-19 emerged in the US in early 2020, federal, state, and local authorities responded in a variety of ways to try to curb community transmission of the SARS-CoV-2, the coronavirus that causes COVID-19. One of these responses was the wearing of face coverings, or masks. By April 2020 the wearing of masks in public became the recommendation of the Centers for Disease Control and Prevention (CDC), but the federal government did not issue a mandate. Subsequently, many states and counties established their own mandates of who should wear masks and when, and how public masking would be enforced. Between April and September 2020, the timing of these masking mandates was highly variable throughout the country. By the fall of 2020, guidance from local health authorities converged so that universal masking in public locations, alongside recommendations for social distancing, had become the norm in most regions.

The evidence for the effectiveness of masking at the individual level has been strong[1–4] for viruses that, similar to SARS-CoV-2, spread principally through close contact with an infected person.[5–9] However, the incremental benefit of community-level mandates beyond public health guidance has been less certain. Studies that have assessed the influence of community-directed masking mandates often compare states or countries that are quite heterogeneous in demography.[10–12] Smaller studies of county-level masking mandates have relied on pre-post comparisons of case incidence[13] or have compared counties without adjusting for the propensity to mandate[14–17]—approaches that may be fraught with bias from secular trends[18] and unmeasured confounding.

Although randomized clinical trials, considered the "gold standard" for evaluating intervention effects, might not be practical for studying the effect of masking mandates, the variation in timing of the enactment of masking mandates at the county level during the early period of the pandemic offers an opportunity to assess the impact of county-level mandates on subsequent case incidence via a pragmatic quasi-experimental comparative effectiveness study. In this observational study we strove to emulate a trial that randomly assigned counties to a masking mandate versus no mandate by using a matched study design[19] to compare the effectiveness of county-level masking mandates on subsequent county-level case incidence during the period of April–August 2020. The primary objective was to determine the impact of masking mandates on the trajectory of county-level COVID-19 case incidence and to further assess the impact of masking mandates across regions of different urbanicity and political leaning.

Study Data And Methods

COUNTY ELIGIBILITY AND CASE INCIDENCE DATA Data on COVID-19 county-level case incidence for all US counties were obtained from USAFacts.org.[20] To create a representative sample of US areas affected by the pandemic, we designed the eligibility criteria for this analysis to include either densely populated areas or counties with moderate-to-substantial community transmission, including areas affected by the second wave of the pandemic in the summer of 2020. Starting in April 2020, thirty-eight US states and Washington, D.C., issued state-level masking mandates, and thirty-four (nearly 90 percent) of them started before September 2020.[21] To be considered eligible, a county had to have at least three consecutive days of daily case count exceeding five as of August 31, 2020, and meet at least one of the following criteria: contain at least one city with a population exceeding 100,000 people, contain a state capital, be the most populated county in the state, or have an average daily case incidence that exceeded twenty during July 1–August 31, 2020. The incidence threshold of twenty was selected on the basis of empirical distribution of case counts in the observation period. A total of 569 counties, representing forty-six states and Washington, D.C., were eligible for this study.

EXPOSURE The exposure of interest was the date of initiation of a public masking mandate for a county, issued by either the state or county government (time zero). These data were captured from a review of local news reports, executive orders, municipality websites, the COVID Analysis and Mapping of Policies database at Georgetown University,[22] and the *New York Times* state-level reopening tracker.[23] For a county to be considered as exposed for this analysis (that is, under a masking mandate), the mandate order needed to be posted between April 1 and August 31, 2020, and—at minimum—require that masks be worn by employees and customers inside public-facing buildings. If the state or county did not have a masking mandate but we identified any city in the county that had one, that county was deemed to have a masking mandate for this analysis. The first and last masking mandates among the 569 counties were observed April 4 and August 25, 2020. Case incidence data were collected from March 21 to October 20, 2020, to allow for a preintervention observation period of a minimum of two weeks and a post-intervention period of at least eight weeks.

MATCHING ALGORITHM Each exposed county was matched one to one with a county from a risk set of unexposed counties that was assembled from within the same region—but not within the same metropolitan area—as the exposed county, based on information from the Bureau of Economic Analysis.[24] To be eligible to serve as an unexposed county match to an exposed county, a county had to have no masking mandates at least three weeks past the date of the mandate for the exposed county. When more than one unexposed county was available from a risk set, a nearest-neighbor matching was used to select the unexposed match on the basis of the following variables: population density, total population, presidential election voting patterns in 2016, and case incidence and instantaneous reproduction number (R_t) in the two weeks before time zero. The matching procedure proceeded chronologically through each exposed county to identify a corresponding risk set of potential unexposed counties. Because risk sets were defined on the basis of the timing of making mandates, unexposed counties could be matched to multiple exposed counties, and an unexposed county could later become an exposed county (online appendix exhibit S1).[25]

Population density and total population were log-transformed before analysis because of the significant skewness for the largest cities. County-level daily incidence (cases per 100,000) and R_t were categorized as the twenty-fifth, fiftieth, and seventy-fifth percentiles for the two weeks before time zero.[26–28] R_t was estimated using the method of Anne Cori and colleagues,[26] with a moving average window of three days, as commonly used in models of transmission of SARS-CoV-2 and other communicable diseases.[29–34] The generation time was assumed to follow a gamma distribution with a mean of 7.5 days and standard deviation of 3.4 days, according to a previous

Masking mandates likely conferred benefit in reducing community transmission rates and case incidence during the initial months of the pandemic.

epidemiological survey of the first 425 COVID-19 cases in Wuhan, China.[34] Matching on these variables was based on the nearest Euclidean distance.[35,36] All matching variables were standardized by mean and standard deviation to give them equal weight.

