Face Mask Use and Physical Distancing Before and After Mandatory Masking: No Evidence on Risk Compensation in Public Waiting Lines

Gyula Seres (HU Berlin)
Anna Balleyer (University of Groningen)
Nicola Cerutti (Mercator Research Institute on Global Commons and Climate Change)
Jana Friedrichsen (HU Berlin, FU Berlin, WZB Berlin and DIW Berlin)
Müge Süer (HU Berlin)

Discussion Paper No. 300

November 16, 2021

Face mask use and physical distancing before and after mandatory masking: No Evidence on risk compensation in public waiting lines*

Gyula Seres†   Anna Balleyer‡   Nicola Cerutti§   Jana Friedrichsen¶

Müge Süer‖

November 2021

Abstract

During the COVID-19 pandemic, the introduction of mandatory face mask usage triggered a heated debate. A major point of controversy is whether community use of masks creates a false sense of security that would diminish physical distancing, counteracting any potential direct benefit from masking. We conducted a randomized field experiment in Berlin, Germany, to investigate how masks affect distancing and whether the mask effect interacts with the introduction of an indoor mask mandate. Joining waiting lines in front of stores, we measured distances kept from the experimenter in two treatment conditions – the experimenter wore a mask in one and no face covering in the other – in two time spans – before and after mask use becoming mandatory in stores. We find no evidence that mandatory masking has a negative effect on distance kept toward a masked person. To the contrary, masks significantly increase distancing and the effect does not differ between the two periods. However, we show that after the mandate distances are shorter in locations where more non-essential stores, which were closed before the mandate, had reopened. We argue that the relaxations in general restrictions that coincided with the mask mandate led individuals to reduce other precautions, like keeping a safe distance.

Keywords: COVID-19; Face Masks; Social Distancing; Risk Compensation; Field Experiment; Health Policy
JEL Codes: I12, D9, C93

1 Introduction

The novel coronavirus SARS-CoV-2, which quickly spread to almost all countries in the world has – by the middle of August 2021 – led to over 200 million confirmed infections and more than four million deaths (CSSE, 2021). To address the imminent health emergency, to achieve a sub-exponential growth of the virus, and to mitigate hospital overload, most countries implemented complete or partial lockdown policies including stay-at-home orders and travel bans during the first wave of the pandemic in spring 2020, mandated social distancing, and also encouraged personal precautions in terms of hand hygiene and respiratory etiquette. While the conjunction of these policies was proven effective, it has also become clear that lockdown measures have severe consequences for the economy and society (Fernando E. Alvarez, 2020; Thunström et al., 2020; Tull et al., 2020) such that many countries moved to lighter regimes as the pandemic progressed.

In that process, mandated face mask use played a key role as a non-pharmaceutical intervention that was first expected and then shown to be effective in combating the spread of 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; Howard et al., 2021; Chernozhukov et al., 2021). However, health authorities and politicians were initially cautious in advising the community to wear masks with reference to a potential backlash from an induced false sense of security (WHO, 2020a; Synhetsstyrelsen, 2020; Norwegian Institute of Public Health, 2020) or, put more technically, from individuals adjusting their behavior to the perceived reduction in risk.

Indeed, previous research shows that, under reasonable assumptions, individuals may react to improvements in the safety of driving by taking excess risk. This behavioral response is known as 'risk compensation.'[1] We will develop our arguments using a modified version of the model by Blomquist (1986), who shows that risk compensation is the rational adjustment of costly risk-mitigating efforts to improvements in the safety of an environment if individual effort costs increase in exogenous safety.

Risk compensation may thwart efforts to combat the spread of SARS-CoV-2 as mask use can lead to a decrease in the perceived risk of contagion in two ways. First, face coverings are believed to reduce the spread of the virus mostly because they prevent infected individuals from transmitting the virus to others. Accordingly, individuals may feel that their contagion risk is lower if someone near them is masked or if the use of masks is widespread. Second, individuals may perceive the introduction of a face mask mandate as an indication that alternative precautions – e.g., avoiding unnecessary contacts and trips, keeping safe physical distances from others – had become less relevant. Specifically, individuals may perceive using face masks as an effective mean 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) offset.

