Face mask use and physical distancing before and after mandatory masking: Evidence from public waiting lines

Seres, Gyula; Balleyer, Anna; Cerutti, Nicola; Friedrichsen, Jana; Süer, Müge

Seres, Gyula; Balleyer, Anna; Cerutti, Nicola; Friedrichsen, Jana; Süer, Müge

Working Paper

Face mask use and physical distancing before and after mandatory masking: Evidence from public waiting lines

WZB Discussion Paper, No. SP II 2020-305

Provided in Cooperation with:
WZB Berlin Social Science Center

WZB
Wissenschaftszentrum Berlin
für Sozialforschung

Gyula Seres
Anna Balleyer
Nicola Cerutti
Jana Friedrichsen
Müge Süer

Face mask use and physical distancing before and after mandatory masking: Evidence from public waiting lines

Discussion Paper

SP II 2020–305

July 2020

Research Area
Markets and Choice

Research Unit
Economics of Change

Wissenschaftszentrum Berlin für Sozialforschung gGmbH
Reichpietschufer 50
10785 Berlin
Germany
www.wzb.eu

Affiliation of the authors:

Gyula Seres, Humboldt-Universität zu Berlin
(corresponding author; gyula.seres@hu-berlin.de)

Anna Balleyer, University of Groningen

Nicola Cerutti, Berlin School of Economics and Law

Jana Friedrichsen, Humboldt-Universität zu Berlin, WZB, and DIW Berlin

Müge Süer, Humboldt-Universität zu Berlin

Abstract

Face mask use and physical distancing before and after mandatory masking: Evidence from public waiting lines

During the COVID-19 pandemic, the introduction of mandatory face mask usage was accompanied by a heated debate. It was argued that community use of masks creates a false sense of security that could decrease social distancing, thus making matters worse. We conducted a randomized field experiment in Berlin, Germany, to investigate whether masks lead to decreases in distancing and whether this mask effect interacts with the introduction of a mask mandate in Berlin. Joining lines in front of stores, we measured the distance kept from the experimenter in two treatment conditions – the experimenter wore a mask in one and no face covering in the other – both before and after the introduction of mandatory mask use in stores. We find no evidence that mandatory masking has a negative effect on distance keeping. To the contrary, in our study, masks significantly increase distancing and the effect does not differ between the two periods. Further, we find no evidence that the mask mandate affected distancing. However, our results suggest that the relaxation of shop opening restrictions had a negative effect on distancing.

Keywords: COVID-19; Face Masks; Social Distancing; Risk Compensation; Field Experiment; Health Policy

JEL classification: I12, D9, C93

1 Introduction

The novel coronavirus SARS-CoV-2 that quickly spread to almost all countries in the world has – as of mid-June 2020 – lead to more than eight million confirmed infections and more than 400,000 deaths (CSSE, 2020; Dong et al., 2020). To address the imminent health emergency, to make the growth rate of the virus sub-exponential (colloquially, to flatten the curve), and to mitigate hospital overload, most countries implemented complete or partial lockdown policies including stay-at-home orders, travel bans, social distancing, and also emphasized personal precautions in terms of hand hygiene and respiratory etiquette. While the conjunction of these policies has been proven effective and death rates are believed to have been substantially higher in their absence, it has also become clear that both the uncontrolled pandemic and the successful lockdown measures have had severe consequences for the economy and society (Fernando E. Alvarez, 2020; Thunström et al., 2020).

Given that SARS-CoV-2 outbreaks might remain a possibility for a long time, societies need to develop alternatives to a strict lockdown that allow for a safe life with the virus even though neither an effective treatment nor a vaccine is available. Mandated face mask use is a non-pharmaceutical intervention that is potentially very potent in combating COVID-19 (van der Sande et al., 2008; Rengasamy et al., 2010; Suess et al., 2012; Saunders-Hastings et al., 2017; Eikenberry et al., 2020; Mitze et al., 2020). However, health authorities and politicians have been cautious in advising universal mask mandates with reference to a potential backlash from an induced false sense of security (WHO, 2020; Synhetsstyrelsen, 2020; Norwegian Institute of Public Health, 2020). Such compensating behavior is found in the context of road safety regulation, offsetting the expected positive effects from regulation (Peltzman, 1975). Subsequently, although risk compensation is studied in the context of HIV prevention (Eaton and Kalichman, 2007; Marcus et al., 2013; Wilson et al., 2014), bicycle helmets (Adams and Hillman, 2001), and seat-belt laws (Houston and Richardson, 2007; Evans and Graham, 1991; Cohen and Einav, 2003), among others, there are mixed results on the existence of risk compensatory behavior.

As robust evidence on the existence and extent of risk compensation in response to face mask use is missing, citing risk compensation as an argument against the mandatory use of masks relies on two implicit assumptions. First, this argument assumes that risk compensation will actually happen in the context of the current epidemic and in response to mask use. Second, risk compensation only matters if its effect is larger than the presumably positive direct effects of a greater prevalence of masks in the community that would follow from a mask mandate. Seres et al. (2020) run a field experiment measuring the effect of face masks on distancing in outside waiting lines in Berlin, Germany in April 2020. Their study provides evidence against risk compensation in a context where masks were not mandatory. Subjects were observed to even stand further away from an experimenter who was masked than from an unmasked one. Using an additional survey, they show that this behavior might be triggered by second order beliefs, meaning that people expect individuals who wear a mask to prefer others to stay further away from them. Seres et al. (2020) yields insights about behavioral effects of masks, but their results cannot be easily extrapolated to a situation with a mask mandate. Therefore, we address the question of how a mask mandate influences the effect of masks on distancing behavior.

Using the same methodology as Seres et al. (2020), we run a field experiment outside waiting lines in Berlin before and after a mask mandate was put into place. We find that individuals stand further away from someone wearing a mask than from an unmasked person both before and after the introduction of the mask mandate. Thus, the mandate did not crowd out the positive effect of the face mask observed under voluntary masking. While we observe more people wearing masks themselves after the introduction of the mask mandate than before, we also find that average distances to other persons are shorter after the mandate than before. Using contextual data in the form of the number of open shops in the surrounding, we argue that this effect is not driven by the mask mandate but by concurrent changes in the perceived risk from the virus. This is in line with an array of studies showing that the adoption of precautionary behavior against COVID-19 crucially depends on the perceived risk of becoming severely ill from the virus (Ajzenman et al., 2020; Allcott et al., 2020; Grossman et al., 2020; Harper et al., 2020; Larsen et al., 2020; Rosenfeld et al., 2020; Wise et al., 2020).

Our results complement further evidence from Germany, Italy, and the US. Empirical studies examine social distancing in terms of time spent outside and proximity during this period. Kovacs et al. (2020) use location data from Germany to show that the introduction of face mask mandates in Germany did not lead to a compensatory effect in individuals' mobility patterns in terms of time spent outside. In contrast to these, Yan et al. (2020) argue that US Americans spent more time outside their homes after masks became mandatory in public spaces. Their empirical strategy does not preclude that mobility would have changed in this way also in the absence of masks. Another key dimension is observed distancing in community settings. A very similar result to that of Seres et al. (2020) was obtained in a field experiment in Italy: Marchiori (2020) shows that wearing a face mask can substantially improve adherence with the physical distancing regulations on pavements both in the absence of a mask mandate and after its introduction. Based on this body of evidence, a direct behavioral backlash on distancing from making face masks mandatory appears unlikely. For maximal support from the population, such mandates should be clearly communicated as necessary and as an additional safety measure.[1]

The rest of this paper proceeds as follows: Section 2 provides the setup, including the local progress of the epidemic, and provides a general introduction to the policy environment. Section 3 describes the experimental design. Section 4 formally states the hypotheses. Section 5 provides the main results of this paper. Section 7 concludes with further interpretations of the main results and a discussion.

2 Background

A face mask mandate was introduced in all German federal states toward the end of April and coincided with the relaxation of other regulations. In Berlin, starting in mid-March 2020, only supermarkets and stores selling basic necessities were allowed to open. Then, from April 22, 2020 onward, small retail stores ($< 800m^2$) were allowed to reopen under certain restrictions (e.g., limited number of customers). While masks had been previously dismissed as an attractive policy option, the expectation of the increased movement of citizens and potential crowding in cities as well as increasing public pressure, led to the introduction of mandatory masking policies in all federal states with only slight variations regarding the exact starting dates. The objective of these policies was to reduce the risk of contagion in places that became increasingly frequented but where physical distance recommendations are harder to uphold, such as shops and public transport. However, individuals may adjust their precautions in other dimensions in response to such a mandate so that the net effect is, *a priori*, not clear. Therefore, we follow up on Seres et al. (2020) with an identical field experiment, conducted in Berlin *after* the introduction of compulsory masking, evaluating the effect of masks and the interaction with other policy changes. The detailed timeline of the experiment and the restrictions are stated in Table 1.