OUTCOME The primary outcome was the county-level daily COVID-19 case incidence over the course of six weeks after masking mandate initiation. The daily incidence (cases per 100,000) was captured from time zero to six weeks postmandate (primary and secondary analyses) and from time zero to eight weeks postmandate (sensitivity analysis in a subset of counties). The reported daily incident COVID-19 case counts were smoothed using a three-day rolling average to reduce extra noise that was a result of batch reporting.

COVARIATES County-level covariates with the potential to confound the analysis results were considered for each model, including social distancing, population density, wet-bulb temperature, proportion of county residents with diabetes, and proportion of county residents earning less than 200 percent of the federal poverty level. Social distancing was quantified using daily cellphone movement provided by Unacast (which collects and aggregates human mobility data),[37] measuring the percentage change in visits to nonessential businesses within each county compared with visits in a four-week prepandemic baseline period between February 10 and March 8, 2020. Regarding wet-bulb temperature, it has been demonstrated that humidity and temperature both play a role in the seasonality of influenza, and earlier studies also reported significant effect of temperature and humidity on SARS-CoV-2 transmission.[34,38,39] For this reason, we used wet-bulb temperature, a metric that captures the complex thermodynamic relationship of temperature and humidity, has been shown to predict human health events with more precision than temperature and humidity separately, and avoids the associated problem of collinearity.[40,41] We used a rolling average of the social distancing metric and wet-bulb temperatures from four to fourteen days before case identification to account for the incubation period of COVID-19.[34] Nonlinear effects for wet-bulb temperatures, obtained from National Oceanic and Atmospheric Administration Local Climatological Data,[42] were included, as described in previous work modeling the effect of climate on SARS-CoV-2 transmission.[34] Diabetes and poverty covariates were abstracted from the Census Bureau's American Community Survey,[43] Behavioral Risk Factor Surveillance System,[44] ArcGIS Business Analyst,[45] and Multi-Resolution Land Characteristics Consortium.[46] County-level covariates were standardized by mean and standard deviation before analysis.

PRIMARY ANALYSIS Data were assessed for outliers relative to the R_t before analysis. Days with an R_t value outside of the 2.5–97.5 percentiles (R_t of 0.33 at the 2.5 percentile and 3.51 at the 97.5 percentile) were excluded from the analysis. Generalized linear mixed effects models were fit in a matched pair analysis to evaluate the association between exposures and outcomes, using a log link function while adjusting for covariates. Hierarchical random intercept and slope were used to account for correlations within the matched group and each individual county separately. We modeled the nonlinear trajectories of disease transmission over time, using cubic B-spline functions with no interior knots, and we compared trajectories between exposed and unexposed counties, using interactions between time and exposure. The time variable was defined as the difference between the date of case identification and the date of onset for the masking mandate. The adjusted ratios of moment-in-time case counts were estimated at two, four, and six weeks and averaged over the course of six weeks after mandate initiation.

SECONDARY ANALYSIS The ratio of daily case counts was also assessed across two subgroups. The first subgroup analysis examined counties by presidential voting behavior in the 2016 election,[47] contrasting counties that voted majority Republican (more than 50 percent of the vote went to the Republican presidential candidate) versus counties voting majority Democrat (up to 50 percent of the vote went to the Republican candidate). The second examined counties with high population density, which included more urban settings with 200–2,000 people per square mile, versus those with low population density, which included the most rural or suburban settings in which population density was lower than 200 people per square mile.

SENSITIVITY ANALYSIS Our analysis of model robustness considered different criteria for selecting unexposed counties during matching. First, we removed pairs whose intercounty distance was greater than 1,000 miles to reduce heterogeneity between counties (reducing our sample by 3.7 percent of matched pairs), as well as counties that were within 100 miles of each other to reduce interdependence of effects within metropolitan areas (reducing our sample by 9.4 percent of matched pairs). Second, we examined the additional longitudinal effects of extending the postmandate period to eight weeks. For this analysis, the minimum time to be without a masking mandate to be considered an unexposed county was extended from three to five weeks after mandate inception in the exposed county. We also added constraints to matching with replacement to ensure that a selected unexposed county could not be chosen within eight weeks of its inclusion in the study. By doing this, we could still use a county multiple times as a control, but there were no overlapping periods included within counties. This analysis reduced the matched pairs to 45.9 percent of the original matched pairs.