However, evidence on the existence and extent of risk compensation is limited and hardly available for the specific context of the pandemic. In a seminal paper, Peltzman (1975) finds that individual behavioral changes almost completely offset the expected positive effects on traffic deaths from

---

[1]In the literature, the term risk compensation is used interchangeably with risk homeostasis by Wilde (1982). However, the models behind these two concepts are different. Here, we adopt the definition and the model of risk compensation from Blomquist (1986). Wilde (1982) assumes that agents' preferences are only on risk levels, not accounting for effort costs. As we test for the occurrence of risk compensation, the positive framework of Wilde (1982) is not applicable.

mandated design changes in automobiles, e.g. seat belts (Peltzman, 1975).[2] Subsequently, risk compensation is also 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. The empirical results are mixed and the discussion continues as to the reliability of results from individual studies on risk compensation and the general applicability of the underlying theory (Pless, 2016; Radun et al., 2018; Mantzari et al., 2020).

We take the formal model by Blomquist (1986) as a starting point. The model describes a setting appropriate to risk associated with traffic safety. We argue that two key assumptions in the model may be violated in the context of the pandemic. First, the model assumes that improvements in external safety conditions reduce the size of the loss if an accident occurs, e.g., using a seat belt protects the one using it in an accident. Such a protective effect is less clear when it comes to masks even though mask wearing reduces the amount of virus shed: While the initial viral load is found to affect transmission (Marks et al., 2021), i.e. the probability of an accident, it remains unclear for SARS-CoV-2 whether the initial viral load of a patient affects the severity of the disease (He et al., 2020; Liu et al., 2020; Spinelli et al., 2021; Van Damme et al., 2021), i.e. the size of the loss. Second, Blomquist (1986) assumes that improved external safety conditions increase the (marginal) cost of engaging in own safety effort. However, mask use may actually decrease one's effort cost. For example, masks can serve as a reminder, thereby reducing the cognitive costs of adhering to other precautions like social and physical distancing rules. Further, masks may signal a desire for additional distance, which reciprocal individuals prefer to respect (cf. Seres et al., 2021). This is not to say that risk compensation does not occur during a pandemic in reaction to community use of masks, but the literature relies on assumptions that do not necessarily hold in this scenario.

Our study builds on the field experiment by Seres et al. (2021) but we extend the focus to the mask mandate as another potential determinant of risk compensation and the interaction between the mask mandate and the experimental variation in mask wearing. Seres et al. (2021) measured the effect of face masks on distancing in outside waiting lines in Berlin, Germany, in April 2020 before any mask mandate. Their study provides evidence against risk compensation in a context where masks were not mandatory.[3] Subjects stood further away from a masked experimenter than from an unmasked experimenter. Using an additional online survey, Seres et al. (2021) argue that this behavior is related to masks sending a social signal: People expect individuals who wear a mask to prefer others to stay farther away from them, which they then do. However, these results cannot be easily extrapolated to a situation with a mask mandate, which has become a common policy. The mandate might change the perceived exogenous safety outside. It increases the fraction of individuals wearing a face mask and those wearing face masks have a lower probability of transmitting the virus if they are infected (Chernozhukov et al., 2021), thus reducing overall infection risk. If risk compensation occurs, subjects are expected to exert less effort in maintaining a distance with a mandate and the treatment difference in distances kept to masked and unmasked others may change as a result.

---

[2]Note that accident risk when driving and desirable driving outputs like reduced travel time are arguably complements. A condition that is akin to effort cost increasing in exogenous safety.

[3]The literature also does not provide support for risk compensation on hand hygiene, the other main preventive measure. Mantzari et al. (2020) present a meta-study of six empirical studies on the effect of community use of face masks on hand hygiene. They find that all cited works find either no effect or a positive significant effect on hand washing and sanitizer use.