[1] Settele and Shupe (2020) provide evidence from survey data that support for policies critically depends on the information and perceptions that individuals hold.

14.03.2020	**Beginning of Corona Related Restrictions**
22.03.2020	**Tightest Restrictions in Place**
Start: Data Collection 1	18.04.2020
20.04.2020	**Retail** $< 800m^2$ **Reopen**
End: Data Collection 1	24.04.2020
27.04.2020	**Mask Mandate in Public Transportation**
29.04.2020	**Mask Mandate in Shops**
02-04.05.2020	**Big Retails Reopen, Gatherings of up to 50 People Allowed**
Start: Data Collection 2	12.05.2020
15.05.2020	**Restaurants and Cafes Reopen with Limited Hours**
End: Data Collection 2	20.05.2020

Table 1: Berlin COVID-19 Restrictions and Experiment Timeline

3 Experiment

3.1 Location

The field experiment took place in 2020 during the COVID-19 pandemic in Berlin, Germany, a place that was not an outlier in terms of infection rates in Germany.[2] The first part of the data was collected before the face mask mandate and the second half after the introduction of the mandate. During the first data collection period, acceptable reasons to leave the place of residence were defined at the state level, limiting the mobility of the experimenters. To comply with public health recommendations, the choice of stores was made to avoid long commuting from the experimenters' homes. Figure 1 shows the locations of businesses visited. There was no overlap in the list of stores between experimenters, therefore, only one of them visited each store in the sample. To better profit from the natural experiment setting created by the mask mandate, in May, the experimenters revisited the same stores as those in April. The store types, where observations took place, were previously restricted to supermarkets, drug stores (except pharmacies), and post offices to observe a sample representing the population visiting public areas. During the pandemic, lines in front of businesses were frequent in Berlin, but irregular. Therefore, only the existing lines at the moment of data collection could be utilized for our experiment. We address potential randomization concerns regarding store selection in Section 5.

[2]According to Robert Koch Institute, one of the central bodies for the safeguarding of public health in Germany (`https://www.rki.de/`), the state of Berlin had the seventh highest number of SARS-CoV-2 infections per 100,000 population of the 16 German states as of May 1, 2020, when the incidence of COVID-19 cases in Berlin was 157 per 100,000 inhabitants; close to the federal average 197 per 100,000 inhabitants.

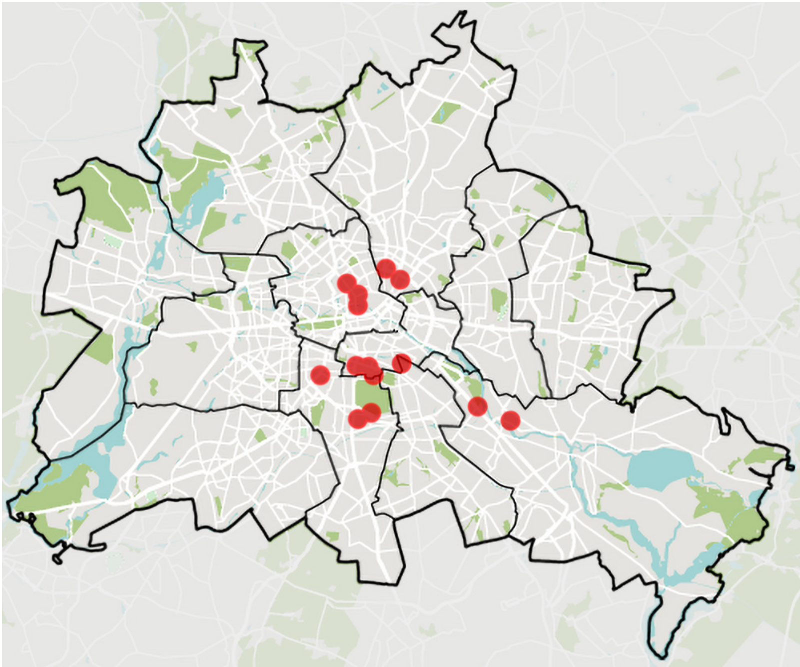


Figure 1: Map indicating the observation sites

3.2 Experimental Design

Our experiment has a 2×2 between-subject design with respect to using a mask and time period. To study the effect of masks on distancing, we use a between-subject design with randomized face covering. In the MASK treatment, the experimenter was wearing a mask, whereas in the NOMASK treatment, no face covering was used. To investigate whether the treatment effect interacts with a set of policy changes implemented at the end of April 2020, we ran the same design twice, once before and once after the introduction of a mask mandate in Berlin.

Our experiment was carried out by experimenters who measured the distance between themselves and others in lines in front of businesses. Data was collected in two periods between April 18-24 and May 12-20. In both periods, 60 observations each were recorded by four experimenters, adding up to 240 in each period and 480 in total. The pre-registered experimental protocol is in Appendix A.[3]

The experimenters are independent researchers, two women and two men, aged between 31 and 35, who participated voluntarily and are credited as co-authors of this paper. Each recorded the observations individually in their own neighborhood. Two measures were taken to reduce as much as possible potential noise from different appearances of the experimenters. First, each member of the team used a white FFP2 respiratory protection mask, which was the most easily accessible type of mask in pharmacies during the first period of data collection.[4] Second, the dress-code was standardized to a pair of blue jeans and a dark colored top (Balafoutas and Nikiforakis, 2012).

Each experimenter independently located a line outside a shop in their neighborhood and determined an even number of observations to be collected there. The experimenter wore a mask (treatment MASK) or not (treatment NOMASK) based on the result of a coin toss.

[3]The pre-mandate data was used in Seres et al. (2020), which had five experimenters. In this paper, we use data of four experimenters who participated in both measurement periods. This study was pre-registered with five experimenters; however, one was unable to participate in the second period.

[4]An FFP2 mask is a mechanical filter respirator as defined by the EN 149 standard, similar to the N95 design. FFP2 and surgical masks are not visibly different and we do not expect that carrying out the experiment with surgical masks would have altered the results.

Then, the experimenter joined the line, maintaining a distance of 150 cm from the previous person, measured with a mobile device. While waiting for the subject, meaning the next person arriving and joining the line behind the experimenter, she/he assumed a sideways position in the line, thus ensuring her/his face would be visible to the next person but avoiding eye contact. Upon arrival of the subject, the experimenter measured the distance between her/his own feet and the subject's, subsequently left the line, and input the measured distance and demographic data of the subject into a previously prepared table. Particular cases (e.g. groups of people, strollers) were uniformly measured according to the protocol (Appendix A). The distance was recorded via a mobile augmented reality application, which provides 1-centimeter precise measurements. No visual or audio recordings were taken to comply with privacy laws. A measurement took about 5-20 seconds to complete. Distance was only recorded if the subject assumed a steady position for the time of measurement and it was clear for them where to stand. For groups, the measured subject was the person closest to the experimenter. We made note of no case when a subject recognized the measurement or reacted to it by moving away. Having completed the input of data, the experimenter returned to the end of the line.

At any store and period, an equal number of observations with and without a mask were collected. At each visit, the experimenter used a coin toss to determine with which of the two treatments to start.

The experimenters also collected information on the subjects' demographic profile. In particular, subject's age group, gender, the number of accompanying children and adults were recorded as well as whether the subject was wearing a mask at the time of measurement. Note that during the second round of data collection, i.e. after the introduction of the mask mandate, all subjects presumably had a mask with them as a prerequisite to enter the store, unlike before the mandate. However, no law mandated using the mask while waiting outside.

Additional controls for the setting include the length of the queue, store type and exact location. In order to control for the impact of the store closure policy, we recorded the number of businesses within a 50-meter radius around the location that were open at the time of measurement during the second round of data collection in May that were legally closed during the first round of data collection in April. This variable shows substantial variance as it ranges from 0 to 6 in the May sample. As we argue above, changes in distancing may also be influenced by the general perception about the epidemic. As a measure of this factor, we gathered daily data from Google trends that shows the relatively number of searches for the novel coronavirus in Berlin.[5]

4 Hypotheses

The study consists of two main observational periods, the first taking place before the exogenous policy changes, including the introduction of mandatory mask wearing in stores, and the second one afterwards. Each period has a balanced number of observations per treatment group. The introduction of a mandate by the state creates a natural experiment setting and lets us understand the impact of the policy by analyzing pre- and post-intervention periods. As the mandate was brought into force at the same time as the aforementioned measures, it is hard to isolate its pure effect. However, we believe that the mask mandate and relaxation measures have different effects on our dependent variable, kept distance. From this point on, the exogenous difference between periods is referred to as policy change and the different policies are underlined separately when necessary.