LIMITATIONS Our approach was not without limitations. First, unobserved heterogeneity across counties may have contributed to residual confounding. For example, exposed and unexposed counties could have had differential rates of local gathering events or different nonpharmaceutical interventions such as business closures. However, a tipping-point analysis,[48,49] described in the appendix,[25] revealed that the magnitude of effect would have needed to be quite large to overturn the results. It would also be difficult to invoke other community or state mitigation strategies as potential confounders, given that many businesses were reopening during this period but schools had not yet opened across the country. Second, because publicly available information on county mask mandates was sometimes contradictory, unclear, or incomplete, misclassification of the exposure was possible, which could bias estimates of effect. To address this concern, we repeated the process for county exposure assignment using updated public records available on local municipality websites and local news media. This review did identify some misclassification of exposure, but repeated analyses with updated exposure classification revealed that the impact of this bias was likely small and in the direction of biasing toward the null. Similarly, our measurement of social distancing did not fully capture the impact of business reopenings during this period; thus, we might have biased results to the null.

Third, exposed and unexposed counties could have sustained differential SARS-CoV-2 transmission rates related to higher or lower incidence rates in neighboring counties and regions. However, our findings were robust in sensitivity analyses that excluded matched pairs that were either close or distant in proximity. Fourth, it is possible that unexposed counties could have instituted a masking mandate in the latter days of the follow-up period, which could have biased estimates toward the null. In the sensitivity analysis that extended the follow-up period to eight weeks postmandate, we extended the criterion for no mask requirement in the unexposed county from three to five weeks; the analysis revealed similar findings. Finally, our results might not be transportable to counties with smaller population densities, given that most of the analyzed counties have higher population density, or generalizable to later periods, when higher population immunity was attained through either natural infection or acquired vaccination.

Study Results

The matching procedure, which leveraged county exposure assignment based on our review of public records available on local municipality websites and local news media, produced 351 matched county pairs identified across 412 unique counties (56.0 percent of the US population). A list of the matched pairs and daily COVID-19 case incidence per 100,000 people during the two weeks before masking mandates in the mandate groups is in exhibits S2 and S3.[25] Exhibit 1 demonstrates the distribution of baseline variables between matched counties. Imbalance was still observed after matching. Variables that had more than 0.1 standardized difference were included in the regression model as baseline covariates. Estimates of regression coefficients are in appendix section A4.[25]

PRIMARY ANALYSIS Across time, the estimated daily case incidence per 100,000 people in masked counties was 75 percent of case incidence in unmasked counties (95% confidence interval: 67, 83) at four weeks and 65 percent (95% CI: 58, 74) at six weeks postintervention (exhibit 2 and appendix exhibit S4).[25] Aggregated across all six weeks, the average case incidence in masked counties was 82 percent (95% CI: 75, 90) of the case incidence experienced in unmasked counties. Exhibit 3 shows predicted postestimation trajectories of daily case counts per 100,000 people over the period

EXHIBIT 1

Comparison of county demographics and SARS-CoV-2 transmission rate at baseline (two weeks before masking mandate) between matched US counties without and with mandates, 2020

	Without mandate (129 counties, 351 baseline intervals)		With mandate (351 counties, 351 baseline intervals)		Standardized difference[a]
	Median	IQR	Median	IQR	
County characteristics					
Population density, people per square mile	380.0	190.6, 786.9	374.4	176.8, 909.9	0.21
Total population, in 1,000s	196.0	114.6, 396.8	253.3	133.2, 527.3	−0.26
Diabetes, % of population	10.0	8.5, 10.9	10.2	9.0, 11.5	−0.22
Low income, % of population[b]	33.7	27.0, 38.3	33.0	26.2, 38.6	0.05
Votes for Republican in 2016 presidential election, %	50.7	43.5, 64.8	48.0	38.1, 59.1	0.33
Over the course of the 2 weeks before intervention:					
Instantaneous reproduction number (R_t)					
25th percentile[c]	1.0	0.8, 1.2	1.0	0.8, 1.2	−0.05
50th percentile[c]	1.2	1.0, 1.4	1.2	1.0, 1.5	-0.07
75th percentile[c]	1.4	1.2, 1.8	1.5	1.2, 1.9	-0.13
Daily case counts per 100,000 people					
25th percentile[c]	7.4	3.3, 14.4	9.4	4.2, 17.3	−0.19
50th percentile[c]	9.1	4.8, 18.3	11.7	5.9, 21.8	−0.17
75th percentile[c]	11.5	6.4, 22.9	14.6	7.5, 26.5	−0.16
Social distancing					
25th percentile[c]	−0.22	−0.56, −0.10	−0.27	−0.59, −0.14	0.15
50th percentile[c]	−0.20	−0.53, −0.09	−0.26	−0.58, −0.12	0.14
75th percentile[c]	−0.19	−0.51, −0.07	−0.25	−0.56, −0.11	0.15
Daily wet-bulb temperature , degrees C[d]					
25th percentile[c]	17.8	10.5, 20.5	17.7	8.9, 20.8	0.02
50th percentile[c]	18.3	10.9, 21.0	18.8	9.7, 21.4	−0.02
75th percentile[c]	19.1	11.5, 21.8	20.1	10.8, 22.2	−0.02