To assess whether a mandate affects distancing or the observed change in distancing when queuing after a masked person, we exploit the first-time introduction of an indoor mask mandate in Berlin in late April 2020. We repeated the field experiment from Seres et al. (2021) after the introduction of the first indoor mask mandate in Berlin and analyze the collected data jointly with the original data. As the subjects are unaware of the experiment, not only can they not form beliefs about the study or adjust their behavior concerning any perceived aim, they also cannot self-select into the experiment.[4] By construction, however, our sample only contains subjects who were willing to leave their residence despite the risk posed by the ongoing pandemic in April and May 2020.[5] We should further note that the introduction of the mask mandate was packaged with other regulatory changes, including the (restricted) reopening of non-essential businesses, and that infection numbers, even though relatively low throughout the entire study, were lower in the second phase of data collection than in the first (see footnote 5).

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, we present evidence showing that the presence of a masked person does not create risk-compensating behavior in either context. However, our results also reveal that the average distance that the subject kept is significantly lower in the post-mandate period than in the pre-mandate period. One reason could be that individuals, who understand that the mandate increases mask use, compensate for the implied decrease in general risk by taking more risk individually, for instance by standing closer to each other. However, using contextual data in the form of the number of open shops in the surrounding, we argue that the observed reduction in distances is likely not driven by the indoor mask mandate but by concurrent regulatory changes. In particular, more shops were open, leading to more pedestrians on the streets and potential crowding. In principle, the easing of general restrictions that accompanied the mask mandate might also have reduced the perceived risk from the virus and, thus, led to less cautious behavior.[6] However, data from a snapshot monitoring do not show notable changes in risk perception (Betsch et al., 2020a), which would be driving behavioral adjustments. This might be the case because other indicators did not suggest much relief between April and May 2020. For instance, available hospital capacity for intensive care in Berlin slightly decreased from on average 400 free beds during the first to 390 during the second data collection period (DIVI, 2021).

Our results complement further evidence obtained during the pandemic from other geographic locations mostly showing that risk compensation does not occur in response to masking or mask mandates. Empirical studies examine social distancing in terms of time spent outside and proximity during this period concerning mask wearing. Kovacs et al. (2020) use location data from Germany to show that imposing indoor face mask mandates in Germany did not lead to a compensatory effect in individuals' mobility patterns in terms of time spent outside. Liebst et al. (2021) utilize a natural experiment approach to compare video-observational data from two Dutch cities and find no

---

[4]We also show empirically that the experimental treatment (experimenter wears a mask or not) does not influence the subjects' decision whether to enter the waiting line or whether to wear a face mask themselves (see Subsection 5.2).

[5]Actual risk was arguably quite low at the time: According to data from the Berlin health authorities, the average number of daily new infections was about 63 during the first phase of data collection and 31 during the second phase of data collection, resulting in an average 7-day incidence of new infections per 100,000 inhabitants of 15 during the first and 6 during the second phase (Berlin LAGeSo, 2021).

[6]An array of studies shows 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).

evidence supporting the risk compensation hypothesis. Marchiori (2020) shows that wearing a face mask can substantially improve adherence with the physical distancing regulations on pavements in Italy, both in the absence of a mask mandate and after its introduction. While their study focuses on moving subjects with distance defined as the closest distance reached when the subject passes the experimenter, this result is qualitatively comparable to that of Seres et al. (2021) where subjects were standing in line suggesting that the effect might be robust across settings. A randomized controlled trial in Bangladesh that studies interventions to increase mask uptake also provides evidence that mask use and social distancing are complementary: the interventions that increased mask wearing were found to also improve adherence with distancing recommendations (Abaluck et al., 2021). Furthermore, a meta-analysis (Bakhit et al., 2021) concerning the effects of face masks and an online study (Guenther et al., 2021) investigating risk-taking behavior during the COVID-19 pandemic also do not find any evidence supporting risk compensation. We are aware of only one study supporting risk compensation in response to mandatory masking: Yan et al. (2020) argue that US Americans spent more time outside their homes after masks became mandatory in public spaces. However, their empirical strategy does not preclude that mobility would have changed similarly even with the absence of masks.