In the pre-policy period, wearing a mask was voluntary; whereas during the post-policy, all subjects had to carry a mask with them as they were expected to wear it in the store. Therefore, we hypothesize that as masks became a common sight of the city in the second observational period, a mask mandate would increase the general public awareness of the

[5]The chosen keyword is "Coronavirus", as it is most commonly called colloquially in the German-speaking online community.

health hazard. Based on the two-process theory of reasoning (Stanovich and West, 2000; Kahneman, 2011), we argue that compliance with distancing requires mental effort. In our context, both the mandate and seeing a masked experimenter may serve as a reminder inducing a conscious decision-making process, System 2.[6] Thus, the mandate would not have a significant impact on the subjects entering the line behind the masked experimenter, whereas it would increase the distance kept by the subjects behind the experimenter without a mask by already triggering their System 2 through the higher presence of masked people on streets. Keeping the natural experiment setting and our expectations in mind, we formed and preregistered 3 hypotheses.[7]

Seres et al. (2020) use a survey to understand the mechanism behind the mask increasing physical distancing and conclude that it might be the following: people tend to believe that a person wearing a mask prefers others to keep a greater distance. This mechanism could still be at play after the introduction of the mask mandate because our experiment takes place outdoors and masks are only obligatory in stores. Another explanation for a distance increase in response to face masks could be a reminder effect: people seeing other individuals wearing masks may be reminded of the health risk from COVID-19 and, thereby, of the appropriate measures to take to prevent an infection. Seres et al. (2020) find no evidence in this regard. However, the introduction of the mask mandate required people joining lines to carry a mask with them, led to an increase in general mask usage in the waiting line, and could potentially create a reminder effect through the larger presence of masks. Put differently, masks might work as a trigger activating System 2 cognition, thus leading people to keep a greater distance just because of the fact that they adjust their judgment based on the severeness of the situation. Taking both mechanisms into consideration, the introduction of a mask mandate may not change the distancing behavior of the subjects behind the masked experimenter. If a mask still works as a respect signal, then they continue to signal the same. On the other hand, if masks work as a general reminder, then the effect of this reminder would spread to the entire population without affecting the behavior of the subjects joining the line behind a person with a mask.

Hypothesis 1.A. *Distance kept toward the masked experimenter in treatment* MASK *in the waiting line does not change with the policy.*

Hypothesis 1.B. *Distance kept toward the unmasked experimenter in treatment* NOMASK *in the waiting line is greater after the policy change.*

Unlike people behind the masked experimenter, based on the two-process theory, the mandate can be expected to have a positive effect on the distancing behavior of people behind a person without a mask. According to this explanation, the mandate itself as well as its direct consequence of seeing masks more often on the streets as well as in the waiting line could potentially increase the general public awareness regarding COVID-19-related risk and mitigation measures. The heightened awareness may then induce subjects to increase their precautions and, thus, their distancing.

Hypothesis 2. *After the policy change, distance kept toward the experimenter in the waiting line is the same in treatments* MASK *and* NOMASK*, i.e., subjects keep the same distance from the masked and unmasked experimenter.*

As a consequence of hypotheses 1.A and 1.B, we expected the subjects in the post policy sample to keep on average the same distance from the experimenter in both treatment conditions. The convergence of distances in the two treatments is driven by the expectation that the mandate might increase distancing for people behind a person without a mask but leave unchanged the distances behind a person wearing a mask. A convergence in distances across

[6] System 1 is a cognitive process defined in the psychology literature as automatic, largely unconscious, relatively fast, and undemanding of computational capacity. In contrast, System 2 is demanding and relatively slow.

[7] For narrative purposes, the order and wording of hypotheses is different from that of the preregistered list.

the two treatments could alternatively result if masks lose their informational value with the introduction of the mandate but awareness is unchanged. If wearing a mask is no longer perceived as signaling a preferred larger distance, people behind the masked experimenter might keep a shorter distance after the mandate than before and on average the same as if standing behind the unmasked experimenter. This explanation is not hypothesized separately as it directly contrasts with Hypotheses 1.A and 1.B.

Hypothesis 3. *After the policy change, subjects wearing a mask do not keep a greater distance from the experimenter than unmasked subjects (treatment conditions* Mask *and* NoMask *pooled).*

Using the pre-mandate sample, Seres et al. (2020) conclude that subjects wearing a mask keep a larger distance in general. However, they also argue that it might be due to a selection effect. Therefore, the mask mandate is expected to increase the representation of those who wear a mask in the line, creating a different selection compared to the pre-policy period. Hence, distancing behavior of this post-policy sub-sample might resemble that of the rest of the post-policy sample, therefore, we expect no difference in distancing behavior.

5 Empirical analysis

5.1 Sample Characteristics

The data set contains 480 observations. For descriptive statistics on subject characteristics, see Table 2. Compared to the city's age groups, our sample underrepresents 60+ population (10.6% vs. 24.7%). This is not surprising due to the asymmetric effect of the disease on elderly (Verity et al., 2020). However, considering that social distancing is crucial in public, our study aims to measure the effect on people who leave their homes. Our sample is representative in terms of gender according to a Chi-Square Goodness of Fit test (54.4% vs. 50.8% in the population, $\chi^2 = 2.45$, p = 0.117). Most subjects in our sample arrive at the store alone, only 10.6% come with adult, and 6.1% with minor, companions. There is a clear increase in mask use after the policy change as it soars from 17.1% pre-mandate to 40.1% post-mandate. We also recorded the length of the line as the number of people standing outside in front of the experimenter. The mean length is 5.63 individuals with a standard deviation of 3.83.

5.2 Estimation Strategy

Our analysis seeks to understand whether and how the introduction of a mask mandate changes the physical distancing behavior that face masks create. In doing so, we exploit a natural experiment setting formed after the mask use in stores in Berlin was mandatory starting from April 27^{th}, 2020. Randomization of mask use by the experimenters allows us to use a difference-in-differences approach, comparing pre- and post-policy periods for subjects behind a masked or an unmasked experimenter. Due to the randomization of treatments, we can assume parallel trends between treatment groups, thus arguing in favor of causal evidence regarding permanence of masks' behavioral effect. On the other hand, we are also well aware of the fact that the introduction of a mask mandate came along with other relaxation measures. Thus, we interpret the pre- and post-policy differences as the joint effect of the mandate and relaxations, naming it accordingly in our model.

Pooling the entire sample, we estimate the following equation to identify the effect of the policy on distancing and its interaction with our Mask treatment:

$$\begin{aligned} Distance_i = \beta_0 + \beta_1 MaskE_i + \beta_2 Policy_i + \beta_3 MaskE_i \times Policy_i \\ + \beta_4 MaskS_i + \phi X_i + \varepsilon_i \end{aligned} \tag{1}$$

Count	Pre-Mandate NoMask	Pre-Mandate Mask	Post-Mandate NoMask	Post-Mandate Mask	∑
Subject Without Mask	102	97	77	65	341
Subject With Mask	18	23	43	55	139
Accompanying Adult =0	107	105	108	109	429
Accompanying Adult =1	11	13	12	10	46
Accompanying Adult >1	2	2	0	1	5
Accompanying Child =0	111	112	113	116	451
Accompanying Child =1	7	7	7	4	25
Accompanying Child >1	2	2	0	0	4
Female Subject	61	65	65	70	261
Male Subject	59	55	55	50	219
Aged under 15	0	1	1	0	2
Aged between 15 and 25	13	19	14	15	61
Aged between 25 and 35	38	34	42	40	154
Aged between 35 and 45	35	29	33	33	130
Aged between 45 and 60	20	20	21	21	82
Aged above 60	14	17	9	11	51
Total	120	120	120	120	480

Table 2: Number of Subjects in different treatment conditions.