SOURCES All characteristics were obtained from the Census Bureau, American Community Survey, 2018 (see note 43 in text) except health data, which were obtained from the Centers for Disease Control and Prevention, Behavioral Risk Factor Surveillance System, 2017 (see note 44 in text); data on social distancing, which were obtained from Unacast (see note 37 in text); and data on wet-bulb temperature, which were obtained from the National Oceanic and Atmospheric Administration (see note 42 in text). **NOTES** Population density, diabetes, low income, social distancing, wet-bulb temperature, and baseline case count defined as the median case counts during two weeks before mandate were included in the analysis as covariates. Population size was not included because of its high correlation with population density. IQR is interquartile range. [a]The standardized difference was calculated by standardized difference in mean. [b]Low income was defined as percent of county residents with income less than 200 percent of the federal poverty level. [c]Percentiles of variables within each county over the course of 14 days before time zero. [d]Daily wet-bulb temperatures were calculated by averaging the hourly recordings from weather stations that contribute to the National Oceanic and Atmospheric Administration's Local Climatological Data.

of two weeks premandate to six weeks postmandate, conditioned on all counties having or not having the mandate. Time to peak incidence was reduced (sixteen versus twenty-two days between counties with mandate and those without), alongside a widening difference in daily case incidence over time in the setting of masking mandates.

SECONDARY ANALYSES Across subgroups, the daily case incidence per 100,000 people among Republican-leaning counties with a masking mandate was reduced to 66 percent (95% CI: 57, 76) and 61 percent (95% CI: 52, 73) relative to Republican-leaning counties without a masking mandate at four and six weeks after onset of the intervention, respectively (exhibit 2 and appendix exhibit S4).[25] The daily case incidence per 100,000 people among Democrat-leaning counties with a masking mandate was reduced to 79 percent (95% CI: 67, 93) and 65 percent (95% CI: 53, 80) relative to Democrat-leaning counties without a masking mandate at four and six weeks after intervention, respectively.

Urban counties were also uniquely influenced by masking mandates compared with more suburban and rural counties. The daily case incidence per 100,000 people for urban counties with a masking mandate was reduced to 71 percent (95% CI: 63, 81) and 64 percent (95% CI: 55, 74) relative to urban counties without a masking mandate at four and six weeks after intervention, respectively. This compares with moment-in-time effects among rural and suburban counties of 81 percent (95% CI: 66, 101) and 64 percent (95% CI: 51, 82) at four and six weeks after intervention, respectively. Postestimation adjusted daily case incidence per 100,000 people across subgroups is graphically presented in appendix exhibit S5.[25]

SENSITIVITY ANALYSES Findings remained robust to more stringent inclusion and exclusion criteria and matching procedures (appendix ex-

EXHIBIT 2

Ratios of daily COVID-19 case incidence per 100,000 people between matched counties with and without masking mandates, overall and by county type, 2020

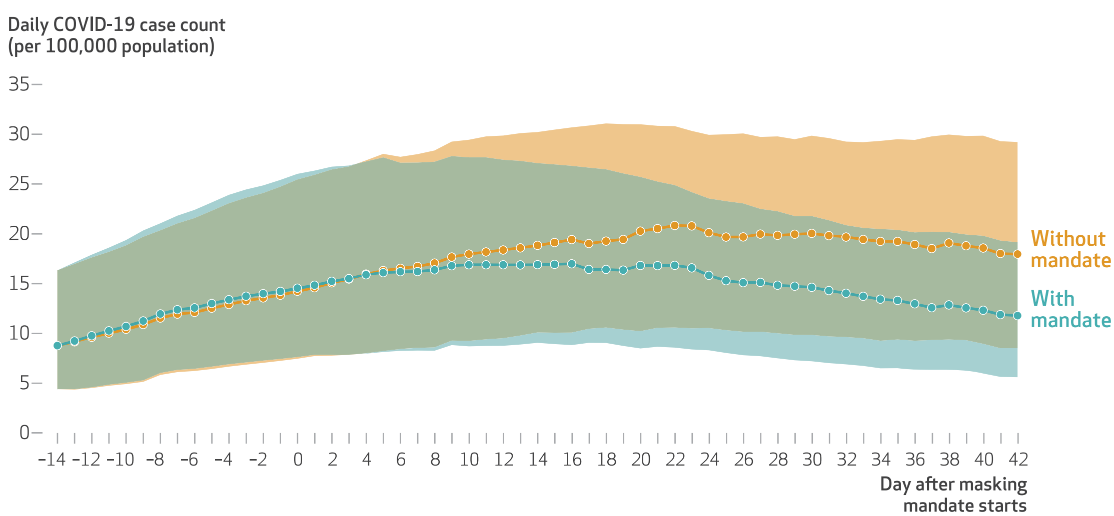
	Daily case counts ratio estimates between exposed and unexposed counties after intervention				
	At 2 weeks	At 4 weeks	At 6 weeks	Avg. over 6 weeks	*p* value[a]
All counties (351 pairs, 412 counties)	0.90	0.75	0.65	0.82	<0.001
Subgroup analysis[b]					
Voting behavior[c]					
Republican (136 pairs, 174 counties)	0.82	0.66	0.61	0.75	<0.001
Democrat (142 pairs, 178 counties)	0.92	0.79	0.65	0.85	<0.001
Population density[d]					
Urban (180 pairs, 224 counties)	0.85	0.71	0.64	0.79	<0.001
Suburban and rural (68 pairs, 94 counties)	1.00	0.81	0.64	0.89	<0.001