The rest of this paper proceeds as follows: Section 2 provides the setup, including the local progress of the pandemic, and a general introduction to the policy environment. Section 3 describes the experimental design. Section 4 introduces a formal model of risk compensation and states the hypotheses. Section 5 provides the main results of this paper. Section 6 concludes with further interpretations of the results and a discussion.

2 Background

The first indoor face mask mandate was introduced in all German states toward the end of April 2020 and coincided with the relaxation of other regulations against the spread of COVID-19. In Berlin, starting in mid-March 2020, only supermarkets and stores selling basic necessities were allowed to open. April 22, 2020 witnessed the reopening of small retail stores ($< 800m^2$) under certain restrictions (e.g., a limited number of concurrent customers). Although 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 starting dates. Berlin introduced mandatory masking in shops on April 29. In outdoor areas, the masks were (and, as of August 2021, still are in Berlin) recommended in places where keeping a minimum of 150cm distance is not possible. The objective of the indoor mask mandates was to reduce the risk of contagion in increasingly frequented places where physical distance recommendations are harder to uphold, e.g., shops and public transport. However, individuals may adjust their precautions in other dimensions in response to this mandate such that the net effect is, *a priori*, not clear. Therefore, we follow up on Seres et al. (2021) with an identical field experiment, conducted in Berlin *after* the introduction of compulsory masking in shops and public transport, evaluating the effect of masks and the interaction with other policy changes on distance keeping in waiting lines. Table 1 displays the detailed timeline of the experiment and the restrictions.

	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	**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 and study design

3.1 Location

The study took place in 2020 during the COVID-19 pandemic in Berlin, Germany.[7] The first field experiment was conducted before the face mask mandate and the replication was carried out after its introduction. During the first data collection period, acceptable reasons to leave the place of residence were defined at the state level and limited the mobility of residents, including the experimenters. To comply with public health recommendations and mobility restrictions in Berlin, store selection aimed at avoiding long commuting from the experimenters' homes.[8] There was no overlap in the list of stores between experimenters, hence, only one of them visited each store in the sample. During the second phase of data collection in May, the experimenters revisited the same stores as in April. The experimental protocol restricted the store types, in front of which observations took place, to supermarkets, drug stores (except pharmacies), and post offices to collect data from a sample that represents the population visiting public areas. During this first wave of the pandemic in Berlin, lines in front of businesses were frequent but occurred randomly

---

[7]According to the 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 rate in Berlin was 157 per 100,000 inhabitants; close to the federal average 197 per 100,000 inhabitants.

[8]Figure 4 in Appendix C shows the locations of businesses visited.

depending on traffic. Therefore, only the lines that existed at the moment of data collection could be utilized. We address potential randomization concerns regarding store selection in Section 5.

3.2 Methodology

Our study consists of a repeated field experiment that we conducted before and after the introduction of an indoor mask mandate in Berlin. In the field experiments, we randomized the use of face covering. In the MASK treatment, the experimenter was wearing a mask, whereas, in the NOMASK treatment, no face covering was used. Data was collected between April 18–24 (pre-mandate data collection) and between May 12–20 (post-mandate data collection). In both periods, 60 observations were recorded by each of the four experimenters. This adds up to 240 in each period and 480 in total. The pre-registered experimental protocol is included in Appendix A.[9]

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.[10] With this setup, following the methodology of Balafoutas and Nikiforakis (2012) where the researchers were also the experimenters, we avoid potential ethical concerns from financial or hierarchical obligations of the experimenter toward the authors. Each experimenter recorded the observations individually in their own neighborhood, widely spread throughout Berlin (see Figure 4). Two techniques were employed to reduce 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.[11] Second, the dress-code was standardized to a pair of blue jeans and a dark-colored top (Balafoutas and Nikiforakis, 2012).

Procedure Each experimenter independently located a line outside a shop in their neighborhood and determined an even number of observations to be collected there, denoted by N. The experimenter started data collection wearing a mask (treatment MASK) or not (treatment NOMASK) based on the result of a coin toss. Then, the experimenter joined the line, maintaining a distance of 150cm 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, they assumed a sideways position in the line, thus ensuring their face would be visible to the next person but avoiding eye contact. Upon the subject's arrival, the experimenter measured the distance between their 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 a 1-centimeter precise

---

[9]The pre-mandate data was used by Seres et al. (2021), which had five experimenters. In this paper, we use data from four experimenters who participated in both measurement periods. This study was pre-registered with five experimenters but one was then unable to participate in the second phase of data collection.