Notes: Values show the number of observations with the given characteristics for categorical variables. Age groups and gender reflect the experimenters' impressions and are not to be interpreted as point estimates. Subjects are counted with a mask if they were wearing one at the time of measurement.

in which $Distance_i$ is the distance kept by subject i, $MaskE_i$ is the indicator of the experimenter wearing a mask, $Policy_i$ is an indicator for data collected after the policy changes took place, thus distinguishing the two periods of data collection, and $MaskS_i$ is an indicator for the subject wearing a mask. Hence, β_2 captures any effect in distancing that results from the conjunction of policy changes between the first and the second data collection period but does not relate to the treatment, whereas the effect of the mask mandate jointly with other policy changes on the effect from mask wearing is identified by β_3. If we cannot reject $\beta_3 = 0$, this implies that the face mask effect on distancing is not significantly different between two periods. X_i is a vector of all other covariates and controls used in different model specifications. Standard errors ε_i are clustered according to store and date in order to mitigate any potential correlation in error terms.[8] As the experimenter locations are not overlapping in our sample, this approach also covers experimenter related correlations.

5.3 Main results

We structure the discussion of results along the hypotheses laid out above. In the raw data, we observe that subjects keep a shorter distance from the experimenter in the data collected after the policy change than before it on average. This also holds true in each treatment condition separately. While subjects kept an average distance of 151.14 cm (SD=29.62) to the unmasked experimenter in our pre-mandate sample, the average distance to the unmasked experimenter is only 143.35 cm (SD=31.79) in the post-mandate sample. Similarly, subjects kept an average distance of 159.85 cm (SD=31.79) to the masked experimenter in the pre-mandate sample, but only 151.41 cm (SD=34.08) in the post-mandate sample. Thus, distances kept are, on average, 7.79 cm (NoMask) and 8.41 cm (Mask) shorter in the post-policy period than in the data collected before the policy change. Regression analysis confirms this observation: from Table 3, we see that the coefficient of the policy change

[8]The clustered standard errors are used to mitigate any potential serial correlation in the error terms due to clustered sampling. As we are considering relatively small number of clusters (55 in total), we also perform wild cluster bootstrap method as a robustness check following Cameron et al. (2008). Please see Appendix for details.

is negative if we regress observed distances on the policy change and controls for the subsample of subjects facing the masked (column 1) and unmasked (column 2) experimenter. Therefore, we reject both hypotheses 1.A and 1.B that subjects do not change or increase their distancing toward the masked respectively unmasked experimenter due to the policy change.

Result 1. *Distance kept toward the experimenter in the waiting line is weakly shorter after the policy change in treatments* Mask *and* NoMask.

In order to better understand whether the observed change in distancing between the pre- and post-policy samples are driven by the mask mandate or by the relaxations in the restrictions in place, we estimate different specifications of equation 1. The results and summarized in columns 3 to 6 of Table 3.

Table 3: Distance OLS

	(1)	(2)	(3)	(4)	(5)	(6)
Sample	Mask	NoMask	Pooled	Pooled	Pooled	Pooled
MaskE			9.222*	9.449*	9.173*	9.392*
			(4.216)	(4.266)	(4.221)	(4.273)
Policy	-12.26**	-8.213	-9.294*	-0.0833	-3.055	8.082
	(4.380)	(4.546)	(4.610)	(6.407)	(9.891)	(10.88)
MaskE×Policy			-1.614	-2.250	-1.481	-2.098
			(5.557)	(5.458)	(5.610)	(5.502)
MaskS	13.57*	1.545	7.376*	7.623*	7.499*	7.785*
	(5.838)	(4.052)	(3.251)	(3.298)	(3.302)	(3.356)
Stores				-3.173**		-3.255**
				(1.177)		(1.198)
Online Search					0.170	0.216
					(0.236)	(0.209)
Pop. Density	-2.282***	-0.610	-1.423***	-1.066**	-1.435***	-1.071**
	(0.411)	(0.343)	(0.309)	(0.330)	(0.300)	(0.315)
Acc. Adult	-5.087	-4.387	-5.442	-5.788	-5.403	-5.748
	(6.952)	(8.273)	(4.999)	(4.944)	(4.992)	(4.933)
Acc. Child	-0.382	-5.921	-4.444	-3.789	-4.482	-3.821
	(6.339)	(4.175)	(3.126)	(3.254)	(3.163)	(3.296)
Ppl in Line	0.625	0.904	0.865**	0.938**	0.860**	0.934**
	(0.731)	(0.484)	(0.304)	(0.309)	(0.296)	(0.303)
Constant	180.7***	158.7***	165.2***	156.5***	151.5***	138.9***
	(7.179)	(5.412)	(5.932)	(6.701)	(20.15)	(18.32)
Demographics	Yes	Yes	Yes	Yes	Yes	Yes
Observations	240	240	480	480	480	480
R^2	0.146	0.091	0.107	0.125	0.109	0.127

Notes: Ordinary least squares estimates. Dependent variable is distance kept from the experimenter. Standard errors in parentheses are clustered by day and store. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. MaskE and MaskS are indicator variables for whether the experimenter or subject, respectively, used a face mask. Acc. Adult and Acc. Child indicate whether the subject was accompanied by at least one other adult or child, respectively. Density is population density based on the 2011 German Census data. Controls include gender and age dummy variables. Standard errors are clustered in day and store level.

In line with the observed raw difference in distancing, column (3) indicates that distances toward both the unmasked and the masked experimenter are about 9 cm shorter after the policy changes had taken effect. While the coefficient on the treatment dummy *MaskE* is positive and significant, the coefficient of the interaction between *MaskE* and the policy change is not significant, suggesting that subjects keep significantly larger distances to the masked experimenter both before and after the policy changes took effect.

We conclude that any difference from the policy changes must have affected subjects facing the masked and the unmasked experimenter equally. Using additional specifications, we argue that the observed shift in behaviors can be explained by a combination of factors including a change of general perception about the pandemic and relaxation of business openings. Specifically, we include the variable *Stores*, measuring the number of businesses that were legally closed in April but were open in May at the time of measurement within a 50-meter radius of the point of data collection in May, and the variable *Online Search*, representing the relative number of Berlin specific hits on Google search for the novel coronavirus on the day of measurement.

Models (4)-(6) in Table 3 include the additional covariates individually and in combination. These results suggest that the decline in distancing after the policy change can be explained by reopening stores rather than the introduction of the mask mandate.[9]. According to model (6), an additional newly open store near the location of measurement is related to a decreased in distancing of 3.255 cm on average (p=0.002). However, we find no evidence that online search for pandemic-related content predicts the difference between distances in the pre- and post-mandate samples well. Virus-related internet traffic positively correlates with greater distancing but the effect is not significant (p=0.257).

Next, we investigate the effect from the mask intervention in the post-mandate sample (hypothesis 2). From table 3, it can be seen that the marginal effect of *MaskE* is at least 9 cm in all specifications and significant. Further, the effect is robust across time: the interaction with the policy change is negative, as predicted, but small and not statistically different from zero. Thus, we find no evidence suggesting that the effect has vanished with the introduction of a mask mandate and thereby reject hypothesis 2. Further, in the post-mandate sample, subjects kept greater distances from the masked than from the unmasked experimenters, on average.

Result 2. *In the post-mandate sample, subjects maintain a significantly larger distance from the experimenter wearing a mask.*

We now turn to hypothesis 3 that, in the post-mandate period, the sub-sample of subjects wearing a mask themselves does not react differently to the masked experimenter than the rest of the sample. We first note that the share of masked subjects is indeed much higher post-mandate than pre-mandate, even though masks were never mandatory in the outside waiting lines we studied (see table 2).[10] Indeed, this increase suggests that more people regularly wear masks, such that the subsample of mask wearers post-mandate is less selected and more similar to the population of unmasked individuals. We find that subjects with mask keep a significantly larger distance than unmasked subjects in the pre-mandate sample (two-sample t=-2.3788, p=0.0091), but this difference vanishes in the post-mandate sample (two-sample t=-0.6327, p=0.2638). Thus, our data is fully consistent with hypothesis 3.

Result 3. *In the post-policy sample, subjects wearing a mask do not keep a larger distance from the experimenter.*

Even though the self-selection into mask-wearing does not allow for a causal interpretation of the estimate coefficient on *MaskS*, we further note that the coefficient is significantly positive in specifications (3) to (6), contradicting the hypothesis that subjects may engage in risk compensation and reduce their distancing in response to the protection offered by a mask.