SOURCES USAFacts, US COVID-19 cases and deaths by state (see note 20 in text); COVID Analysis and Mapping of Policies, Data access (see note 22 in text); the *New York Times* state-level reopening tracker (see note 23 in text); Bureau of Economic Analysis (see note 24 in text); and MIT Election Data and Science Lab (see note 47 in text). **NOTES** Data were estimated from the primary analysis between March 21 and October 6, 2020. Ratios indicate the average case incidence in counties with a masking mandate compared to that in counties without one (exposed and unexposed). The analysis included 351 pairs of analysis windows drawn from 412 unique US counties between April and August 2020. The 95% confidence intervals of the estimates are in appendix exhibit S4 (see note 25 in text). [a]*p* values obtained from testing for null hypothesis of no significant difference in trajectory of incidence case over time between masked and unmasked counties, using Wald test on the regression coefficients of three interactions between masking mandate and polynomials of time (linear, quadratic, and cubic time). [b]Results stratified by political leaning and population density. [c]Voting behavior defined by the percentage of votes for the Republican candidate in the 2016 presidential election. The Republican stratum included pairs with more than 50 percent votes for the Republican candidate in both counties in the 2016 presidential election, and the Democratic stratum included pairs with up to 50 percent votes for the Republican candidate in both counties. [d]The urban stratum included pairs of counties with population density between 200 and 2,000 people per square mile, whereas the rural stratum included pairs of counties with population density less than 200 people per square mile.

hibit S6).[25] Extending the follow-up window to eight weeks also revealed similar findings, albeit in a smaller group of 161 matched pairs from 297 unique counties that were somewhat more densely populated, were from colder regions, had more significant social distancing, and had lower baseline transmission rates than the original sample (appendix exhibit S7).[25] At

EXHIBIT 3

Adjusted estimated daily COVID-19 case incidence per 100,000 people in US counties, under counterfactual scenarios assuming that all counties were either exposed or not exposed to masking mandates



SOURCES USAFacts, US COVID-19 cases and deaths by state (see note 20 in text); COVID Analysis and Mapping of Policies, Data access (see note 22 in text); and the *New York Times* state-level reopening tracker (see note 23 in text). **NOTES** The analysis included all of the analysis windows from 412 unique US counties between April and August 2020. "Without mandate" assumes that a mandate would not have been implemented, and "With mandate" assumes that a mandate would have been implemented. Time is aligned at the intervention time (time zero). Solid lines indicate medians. Shaded areas indicate interquartile ranges. The plot of the mean is in appendix exhibit S5 (see note 25 in text).

eight weeks of follow-up, counties with a masking mandate had a daily case incidence per 100,000 individuals that was reduced to 57 percent (95% CI: 45, 72) of the daily case incidence per 100,000 people in counties without masking mandates (exhibit 4, appendix exhibits S8 and S9).[25] This was comparable to the reduction noted at six weeks from this analysis (58 percent; 95% CI: 48, 72); the overall case incidence reduction across eight weeks was 29 percent (95% CI: 17, 39). Aggregate reductions over eight weeks remained larger in Republican-leaning versus Democratic-leaning counties and in urban versus suburban/rural counties, although moment-in-time ratio estimates of case incidence between exposed and unexposed counties at six and eight weeks were similar in subgroup analyses of Republican- and Democratic-leaning counties. Case incidence reductions were attenuated in suburban and rural counties with population density below 200 people per square mile (exhibit 4, appendix exhibit S8).[25]

Discussion

In this matched comparison of US counties with and without masking mandates in the early months of the COVID-19 pandemic in the US, the mandates were associated with reduced case incidence six weeks after the onset of the mandates. Reduction in case incidence ranged from 11 percent to 25 percent during the six-week interval (exhibit 2), with peak impacts between three and four weeks after the initiation of mandates. The benefit of masking mandates was observed across all subgroups, with peak moment-in-time reductions noted at six weeks post-intervention. The effect was strongest among Republican-leaning counties, where reductions, by inference, reached 25 percent (95% CI: 14, 34) across six weeks, and among urban counties, where reductions, by inference, reached 21 percent (95% CI: 13, 29) by six weeks. Findings in the study were robust to a sensitivity analysis that extended the follow-up period to eight weeks. Moment-in-time case incidence reductions were sustained but plateaued between six and eight weeks in analysis of all counties and subgroups, including Republican-leaning and rural counties, whose moment-in-time risk reduction plateaued at 50 percent and attenuated from 53 percent to 50 percent, respectively, between six and eight weeks, respectively.