[10]Experimenters being co-authors of the study might raise questions regarding the conscious or unconscious effects on outcomes. We believe these concerns are unfounded as the study design of Seres et al. (2021), which we simply replicate after the mandate, was preregistered stating no expectations regarding the outcome. The post-mandate data mirrors the findings of the first period, partly contradicting our hypotheses, which goes against potential concerns of experimenter influence on the outcome.

[11]An FFP2 mask is a mechanical filter respirator as defined by the EN 149 standard, similar to the N95 design. At this time, surgical masks were in short supply. As the public use of both FFP2 and surgical masks was uncommon before the outbreak, we do not expect that carrying out the experiment with surgical masks would have altered the results.

measurement. 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 if the continuation of the line was unobstructed. Having completed the input of data several meters away from the line, the experimenter returned to the end of the line to wait for the next subject arriving and continued with the same treatment until he or she had collected $N/2$ observations, switch to the other treatment and collect another $N/2$ observations to achieve balance.

Thus, at any store, an equal number of observations with and without a mask were collected. At each visit, the experimenter used a coin toss to determine which of the two treatments to start with. Suppose the experimenter had decided to collect N observations and that they would start with mask on if the coin toss turned out to be tails. Then, if the coin showed tails, they would start with the mask on and collect $N/2$ observations masked. Then, they would take off the mask and collect another $N/2$ observations unmasked, concluding a session. This procedure ensured a balanced set of observations across treatments.

Study design   The study consists of two main observational periods, the first taking place before a set of policy changes, including the introduction of mandatory mask-wearing in stores, and the second one afterward. This two-phase study design allows us to assess the impact of the policy change on distancing by comparing distancing in both the pre- and post-policy periods. As the mandate was brought into force at the same time as several other measures and could not be experimentally varied (see Table 1), it is hard to isolate its pure effect, which is what we are most interested in. However, we argue that the mask mandate and relaxation measures affect our dependent variable, kept distance, in different ways and try to separate these in the empirical analysis. From this point on, the difference between periods is referred to as the effect of policy changes, with the different policies (mask mandate, shop re-openings) underlined separately as necessary.

To control for the impact of the relaxation in store closure policy, the experimenters recorded – within a 50-meter radius around their location of data collection – the number of non-essential businesses that were open at the time of measurement during the second round of data collection in May but were legally closed during the first round of data collection in April.[12] In the May sample, this variable shows substantial variance as it ranges from 0 to 6 (with $\mu =2.54$ SD$=2.26$). By construction, its value is always 0 in the April sample as all non-essential businesses had to be closed at the time.

The experimenters also collected information on the subjects' demographic profiles. In particular, in addition to guessing the subjects' age (coded in age groups) and their gender, experimenters recorded the number of accompanying children and adults 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 indoor mask mandate, all subjects presumably had a mask with them as a prerequisite for entering the store, unlike before the mandate. However, no law mandated using the mask while waiting outside at the time.

We also collected data on the length of the line (number of people in the line in front of the experimenter), population density at the postal code level, store type, and the store's address.

---

[12]This variable was preregistered as "*Local*." For the sake of comprehensibility, we renamed the variable as "*Stores*" when writing this manuscript.

4 Theoretical Background and hypotheses

Regarding the COVID-19 safety measures, critics argue that community use of face mask usage might create a false sense of security reducing compliance with other precautionary measures. In the context of our experiment, from the subject's point of view, both the indoor mask mandate and the experimenter's face mask can be interpreted as determinants of exogenous safety (or put differently, as determinants of the riskiness of the situation). In the setting of a viral pandemic, however, we argue that it is unclear whether an increase in exogenous safety decreases individual effort in risk prevention, as is observed with car safety (Peltzman, 1975), or whether both may be complements.