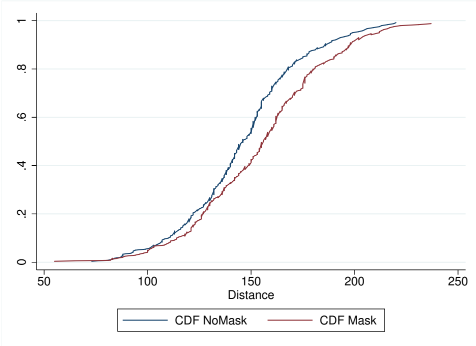
Table 3 reveals further interesting patterns. Population density of the neighborhood decreases distancing, 1000 inhabitants/km^2 decreases distancing by more than 1cm.[11] Subjects arriving in a group keep a shorter distance, but the difference is not significant in any specification, neither for adult nor for minor companions.[12] The number of people in line

[9]We do not include experimenter fixed-effects here because they are collinear with the variable local and the model would be overidentified including both.

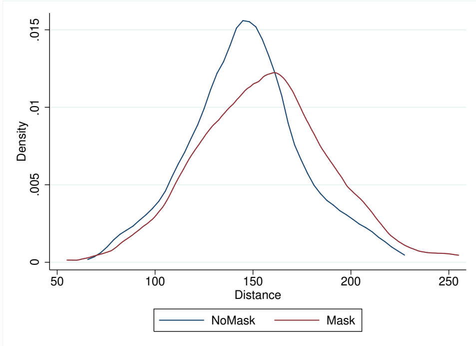
[10]The use of masks is significantly higher post mandate, γ_2=0.9036 p=0.014.

[11]Area is defined by postal code.

[12]Further demographic controls gender and age dummies are not significant.



(a) Cumulative distribution function.



(b) Kernel density estimates

Figure 2: Cumulative distribution functions of distances kept by the subject from the experimenter in NoMask (blue) and Mask (red) conditions (in centimeter). Cumulative distributions are exact and densities are estimated univariate Epanechnikov kernel density functions.

in front of the experimenter has a small but significant effect on distancing (in specification (6): -0.93, p=0.003). To test if wearing a mask makes subjects not to stand behind the experimenter, we test if the sample correlation coefficient between this and the treatment variable is significant. The reasoning is that this behavior would increase the time between observations, resulting in shorter lines. This claim is rejected ($r = 0$, p = 0.117).

5.4 Further Results

German health authorities and official mandates to limit the spread of the coronavirus specify that individuals should keep a distance of at least 150 cm to each other. In addition to our main analysis, we investigate how our treatment of masking the experimenter and the introduction of the mask mandate in Berlin affect compliance with this required minimum distance. We find that compliance is higher toward the masked experimenter in both observation periods. Before the mask mandate, compliance is 54.17% if the experimenter does not wear a mask and 69.17% if she/he does. After the mandate is introduced, compliance is 40% if the experimenter does not wear a mask and 49.17% if she/he does.

The 150cm rule may look arbitrary as the recommendations of safe distances vary substantially between countries.[13] Hence, we also consider if compliance with alternative threshold values increases with masking the experimenter. Figure 2a demonstrates that the choice of the critical value does not change our conclusion that mask improve distancing from the experimenter. It is evident from the figure that subjects in the Mask condition are more likely to exceed any relevant threshold value, i.e. compliance is generally higher there than in the NoMask condition. Using non-parametrically estimated kernel density functions, we confirm a positive shift in distancing (Fig. 2b, D=0.175, P =0.01, two-sided Kolmogorov–Smirnov test).[14]

Equation (1) is correctly specified only if the treatment dummy *MaskE* does not influence the subject's decision of putting on a mask, *MaskS*. We believe that the exogeneity of *MaskS* is given as subjects decide about their use of a mask before seeing the experimenter. However, this decision may be reversed upon seeing the experimenter. We therefore test the

[13]For example, as of June 2020, the U.S. Center for Disease Control and Prevention (CDC) recommends a 6-feet distance (=182.88cm).

[14]A parametric test yields similar results. We estimated a logit model analogous to model (6) in Table 3 where now the dependent variable is compliance with the 150cm threshold. The estimated coefficient of *Mask Experimenter* is positive and significant (β_1 =0.7145, p=0.02). The interaction between *Mask Experimenter* and the policy change is insignificant as in the main model (β_3=-0.2893, p=0.535). See appendix for details.

independence claim is tested with the following logit binary choice model:

$$Pr(MaskS = 1) = \frac{\exp(\gamma_0 + MaskE + \gamma_2 Mandate + \gamma_3 MaskE \times Mandate \times Mandate + \phi X_i + \varepsilon_i)}{1 + \exp(\gamma_0 + MaskE + \gamma_2 Mandate + \gamma_3 MaskE \times Mandate + \phi X_i + \varepsilon_i)} \quad (2)$$

Using the same set of covariates as in Model (4) in Table 3, we find that the coefficient of *MaskE* is not significant (γ_1=0.2358, p=0.524). We conclude that the subjects decide about wearing a mask independently of whether the experimenter wears a mask or not.

5.5 External Validity

According to the medical literature, airborne contagion is a primary source of transmission of SARS-CoV-2 and, thus, distancing between individuals is important to prevent the spread of the virus (Anderson et al., 2020; Liu et al., 2020). Hence, a successful mitigation strategy needs to understand and take into account how policy affects distancing patterns. In this study, we analyze how the introduction of a mask mandate affected distancing in order to contribute knowledge in this respect. An overall evaluation of policies on distancing needs to take into account as many facets of individual behavior as possible because restrictions as well as re-openings alter the choice set of customers, resulting in changes in behavior change that affects the potential exposure to infectious particles.

Putting our design in perspective, it is crucial to note that the data was collected in an environment where transmission is possible (Qian et al., 2020)[15] but at the same time, wearing a mask is optional even under the mandate. The places where we collected data fall into a category of settings in which distancing is recommended by authorities with the pretext that it helps preventing contagion. Our findings do not necessarily generalize to the effect of the mask mandate in stores or high-risk areas where the mask mandate made wearing a mask mandatory. However, our results are fully in line with evidence from mobility patterns in Germany, which have not changed to the negative with the introduction of a mask mandate according to Kovacs et al. (2020).

6 Discussion

We seek to extend the literature on masks by investigating how face mask policies affect distancing and, specifically, how they interact with the effect that face masks have on distancing behavior. We follow up on the claim that face masks make individuals prone to less rigorous compliance with other contagion prevention recommendations, such as physical distancing, a claim that is frequently heard in the discussion on face masks but has so far not received empirical support. Seres et al. (2020) show that individuals keep a greater distance from someone who is masked, contradicting a negative effect of masking. However, *ex ante*, it is not clear that this would still be true in the presence of a mask mandate. With data from periods before and after the introduction of a mask mandate in Berlin, we show that the positive effect of masking observed in Seres et al. (2020) persisted in the presence of a mask mandate. In this section, we evaluate our findings in the light of motivation crowding, risk compensation, two-process theory, and cognitive dissonance.

Based on motivation crowding theory, one may argue that the introduction of a face mask policy may ultimately alter the behavioral response of individuals beyond the wearing of masks, e.g. distancing behavior. We argue that a large part of the precautions that individuals engage in to flatten the curve are intrinsically motivated, this is in line with evidence on the effect of perceived risk on precautions taken and also with the observation that mobility in many place was substantially reduced even before official stay-at-home orders became effective. According to motivation crowding theory, a mask mandate may then crowd in

[15] Airborne lifetime of small speech droplets can reach 8-14 minutes according to Stadnytskyi et al. (2020), but air movements can dilute the concentration of virus and make transmission substantially less likely.

or crowd out intrinsic motivation to comply with measures to prevent the virus spread, depending on circumstances (Frey and Jegen, 2001). A crowding-in in motivation and, thus, an effect that reinforces the effectiveness of the mask mandate with respect to virus spread can be expected if individuals perceive the policy as supporting their intrinsic motivation. However, if individuals perceive it as negating their intrinsic efforts or as making those redundant, the policy may induce countervailing behavior change by crowding out intrinsic motivation (Frey and Jegen, 2001; Festré and Garrouste, 2015). Further, policies may crowd out compliance with existing norms because they are perceived as rules that substitute social norms (Ostrom, 2000). Our data does not support the notion of crowding out from the introduction of the mask mandate as we attribute the observed decrease in distancing in the post-mandate sample to the accompanying policy relaxations.

Further, individuals may perceive the introduction of a face-masking mandate as an indication that alternative precautions – e.g., avoiding unnecessary contacts and trips, keeping safe physical distances to others – had become less relevant. Specifically, individuals may perceive face masks as an effective means of reducing the overall infection risk as evidence for this becomes available. If individuals show risk compensation behavior and decrease their compliance with complementary measures such as distance-keeping, the expected beneficial effect from compulsory masking would be (partially) negated. Ironically, introducing a face mask policy (rather than just recommending face mask use) may introduce a general increase in risk compensation behavior, which was the reason why face masks were not recommended as a protective measure in the first place by WHO and other health bodies. Our results speak against such a direct backlash from a mask mandate but also suggest that distancing is sensitive to contextual changes such as increased shop openings.