The design of this study, which matched on baseline community transmission and demographics, was a major strength. The study simulated a pragmatic trial by which to estimate the incremental effect of mandates on community COVID-19 transmission. A large observational study of counties across the United States in 2021 also suggested that masking mandates were beneficial in reducing SARS-CoV-2 transmission,[10] but the approach in our study was arguably stronger, as it reduced the potential influence of secular trends. By matching unexposed and exposed counties, this study reduced potential observed confounding, and the county-level random effects used in the analysis also

EXHIBIT 4

Sensitivity analysis: ratios of daily COVID-19 case incidence per 100,000 people between matched counties with and without masking mandates, overall and by county type, among a subset of counties that could be followed for at least 8 weeks, 2020

	Daily case counts ratio estimates between exposed and unexposed counties after intervention				
	At 2 weeks	At 4 weeks	At 6 weeks	At 8 weeks	Avg. over 8 weeks
All counties (161 pairs, 297 counties)	0.83	0.68	0.58	0.57	0.71
Subgroup analysis					
Voting behavior[a]					
Republican (47 pairs, 94 counties)	0.68	0.56	0.50	0.50	0.59
Democrat (70 pairs, 134 counties)	0.92	0.68	0.51	0.45	0.72
Population density[b]					
Urban (79 pairs, 147 counties)	0.82	0.59	0.47	0.50	0.65
Suburban and rural (28 pairs, 54 counties)	0.70	0.77	0.88	0.93	0.79

SOURCES USAFacts, US COVID-19 cases and deaths by state (see note 20 in text); COVID Analysis and Mapping of Policies, Data access (see note 22 in text); the *New York Times* state-level reopening tracker (see note 23 in text); Bureau of Economic Analysis (see note 24 in text); and MIT Election Data and Science Lab (see note 47 in text). **NOTES** County types based on a subset of counties that could be followed for at least 8 weeks in the sensitivity analysis between March 21 and October 20, 2020. Ratios indicate the average case incidence in counties with a masking mandate compared with that in counties without one (exposed and unexposed). The analysis included 161 pairs of analysis windows drawn from 297 unique US counties between March and October 2020. 95% confidence intervals of the estimates are in appendix exhibit S8 (see note 25 in text). [a]Voting behavior is defined by the percentage of votes for the Republican candidate in the 2016 presidential election. The Republican stratum included pairs with more than 50 percent votes for the Republican candidate in both counties in the 2016 presidential election, and the Democratic stratum included pairs with up to 50 percent votes for the Republican candidate in both counties. [b]The urban stratum included pairs of counties with population density between 200 and 2,000 people per square mile, whereas the rural stratum included pairs of counties with population density less than 200 people per square mile.

allowed estimation of between-county heterogeneity that was not explained by the observed variables. Another challenge in this study is that we were unable to examine the factors that mediate the benefit of public masking mandates. It is possible, for example, that mandates led to higher rates of sustained masking use in public locations; unfortunately, there were no longitudinal surveys with sufficient sampling across the counties to examine this influence, including whether fatigue over time may have contributed disproportionately to the attenuation in effects observed among some counties at eight weeks. This may make it more difficult as well to translate these findings to later periods of the pandemic, when fatigue with public health recommendations was likely heightened and when more transmissible variants (such as Delta and Omicron) were circulating across the population. It is also possible that mandates signal to the community that risk throughout the region is rising and thus may—at least temporarily—affect practices of social distancing or proclivity to stay home among worried residents. In that regard, the degree to which the social distancing measure in this analysis approximated individual behavior more broadly might also have led us to overfit models and somewhat bias results to the null. Finally, benefits of mandates may also vary by commitment of law enforcement, a variable we were unable to study.

Conclusion

This matched analysis, during a period of variable public masking orders across the US, suggests that on average, masking mandates likely conferred benefit in reducing community transmission rates and case incidence during the initial months of the COVID-19 pandemic. Although such benefits were not equally distributed in all regions, it appears that masking mandates may offer broad value in reducing community risk during periods of elevated SARS-CoV-2 transmission in the US. ■

Support for this research was received from the National Institutes of Health (Grant Nos. R01CA-178744, R01CA-244845, and TR-001878). The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health. [*Published online February 16, 2022.*]

NOTES

1 Doung-Ngern P, Suphanchaimat R, Panjangampatthana A, Janekrongtham C, Ruampoom D, Daochaeng N, et al. Case-control study of use of personal protective measures and risk for SARS-CoV 2 infection, Thailand. Emerg Infect Dis. 2020;26(11):2607–16.

2 Brooks JT, Butler JC. Effectiveness of mask wearing to control community spread of SARS-CoV-2. JAMA. 2021;325(10):998–9.

3 Howard J, Huang A, Li Z, Tufekci Z, Zdimal V, Van Der Westhuizen H-M , et al. An evidence review of face masks against COVID-19. Proc Natl Acad Sci. 2021;118(4):e2014564118.

4 Bundgaard H, Bundgaard JS, Raaschou-Pedersen DET, von Buchwald C, Todsen T, Norsk JB, et al. Effectiveness of adding a mask recommendation to other public health measures to prevent SARS-CoV-2 infection in Danish mask wearers: a randomized controlled trial. Ann Intern Med. 2021;174(3):335–43.

5 Ueki H, Furusawa Y, Iwatsuki-Horimoto K, Imai M, Kabata H, Nishimura H, et al. Effectiveness of face masks in preventing airborne transmission of SARS-CoV-2. mSphere. 2020;5(5):e00637–20.