In this section, we present a stylized model of utility maximization that is adopted from Blomquist (1986). We show that, indeed, a decrease in infection risk through improvements in exogenous safety, e.g. in the form of a mask mandate, need not lead to risk compensating behavior. Instead, we argue that improved safety may even lead individuals to increase their own safety efforts if the mandate reduces an individual's cost of complying with distancing regulations by making the risk from the virus more salient. Without clear evidence on the effect of masking on the cost of adhering to social distancing rules, the theoretical prediction is ambiguous and our empirical study provides the direction of the actual effect. Lastly, we elaborate on the research questions that we test empirically and lay out the preregistered hypotheses.

4.1 A Modified Risk Compensation Model

In the model of Blomquist (1986), which is part of the larger literature on risk compensation in traffic safety measures, an agent faces disutility from her own effort e and from exogenous safety measures s. She might also incur a further utility loss (L) from an "accident" – a stochastic event that occurs with probability p(e,s), decreasing in e and s. The agent chooses her personal effort level to maximize her utility. The utility function of the agent is the expectation

$$U = I - D(e, s) - p(e, s)L \tag{1}$$

where $D(e, s)$ is the value of discomfort and I is income.

For simplicity, we assume that the loss L, which decreases in e and s in Blomquist (1986), is a constant: independent of individual effort and safety measures. While a traffic accident's severity can be mitigated through effort and exogenous safety measures, this is less clear when it comes to the severity of an infection with SARS-CoV2.[13] Specifically, it is unclear whether a reduction in the initial viral load of a patient, which could be achieved e.g. through masking, actually leads to a less severe progression, conditional on becoming infected (Liu et al., 2020; He et al., 2020).[14] Blomquist (1986) assumes that exogenous security measures do not affect the disutility or marginal cost of own effort. On the contrary, in the context of the COVID-19 pandemic, these

---

[13]There is little research in this direction. Gandhi and Rutherford (2020) provide empirical evidence that masks lower severity, Courtney and Bax (2021) claim that this effect is caused by higher hydration of the respiratory track due to face masks.

[14]Note that the argument is about the severity of disease among the infected. There appears to be consensus that community masking reduces the probability of becoming infected as discussed in the introduction. New findings about the effect of viral load on the severity of infection might alter the model as $L_e = 0$ and $L_s = 0$ may no longer hold.

exogenous measures might help individuals to comply with the safety recommendations as they act as a reminder of the severity of the situation. Thus, in our setting, endogenous safety efforts and exogenous measures could become complements.

We derive the conditions under which an increase in exogenous safety will lead to a decrease in individual effort to remain safe (condition for risk compensation) under the modified model assumptions. Let the probability of becoming infected be a continuous and twice differentiable probability function $p(e, s)$ with $p_e < 0$, $p_{ee} > 0$, $p_s < 0$, $p_{ss} > 0$ and $p_{es} > 0$, analogous to Blomquist's (1986) accident probability, and be I an exogenous income. Let further the loss from getting infected be given by the constant L and denote the disutility from effort by $D(e, s)$ with $D_e > 0$, $D_{ee} > 0$, $D_s \geq 0$, and $D_{ss} \geq 0$, noting that we here differ from Blomquist (1986).

The first-order condition from utility function (1) with respect to the endogenous effort e is:

$$- D_e = p_e L. \tag{2}$$

From (2), the effect of a change in exogenous safety conditions on personal effort as an implicit function is:

$$\frac{de}{ds} = \frac{D_{es} + p_{es}L}{-D_{ee} - p_{ee}L}, \tag{3}$$

where the denominator is negative and coincides with the second order condition of utility maximization. Hence, the condition for an individual to be risk compensating ($\frac{de}{ds} < 0$) becomes:

$$D_{es} + p_{es}L \geq 0. \tag{4}$$

While risk compensation in the original model of Blomquist (1986) follows from the assumptions, it may not occur here. Intuitively, if higher external safety level s reduces the marginal cost of effort e to an extent large enough ($D_{es} < 0$ and sufficiently large in absolute value) to outweigh the reduced effectiveness of effort ($p_{es}L > 0$), the expression in (4) becomes negative, so that the optimal choice of e may increase in s ($\frac{de}{ds} > 0$ in (3)). In other words, risk compensation does not occur.[15]