On a different note, we stated our hypotheses on policy effects based on the premise that both the masked experimenter and the mask mandate, which implies that subjects in the waiting line carry a mask with them, serve as triggers: According to the two-process theory of reasoning (Stanovich and West, 2000; Kahneman, 2011), these triggers would induce the necessary mental effort to comply with the recommended distancing. What we find is only partially consistent with this. The positive effect of the mask-wearing experimenter on distancing is present before and after the mask mandate, suggesting that the mask in treatment Mask has a trigger effect. But we do not find evidence that the mask mandate serves as a trigger as distancing if anything decreases after the mandate.

Another theory that may help better understand our findings is cognitive dissonance (Festinger, 1957). According to this theory, the interaction of people with the outside world highly depends on mental inconsistencies. In our case, the relaxations that came with the policy change at the end of April 2020 have plausibly suggested to people that the severity of the situation had decreased. Thus, avoiding a mask became less costly for human psychology after the policy change and the mask mandate may not have created enough of a general awareness to counteract this change in beliefs. When faced with the masked experimenter directly in front of them, though, subjects in our experiment apparently adjust their perception and better comply with the recommended distancing, leading to the observed positive mask effect.

7 Conclusion

This study utilizes a field experiment in which we measure distancing in lines to stores. The experimenters entered these lines randomly with and without a face masks during the COVID-19 pandemic. Measurement was carried out twice: before and after the mask mandate in stores. The main findings show that the effect of masks worn by the experimenters is positive significant and not significantly different in the two time periods. We do not find evidence of risk compensation: Subjects wearing a mask do not keep a shorter distance and the mask mandate did not change this. Using pre-mandate field data and survey, Seres et al. (2020) conclude that mask-driven distancing is implied by a signaling channel. A mask mandate does not impact this conclusion. Assuming that distancing is an effective measure against transmission of SARS-CoV2, we find no evidence that mandating masks has a negative spillover effect.

References

Adams, J. and M. Hillman (2001). The risk compensation theory and bicycle helmets. *Injury Prevention 7*(2), 89–91.

Ajzenman, N., T. Cavalcanti, and D. Da Mata (2020). More than words: Leaders' speech and risky behavior during a pandemic. *Available at SSRN 3582908*.

Allcott, H., L. Boxell, J. Conway, M. Gentzkow, M. Thaler, and D. Y. Yang (2020). Polarization and public health: Partisan differences in social distancing during the coronavirus pandemic. NBER Working Paper 26946.

Anderson, E. L., P. Turnham, J. R. Griffin, and C. C. Clarke (2020). Consideration of the aerosol transmission for COVID-19 and public health. *Risk Analysis 40*(5), 902–907.

Balafoutas, L. and N. Nikiforakis (2012). Norm enforcement in the city: A natural field experiment. *European Economic Review 56*(8), 1773–1785.

Cameron, A. C., J. B. Gelbach, and D. L. Miller (2008). Bootstrap-based improvements for inference with clustered errors. *Review of Economics and Statistics 90*(3), 414–427.

Cohen, A. and L. Einav (2003). The effects of mandatory seat belt laws on driving behavior and traffic fatalities. *Review of Economics and Statistics 85*(4), 828–843.

CSSE (2020). COVID-19 Dashboard by the Center for Systems Science and Engineering (CSSE) at Johns Hopkins University (JHU).

Dong, E., H. Du, and L. Gardner (2020). An interactive web-based dashboard to track COVID-19 in real time. *The Lancet Infectious Diseases 20*(5), 533–534.

Eaton, L. A. and S. C. Kalichman (2007). Risk compensation in HIV prevention: implications for vaccines, microbicides, and other biomedical HIV prevention technologies. *Current HIV/Aids Reports 4*(4), 165–172.

Eikenberry, S. E., M. Mancuso, E. Iboi, T. Phan, K. Eikenberry, Y. Kuang, E. Kostelich, and A. B. Gumel (2020). To mask or not to mask: Modeling the potential for face mask use by the general public to curtail the covid-19 pandemic. *Infectious Disease Modelling*.

Evans, W. N. and J. D. Graham (1991). Risk reduction or risk compensation? The case of mandatory safety-belt use laws. *Journal of Risk and Uncertainty 4*(1), 61–73.

Fernando E. Alvarez, David Argente, F. L. (2020). A simple planning problem for COVID-19 lockdown. NBER Working Paper 26981.

Festinger, L. (1957). *A theory of cognitive dissonance*, Volume 2. Stanford University Press.

Festré, A. and P. Garrouste (2015). Theory and evidence in psychology and economics about motivation crowding out: A possible convergence? *Journal of Economic Surveys 29*(2), 339–356.

Frey, B. S. and R. Jegen (2001). Motivation crowding theory. *Journal of Economic Surveys 15*(5), 589–611.

Grossman, G., S. Kim, J. Rexer, and H. Thirumurthy (2020). Political partisanship influences behavioral responses to governors' recommendations for COVID-19 prevention in the united states. *Available at SSRN 3578695*.

Harper, C. A., L. P. Satchell, D. Fido, and R. D. Latzman (2020). Functional fear predicts public health compliance in the covid-19 pandemic. *International Journal of Mental Health and Addiction*.

Houston, D. J. and L. E. Richardson (2007). Risk compensation or risk reduction? Seatbelts, state laws, and traffic fatalities. *Social Science Quarterly 88*(4), 913–936.

Kahneman, D. (2011). *Thinking, fast and slow.* Macmillan.

Kovacs, R., M. Dunaiski, and J. Tukiainen (2020). Compulsory face mask policies do not affect community mobility in germany. OSF, https://doi.org/10.31219/osf.io/m3sv8.

Larsen, M. V., M. B. Petersen, and J. Nyrup (2020). Do survey estimates of the public's compliance with COVID-19 regulations suffer from social desirability bias? PsyArXiv, https://doi.org/10.31234/osf.io/cy4hk.

Liu, Y., Z. Ning, Y. Chen, M. Guo, Y. Liu, N. K. Gali, L. Sun, Y. Duan, J. Cai, D. Westerdahl, et al. (2020). Aerodynamic analysis of SARS-CoV-2 in two Wuhan hospitals. *Nature 582*(7813), 557–560.

Marchiori, M. (2020). Covid-19 and the social distancing paradox: dangers and solutions. *arXiv preprint arXiv:2005.12446*.

Marcus, J. L., D. V. Glidden, K. H. Mayer, A. Y. Liu, S. P. Buchbinder, K. R. Amico, V. McMahan, E. G. Kallas, O. Montoya-Herrera, J. Pilotto, et al. (2013). No evidence of sexual risk compensation in the iPrEx trial of daily oral HIV preexposure prophylaxis. *PloS One 8*(12).

Mitze, T., R. Kosfeld, J. Rode, and K. Wälde (2020). Face masks considerably reduce COVID-19 cases in Germany: a synthetic control method approach. IZA Discussion Paper No. 13319.

Norwegian Institute of Public Health (2020). Hand hygiene, cough etiquette, face masks, cleaning and laundry - advice and information to the general public. https://www.fhi.no/en/op/novel-coronavirus-facts-advice/facts-and-general-advice/hand-hygiene-cough-etiquette-face-masks-cleaning-and-laundry/.

Ostrom, E. (2000). Collective action and the evolution of social norms. *Journal of Economic Perspectives 14*(3), 137–158.

Peltzman, S. (1975). The effects of automobile safety regulation. *Journal of Political Economy 83*(4), 677–725.

Qian, H., T. Miao, L. Li, X. Zheng, D. Luo, and Y. Li (2020). Indoor transmission of SARS-CoV-2. medRxiv, https://doi.org/10.1101/2020.04.04.20053058.

Rengasamy, S., B. Eimer, and R. E. Shaffer (2010). Simple respiratory protection—evaluation of the filtration performance of cloth masks and common fabric materials against 20–1000 nm size particles. *Annals of Occupational Hygiene 54*(7), 789–798.

Roodman, D., M. Ørregaard Nielsen, J. G. MacKinnon, and M. D. Webb (2019). Fast and wild: Bootstrap inference in stata using boottest. *The Stata Journal 19*(1), 4–60.

Rosenfeld, D. L., H. Rothgerber, and T. Wilson (2020). Politicizing the covid-19 pandemic: ideological differences in adherence to social distancing. PsyArXiv, doi:10.31234/osf.io/k23cv.