6 Centers for Disease Control and Prevention, National Center for Immunization and Respiratory Diseases, Division of Viral Diseases. Science brief: community use of cloth masks to control the spread of SARS-CoV-2 [Internet]. Atlanta (GA): CDC; 2020 [updated 2021 Dec 6; cited 2022 Jan 26]. Available from: https://www.cdc.gov/coronavirus/2019-ncov/science/science-briefs/masking-science-sars-cov2.html

7 Nishimura H, Sakata S, Kaga A. A new methodology for studying dynamics of aerosol particles in sneeze and cough using a digital high-vision, high-speed video system and vector analyses. PLoS One. 2013;8(11):e80244.

8 Smither SJ, Eastaugh LS, Findlay JS, Lever MS. Experimental aerosol survival of SARS-CoV-2 in artificial saliva and tissue culture media at medium and high humidity. Emerg Microbes Infect. 2020;9(1):1415–7.

9 Liu Y, Ning Z, Chen Y, Guo M, Liu Y, Gali NK, et al. Aerodynamic analysis of SARS-CoV-2 in two Wuhan hospitals. Nature. 2020;582(7813):557–60.

10 Lyu W, Wehby GL. Community use of face masks and COVID-19: evidence from a natural experiment of state mandates in the US. Health Aff (Millwood). 2020;39(8):1419–25.

11 Gallaway MS, Rigler J, Robinson S, Herrick K, Livar E, Komatsu KK, et al. Trends in COVID-19 incidence after implementation of mitigation measures—Arizona, January 22–August 7, 2020. MMWR Morb Mortal Wkly Rep. 2020;69(40):1460–3.

12 Kaufman BG, Whitaker R, Mahendraratnam N, Smith VA, McClellan MB. Comparing associations of state reopening strategies with COVID-19 burden. J Gen Intern Med. 2020;35(12):3627–34.

13 Adjodah D, Dinakar K, Chinazzi M, Fraiberger SP, Pentland A, Bates S, et al. Association between COVID-19 outcomes and mask mandates, adherence, and attitudes. PLoS One. 2021;16(6):e0252315.

14 Van Dyke ME, Rogers TM, Pevzner E, Satterwhite CL, Shah HB, Beckman WJ, et al. Trends in county-level COVID-19 incidence in counties with and without a mask mandate—Kansas, June 1–August 23, 2020. MMWR Morb Mortal Wkly Rep. 2020;69(47):1777–81.

15 Karaivanov A, Lu SE, Shigeoka H, Chen C, Pamplona S. Face masks, public policies, and slowing the spread of COVID-19: evidence from Canada. J Health Econ. 2021;78:102475.

16 Guy GP Jr, Lee FC, Sunshine G, McCord R, Howard-Williams M, Kompaniyets L, et al. Association of state-issued mask mandates and allowing on-premises restaurant din-