Note that external safety condition s can be interpreted as a vector in the context of the pandemic. We revisit this idea in the next subsection. Further, it should be noted that the model is consistent with the understanding that individuals' decisions depend not on actual effort cost, disease severity or safety but rather on subjective perceptions of effort costs and loss from getting infected as well as the subjective riskiness of the situation. Thus, D, L, and s do not need to reflect objectively true values but individuals' subjective beliefs about them. The publicly available information on these concepts that we discuss in this paper constitutes a reference case that we expect to apply to individuals with rational expectations. It is easy to see from the model that differences in beliefs may lead to heterogeneity in individuals' responses to masks as well as policy changes. Due to lack of information on the relevant subjective beliefs of our subjects, we cannot follow-up on this aspect of the model empirically within the context of our experiment. We discuss this issue and implications for further studies in Section 6.

---

[15]New findings about the effect of face masks on the severity of a COVID-19 infection, e.g. that $L_e > 0$ would not alter this conclusion as the sign of $\frac{de}{ds}$ would still depend on the sign of D_{es}. Hence, we omit further generalizations for simplicity.

4.2 Hypotheses

We investigate three research questions.[16] First, we ask whether distancing behavior is different in the pre- and post-policy periods. We expected the effect of a mask mandate to differ between the experimental treatments.

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

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

These hypotheses are partly based on insights from (Seres et al., 2021). There, we used an online representative survey of a German sample to understand why the MASK treatment increases physical distancing. We find a potential explanation in people believing that a person wearing a mask prefers others to keep a greater distance.[17] This signaling mechanism is unrelated to the arguments related to risk compensation and could still be at play in the post-mandate phase as our experiment takes place outdoors and masks are only obligatory in stores. Additionally, face masks may have a reminder effect that may differ pre and post mandate. While Seres et al. (2021) find no evidence in this regard pre-mandate, it is possible that seeing someone wear a mask does remind people of the risk from COVID-19 and induces them to engage in appropriate precautions post-mandate. Specifically, the introduction of a mandate may have informed the population that the situation is still risky (Seegert et al., 2020) and distancing necessary, such that seeing a mask reactivates this knowledge and triggers more distancing.

In the context of the model, it is helpful to think of exogenous safety s as being influenced by the experimenter being masked (s_1) and the mask mandate (s_2). Perceived exogenous safety s that affects behavior increases in both s_1 and s_2. However, the mask mandate does not have an additional effect on perceived safety or a reminder effect if the subject is in condition MASK and the MASK treatment does not further affect perceived safety when a mandate is in place, $\frac{\partial^2 s}{\partial s_i \partial s_j} = 0$ for $i, j \in 1, 2$ and $i \neq j$. Thus, distancing in MASK should not change with the introduction of the mask mandate as it does not further change perceived safety. Intuitively, if the MASK treatment already triggers increased precautions, the mandate should not influence distancing in this treatment condition because the signaling mechanisms already identified by Seres et al. (2021) should also be unaffected by the policy. However, in NOMASK, we would expect distances to be greater post-mandate if the cross-derivative D_{es} is negative because s increases with s_2: the mandate may activate additional precaution as its introduction can be understood as telling the population that, at this stage of the pandemic, individual precautions are crucial.

Second, we expect the subjects in the post-policy sample to keep, on average, the same distance from the experimenter in both treatment conditions. The idea is the following: the introduction of the mask mandate required people joining lines to carry a mask with them and led to an increase in general mask use in the line. A potential reminder or salience effect is then already realized such that the treatment becomes ineffective. In the model context, this is also a consequence from

---

[16]For narrative purposes, the order and wording of the hypotheses are different from those in the preregistration. Two additional hypotheses from the preregistration will be addressed in a separate study as was pre-specified.