Saunders-Hastings, P., J. A. Crispo, L. Sikora, and D. Krewski (2017). Effectiveness of personal protective measures in reducing pandemic influenza transmission: A systematic review and meta-analysis. *Epidemics 20*, 1–20.

Seres, G., A. Balleyer, N. Cerutti, A. Danilov, J. Friedrichsen, Y. Liu, and M. Süer (2020). Face masks increase compliance with physical distancing recommendations during the covid-19 pandemic. Retrieved from osf.io/db8sj.

Settele, S. and C. Shupe (2020). Lives or livelihoods? Perceived tradeoffs and public demand for non-pharmaceutical interventions. CEBI Working Paper 19/20.

Stadnytskyi, V., C. E. Bax, A. Bax, and P. Anfinrud (2020). The airborne lifetime of small speech droplets and their potential importance in SARS-CoV-2 transmission. *Proceedings of the National Academy of Sciences 117*(22), 11875–11877.

Stanovich, K. E. and R. F. West (2000). Individual differences in reasoning: Implications for the rationality debate? *Behavioral and Brain Sciences 23*(5), 645–665.

Suess, T., C. Remschmidt, S. B. Schink, B. Schweiger, A. Nitsche, K. Schroeder, J. Doellinger, J. Milde, W. Haas, I. Koehler, et al. (2012). The role of facemasks and hand hygiene in the prevention of influenza transmission in households: results from a cluster randomised trial; berlin, germany, 2009-2011. *BMC infectious diseases 12*(1), 26.

Synhetsstyrelsen (2020). FAQ. 6. Should I be wearing a mouth or face mask if I am healthy? https://www.sst.dk/corona-eng/FAQ.

Thunström, L., S. C. Newbold, D. Finnoff, M. Ashworth, and J. F. Shogren (2020). The benefits and costs of using social distancing to flatten the curve for covid-19. *Journal of Benefit-Cost Analysis*, 1–27.

van der Sande, M., P. Teunis, and R. Sabel (2008). Professional and home-made face masks reduce exposure to respiratory infections among the general population. *PLoS One 3*(7).

Verity, R., L. C. Okell, I. Dorigatti, P. Winskill, C. Whittaker, N. Imai, G. Cuomo-Dannenburg, H. Thompson, P. G. Walker, H. Fu, et al. (2020). Estimates of the severity of coronavirus disease 2019: a model-based analysis. *The Lancet Infectious Diseases*.

WHO (2020). Advice on the use of masks in the context of COVID-19. WHO reference number: WHO/2019-nCoV/IPC_Masks/2020.3.

Wilson, N. L., W. Xiong, and C. L. Mattson (2014). Is sex like driving? HIV prevention and risk compensation. *Journal of Development Economics 106*, 78–91.

Wise, T., T. D. Zbozinek, G. Michelini, C. C. Hagan, et al. (2020). Changes in risk perception and protective behavior during the first week of the COVID-19 pandemic in the United States. PsyArXiv, https://doi.org/10.31234/osf.io/dz428.

Yan, Y., J. Bayham, E. P. Fenichel, and A. Richter (2020). Do face masks create a false sense of security? A COVID-19 dilemma. medRxiv, https://doi.org/10.1101/2020.05.23.20111302.

Supplementary Materials

There are two appendices. The first appendix contains the pre-registered protocol for the field experiment. (Appendix A). The second consists of robustness tests (Appendix B).

A Experimental Protocol

[16] ***Disclaimer:*** *Experimenters signed up to this experiment on voluntary basis under the condition that they do not belong to any risk groups. In order to prevent posing any further risk on any of the parties, the Robert Koch Institute's health recommendations are strictly followed.*[17]

Introduction

The instructions for the recording of data follow. Please read the whole document and follow all points very carefully.

Code of Conduct

Experimenter Appearance

As experimenter, you will need an FFP2 respiratory protection mask for this experiment. **Each time,** before you go to an experiment location, you will take two full-body (self-)portrait photos of yourself: One with and one without a mask. The primary purpose of the photos is recording variables describing your appearance if this is requested by the reviewers. To decrease the noise due to experimenter appearance, you are expected to wear a pair of blue jeans and a dark colored (black, dark gray, or navy blue) top without any visible text or logo.[18] Your outfit and mask type have to match that you used for the preceding data collection.

Location You may choose a location that satisfies the following list of conditions.

- The establishment is an open supermarket, a drug store (except pharmacy), or a post office.
- There must be a queue outside with people waiting for entering the store. The queue must stand on a flat surface with no obstructing objects. Make sure that the queue is clearly visible and it is clear for the arriving subject that you are the last person in the line and approximately where they should stand.
- You can record the data anytime between May 12-20 between 08:00-20:00 during daylight with good visibility. In order to secure good visibility conditions, **do not** record data when it is raining.
- You should avoid stores that have heavy traffic that would make measurement difficult. For instance, if there is another store or a subway exit next door, people in the queue might change their position frequently, making recording data problematic.
- The time gap between people who are let in the store must be sufficiently long. The measurement may take a couple of seconds, and you may be asked to move forward if the queue moves; the subject can also move before you can record the distance between you. The speed is usually slower at post offices than at supermarkets.
- The location you choose should be limited to those you visited during the previous data collection.

[16] There are minor differences in the two protocols. These changes are clearly marked in the text.

[17] The Robert Koch Institute (RKI) is the German government's key scientific institution in the field of biomedicine. It is one of the central bodies for the safeguarding of public health in Germany. See `https://www.rki.de/`.

[18] Please consult us if you do not own these items.

Data Recording Method You will need a smartphone with an installed augmented-reality tape-measure app that is capable of measuring small distances in centimeters with small measurement errors. The error is measured individually on the same device you use on location. Place two flat objects on the ground at any location with a clear surface exactly 100 cm from each other. Similarly to the protocol on location, measure this distance with the application. Do the same measurement five times with different positions of the objects. You may proceed with this hardware and application if the error is within a 3% margin every time.

Preparation for Data Recording In total, you are expected to perform 60 independent observations. Before each session, you set an even target of observations you are planning to record. Half of them you execute with your mask on, the other half without. The order you decide randomly using a fair coin or any random number generator. Example: You set the number to 20. After tossing the coin, you start with 10 observations with your mask on. After finishing with this, you remove the mask and perform another 10 without it. Finally, you leave the location.

The purpose of changing your appearance only once is to limit the number of times you may accidentally touch your face. You can safely avoid this if you remove the mask by only touching the strings. You should proceed the same way if you start your work without your mask on. To learn about the safe way of wearing a mask, please consult the website of the Robert Koch Institute.

Data Recording Procedure Due to lock-down measures in place, you will work alone and record the data individually. After choosing the location, go to the end of the queue outside and carefully follow this protocol.

1. Go to the queue and stand 150 centimeters (1.5 meter) away from the last person.[19] Measure this using the same application.
2. Turn sideways, neither facing the queue nor the subject arriving after you. Make sure that you can see both.
3. If necessary, calibrate your application such that it is ready for measurement. Do not open other applications at this point.
4. If someone is approaching, turn your back against the queue and face the subject before they arrive. Make sure that your face is visible, but look at your device the whole time. Keep a neutral facial expression and do not make eye contact.
5. The app measures distance by pinning two points on the ground. These two points are the closest points of yours and the subject's shoes. You pin the tip of their shoe first when they arrive, and the tip of your shoe second.
6. Record the length and exit the queue.
7. After this, record all remaining variables, starting with the number of people in the queue who were standing before you outside at the point of measurement. After this, go back to the end of the queue until you reach your target number of observations.

Further Points to Consider

If there is a group, the subject is the person closest to you, irrespective of age. Exceptions: If the closest person is an infant in a stroller or a person in a wheelchair, the closest point is where the front wheel touches the ground. If this reference point belongs to a stroller, the person you record is the one handling the stroller.

[19] Recommended minimum safe distance by the Federal Government of Germany and the Robert Koch Institute.

Do not record an observation if you are unable to pinpoint the position of the subject accurately (i.e. the subject can keep jogging in place, move back or forward before you can finish pinning) or if the subject engages in an activity that would trigger distancing according to local social norms (i.e. smoking, talking on the phone, eating).

There are 3 time slots per day: morning 8:00-12:00, mid-day 12:00-16:00, and early evening 16:00-20:00. Do not record more than 50% of the observations in one period of time (e.g. morning), even if they are recorded on different days.