ing with county-level COVID-19 case and death growth rates—United States, March 1–December 31, 2020. MMWR Morb Mortal Wkly Rep. 2021;70(10):350–4.
17 Dasgupta S, Kassem AM, Sunshine G, Liu T, Rose C, Kang GJ, et al. Differences in rapid increases in county-level COVID-19 incidence by implementation of statewide closures and mask mandates—United States, June 1–September 30, 2020. Ann Epidemiol. 2021;57:46–53.
18 Brauner JM, Mindermann S, Sharma M, Johnston D, Salvatier J, Gavenčiak T, et al. Inferring the effectiveness of government interventions against COVID-19. Science. 2021;371(6531):eabd9338.
19 Dagan N, Barda N, Kepten E, Miron O, Perchik S, Katz MA, et al. BNT162b2 mRNA Covid-19 vaccine in a nationwide mass vaccination setting. N Engl J Med. 2021; 384(15):1412–23.
20 USAFacts. US COVID-19 cases and deaths by state [Internet]. Seattle (WA): USAFacts; [last updated 2022 Jan 18; cited 2022 Jan 5]. Available from: https://usafacts.org/visualizations/coronavirus-covid-19-spread-map/
21 Guy GP Jr, Lee FC, Sunshine G, McCord R, Howard-Williams M, Kompaniyets L, et al. Association of state-issued mask mandates and allowing on-premises restaurant dining with county-level COVID-19 case and death growth rates—United States, March 1–December 31, 2020. MMWR Morb Mortal Wkly Rep. 2021;70(10):350.
22 COVID Analysis and Mapping of Policies. Data access [Internet]. Washington (DC): COVID AMP; 2021 [cited 2022 Jan 5]. Available from: https://covidamp.org/data
23 Lee JC, Mervosh S, Avila Y, Harvey B, Matthews AL, Gamio L, et al. See reopening plans and mask mandates for all 50 states. New York Times [serial on the Internet]. 2020 [last updated 2021 Jul 1; cited 2022 Jan 5]. Available from: https://www.nytimes.com/interactive/2020/us/states-reopen-map-coronavirus.html
24 Bureau of Economic Analysis. How is a metropolitan area defined? [Internet]. Washington (DC): BEA; 2008 Jan 10 [cited 2022 Jan 5]. Available from: https://www.bea.gov/help/faq/459
25 To access the appendix, click on the Details tab of the article online.
26 Cori A, Ferguson NM, Fraser C, Cauchemez S. A new framework and software to estimate time-varying reproduction numbers during epidemics. Am J Epidemiol. 2013;178 (9):1505–12.
27 Fraser C. Estimating individual and household reproduction numbers in an emerging epidemic. PLoS One. 2007;2(8):e758.
28 Grassly NC, Fraser C. Mathematical models of infectious disease transmission. Nat Rev Microbiol. 2008; 6(6):477–87.
29 Pan A, Liu L, Wang C, Guo H, Hao X, Wang Q, et al. Association of public health interventions with the epidemiology of the COVID-19 outbreak in Wuhan, China. JAMA. 2020; 323(19):1915–23.
30 Aylward B, Barboza P, Bawo L, Bertherat E, Bilivogui P, Blake I, et al. Ebola virus disease in West Africa—the first 9 months of the epidemic and forward projections. N Engl J Med. 2014;371(16):1481–95.
31 Faria NR, Quick J, Claro IM, Thézé J, de Jesus JG, Giovanetti M, et al. Establishment and cryptic transmission of Zika virus in Brazil and the Americas. Nature. 2017; 546(7658):406–10.
32 Cowling BJ, Ali ST, Ng TWY, Tsang TK, Li JCM, Fong MW, et al. Impact assessment of non-pharmaceutical interventions against coronavirus disease 2019 and influenza in Hong Kong: an observational study. Lancet Public Health. 2020;5(5):e279–88.
33 Flaxman S, Mishra S, Gandy A, Unwin HJT, Mellan TA, Coupland H, et al. Estimating the effects of non-pharmaceutical interventions on COVID-19 in Europe. Nature. 2020;584(7820):257–61.
34 Rubin D, Huang J, Fisher BT, Gasparrini A, Tam V, Song L, et al. Association of social distancing, population density, and temperature with the instantaneous reproduction number of SARS-CoV-2 in counties across the United States. JAMA Netw Open. 2020;3(7):e2016099.
35 Rubin DB. Matching to remove bias in observational studies. Biometrics. 1973;29(1):159–83.
36 Rosenbaum PR. Modern algorithms for matching in observational studies. Annu Rev Stat Appl. 2020;7: 143–76.
37 Unacast. Social distancing scoreboard [Internet]. New York (NY): Unacast; [cited 2022 Jan 5]. Available from: https://www.unacast.com/covid19/social-distancing-scoreboard
38 Shaman J, Kohn M. Absolute humidity modulates influenza survival, transmission, and seasonality. Proc Natl Acad Sci U S A. 2009;106(9): 3243–8.
39 Lowen AC, Mubareka S, Steel J, Palese P. Influenza virus transmission is dependent on relative humidity and temperature. PLoS Pathog. 2007;3(10):1470–6.
40 Cheng Y-T, Lung S-CC, Hwang J-S. New approach to identifying proper thresholds for a heat warning system using health risk increments. Environ Res. 2019;170:282–92.
41 Ross ME, Vicedo-Cabrera AM, Kopp RE, Song L, Goldfarb DS, Pulido J, et al. Assessment of the combination of temperature and relative humidity on kidney stone presentations. Environ Res. 2018;162:97–105.
42 National Oceanic and Atmospheric Administration, National Centers for Environmental Information. Data tools: local climatological data (LCD) [Internet]. Washington (DC): NOAA; [cited 2022 Jan 5]. Available from: https://www.ncdc.noaa.gov/cdo-web/datatools/lcd
43 Census Bureau. American Community Survey (ACS) [Internet]. Washington (DC): Census Bureau; [cited 2022 Jan 5]. Available from: https://www.census.gov/programs-surveys/acs
44 Centers for Disease Control and Prevention. Behavioral Risk Factor Surveillance System [Internet]. Atlanta (GA): CDC; [cited 2022 Jan 5]. Available from: https://www.cdc.gov/brfss/index.html
45 Esri. ArcGIS Business Analyst [Internet]. Redlands (CA): Esri; [cited 2022 Jan 5]. Available from: https://www.esri.com/en-us/arcgis/products/arcgis-business-analyst/overview
46 Multi-Resolution Land Characteristics Consortium [home page on the Internet]. Washington (DC): MRLC; [cited 2022 Jan 5]. Available from: https://www.mrlc.gov/
47 MIT Election Data and Science Lab. U.S. president precinct-level returns 2016 [Internet]. Cambridge (MA): Massachusetts Institute of Technology; 2018 [cited 2022 Jan 5]. [Harvard Dataverse Version 11.0]. Available from: https://dataverse.harvard.edu/dataset.xhtml?persistentId=doi:10.7910/DVN/LYWX3D
48 Rosenbaum PR, Rubin DB. Assessing sensitivity to an unobserved binary covariate in an observational study with binary outcome. J R Stat Soc Series B Stat Methodol. 1983; 45(2):212–8.
49 Greenland S. Basic methods for sensitivity analysis of biases. Int J Epidemiol. 1996;25(6):1107–16.