Do not attempt to make any media record of the subject or any other individual near you as this may be unwelcome without consent. If you meet hostile or unfriendly reactions or you are questioned by someone, you can reveal your identity and that you are conducting a publicly funded scientific study. If this hinders or influences recording data, or puts you in an uncomfortable situation, leave the location.

You are asked to identify if there is a shop/establishment nearby that is open at the time point of measurement and accepting customers, but was legally not allowed to open in April because of the business type (e.g. nail salon, certain types of retail store). To qualify, it has to be visible and within a 50-meter radius from the point of data collection.

Data and Variables

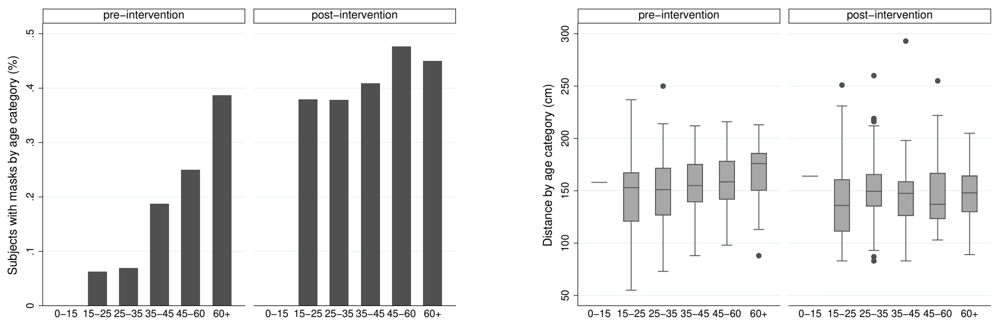
In this part, you can find the list of variables with the corresponding codes. Your task is to complete the spreadsheet for each observation. You will receive the spreadsheet by email. If you finished recording, send the file to gyula.seres@hu-berlin.de.

MaskE	Treatment variable. Experimenter 0=without 1=with mask.
Distance	Distance to the subject. Measured in centimeter (cm).
GenderS	Binary variable. Subject gender 0=male 1=female.
AgeS	Guessed age category of the subject. 0= below 14, 1=14-25, 2=25-35, 3=35-45, 4=45-60, 5=60+. If it is uncertain, write your best guess.
MaskS	Binary variable. Subject 0=without, 1=with a manufactured mask, 2=with homemade mask or improvised cover of mouth and nose (e.g. scarf).
CompanyAdult	Number of accompanying adults, 0=no adult. Adult, if age>14.
CompanyChild	Number of accompanying children, 0=no child. Child, if age<14.
TotalNumofPeople	The total number of people outside in front of you in the queue at the moment of measurement. Do not include people inside.
SocialNormS	The presence of social norm violations (i.e. smoking, food, other).
Address	Address of the experiment. For example, "Spandauer Strasse 1, 10178".
Store	Type of the store. 1=post office, 2=supermarket, 3=drug store, 4=other (please add a note)
Local	At least one business open nearby (50m) that was not allowed to open in April but is open to customers at the time of measurement. 0=no, 1=yes
ID	Surname of experimenter.
Date	Date of the month. E.g. if the date is April 20, write 20.
Time	Time of the day (i.e. 1400, 1430, etc.).
Note 1	Additional remarks, may be left empty.
Note 2	Additional remarks, may be left empty.

B Further robustness checks

B.1 Collection method

The observations were collected in sessions. Each session is defined as the target number of observations that the experimenter aims to obtain when initially approaching the line. The experimenter wore a mask (treatment group) for half of the observations and collected the other half without wearing the mask (control group). The order of the treatment and control group was randomized through a coin toss. To ensure the conditions in which the measurements were taken did not differ, we calculated the time distance between measurements in the same session. The average time between observations was 306 seconds, with a standard deviation of 275 seconds and no significant difference between the treatment group and the control group (Mann-Whitney U test z = -0.854, p = 0.3929). We believe, therefore, that no subject refrained from joining the line because of the experimenter wearing - or not wearing - a mask. On average, 5.63 people (SD=3.83) were present in the line, excluding the experimenter. Pre-mandate, the average line comprised 6.48 (SD=4.11) subjects, while post-mandate the average length dropped to 4.78 (SD=3.33). We did not detect a significant difference between the length of lines between the treatment group and the control group (Mann-Whitney U test z = 0.188, p = 0.8511). The age of subjects wearing a mask pre- and post-intervention, as highlighted in figure 3a, is substantially different. The older portion of the sample was much more likely to wear masks (38.71%) even before their use in shops was made compulsory. The percentage of mask wearers in the other age categories, instead, rose in the post-intervention period, reaching an average of 40.45% subjects aged 0 to 60 from the previous average of 13.88% in the pre-intervention period.



(a) Mask usage by age.

(b) Distance held by the subjects, by age category.

Figure 3: Distance and mask wearing by age categories.

During our measurements, we were not able to accurately record the type of mask used by the subjects. FFP2 masks, while different in substance from a surgical mask due to their filtering properties, are optically difficult to distinguish from other types of masks. We expect the subjects looking at the experimenters to have encountered the same difficulty.

B.2 Wild cluster bootstrap

Cluster-robust standard errors may be inaccurate when calculated on small numbers of clusters (see, e.g., Cameron et al. (2008)). While our data is divided in 55 clusters and, thus, is less prone to this type of inaccuracy, here we report the p-values calculated through a wild cluster bootstrap procedure as implemented by Roodman et al. (2019). This check mostly confirms the robustness of the results shown in section 5.3, table 3.

Table 4: Wild bootstrap p-values

	Distance	Bootstrap p-value
Mask Experimenter	9.392*	[0.048]
	(4.339)	
Mandate	8.082	[0.505]
	(6.014)	
Mask Experimenter ×	-2.098	[0.715]
	(4.335)	
Newly Open Stores	-3.255**	[0.022]
	(0.915)	
Online Search	0.216	[0.339]
	(0.184)	
Mask Subject	7.785**	[0.028]
	(2.514)	
Population Density	-1.071**	[0.004]
	(0.299)	
Accompanying Adult	-5.748	[0.284]
	(4.662)	
Accompanying Child	-3.821	[0.244]
	(2.487)	
# of People in Line	0.934**	[0.009]
	(0.280)	
Female Subject	-0.717	[0.832]
	(3.608)	
Aged under 15	3.835	[0.739]
	(9.351)	
Aged between 15 and 25	-8.159	[0.289]
	(5.392)	
Aged between 25 and 35	-0.986	0.842
	(3.173)	
Aged between 35 and 45	-0.670	[0.891]
	(1.911)	
Aged between 45 and 60	3.626	[0.483]
	(3.597)	
Constant	138.9***	[0.000]
	(16.71)	
Observations	480	
R^2	0.127	

Notes: Ordinary least squares estimates. Clustered errors (day, store) in parentheses. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. Mask Experimenter and Mask Subject are indicator variables for whether the experimenter or subject, respectively, used a face mask. Female Subject=1 if the subject is female. Accompanying Adult and Accompanying Child indicate whether the subject was accompanied by at least one other adult or child, respectively. Population density is based on the 2011 German Census data. Wild cluster bootstrap p-values in square brackets.

Discussion Papers of the Research Area Markets and Choice 2020

Research Unit: **Market Behavior**

Mira Fischer, Rainer Michael Rilke, and B. Burcin Yurtoglu Two field experiments on self-selection, collaboration intensity, and team performance	SP II 2020-201

Research Unit: **Economics of Change**

Anselm Hager and Justin Valasek Refugees and social capital: Evidence from Northern Lebanon	SP II 2020-301
Maja Adena and Anselm Hager Does online fundraising increase charitable giving? A nation-wide field experiment on Facebook	SP II 2020-302
Lawrence Blume, Steven Durlauf, and Aleksandra Lukina Poverty traps in Markov models of the evolution of wealth	SP II 2020-303
Kai Barron, Heike Harmgart, Steffen Huck, Sebastian Schneider, and Matthias Sutter Discrimination, narratives and family history: An experiment with Jordanian host and Syrian refugee children	SP II 2020-304
Gyula Seres, Anna Balleyer, Nicola Cerutti, Jana Friedrichsen Müge Süer Face mask use and physical distancing before and after mandatory masking: Evidence from public waiting lines	SP II 2020-305

Research Unit: Ethics and Behavioral Economics

Tilman Fries and Daniel Parra Because I (don’t) deserve it: Entitlement and lying behavior	SP II 2020-401

All discussion papers are downloadable:
http://www.wzb.eu/en/publications/discussion-papers/markets-and-choice