[17]In the survey experiment, respondents were shown the pictures of the experimenters from the field experiment either in the MASK or the NOMASK treatments and asked to state how far away the pictured person would prefer others to stand from them in a waiting line.

assuming $\frac{\partial^2 s}{\partial s_2 \partial s_1} = 0$ A convergence in distances across the two treatments could alternatively result if masks increase distances mostly through social signaling pre-mandate (cf Seegert et al., 2020) but their informational value with respect to preferred distances is lost with the introduction of the mandate and the associated increase in masking.

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.*

Our last hypothesis concerns the robustness of the positive effect of subject masking on distancing as found by Seres et al. (2021), where subjects wearing a mask kept a greater distance. Pre-mandate, few subjects wore a mask and were seen as preferring larger distances. The mask mandate increased the use of face masks in the waiting line, so that mask wearers post-mandate are supposedly different from mask wearers in the pre-mandate sample. Specifically, preferences and, thereby, distancing behavior of mask wearers might be much more similar to that of non-mask wearers after the mandate because many who did not wear a mask before the mandate decided to do so post mandate.

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).*

5   Empirical analysis

5.1   Sample Characteristics

The data set contains 480 observations, collected in sessions. Each session is defined as the target number of observations that the experimenter aimed 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. The times between individual measurements do not differ by treatment, suggesting that potential subjects did not refrain from joining the line because of the experimenter wearing – or not wearing – a mask.[18]

Table 2 contains descriptive statistics on subject characteristics across conditions, giving no hint of irregularities or imbalances across conditions. Note that our study features a convenience sample of those being outside at the time and should not be expected to be representative. While or sample reflects the general gender balance well (54.4% vs. 50.8% in the population, $\chi^2 = 2.45$, p = 0.117), compared to the city's age distribution, our sample underrepresents the group of those aged 60+ (10.6% vs. 24.7%). This is to be expected as age was identified and publicized as a risk factor for a severe COVID-19 infection, hospitalization and death already early on (Verity et al., 2020; O'Driscoll et al., 2021) meaning that the elderly should rationally stay home more often. Further, a fraction of them may live in retirement homes or not do their own shopping, so they would not have been in these waiting lines even before the pandemic.

---

[18]The average time between observations was 320 seconds, with a standard deviation of 335 seconds and no significant difference between the treatment group and the control group (Mann-Whitney U test z = -0.926, p = 0.3547). We also test whether the sample correlation coefficient between the number of people in the waiting line and the treatment variable is significant. The reasoning is that if joining the line was influenced by the treatment, this correlation would be significant. This claim is rejected (r = -0.0659, p = 0.149).

Count	Pre-Mandate		Post-Mandate		
	NoMask	Mask	NoMask	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.

Most subjects in our sample visited the store alone, only 10.6% came with an adult and 6.1% with minor companions. We also recorded the length of the line measured as the number of people standing outside in front of the experimenter. The mean length is 6.48 (SD=4.11) pre- and 4.78 (SD=3.33) post-mandate. 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).

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; this squares well with survey evidence from Betsch et al. (2020a,b) that mask use in the general public increased from 24 to 73% over the same period. The age of subjects wearing a mask pre- and post-policy, as highlighted in figure 1, 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-policy period, reaching an average of 40.45% subjects aged 0 to 60 from the previous average of 13.88% in the pre-policy period.

5.2 Main results

We structure the discussion of our results around the hypotheses. First, we observe that, on average, subjects in both treatments keep a shorter distance from the experimenter in the data collected after the policy change than before it. While subjects kept an average distance of 151.14cm (SD=29.62) to the unmasked experimenter in our pre-mandate sample, the average distance to the unmasked experimenter is significantly shorter, only 143.35cm (SD=31.79) in the post-mandate sample (two-sample t = -2.1889, p = 0.015). Similarly, subjects kept an average distance of 159.85cm (SD=31.79) to the masked experimenter in the pre-mandate sample, but only 151.41cm (SD=34.08) in the post-mandate sample (two-sample t = -1.8941, p = 0.029). Thus, the distances kept are, on average, 7.79cm (NoMask) and 8.41cm (Mask) shorter in the post-policy period than
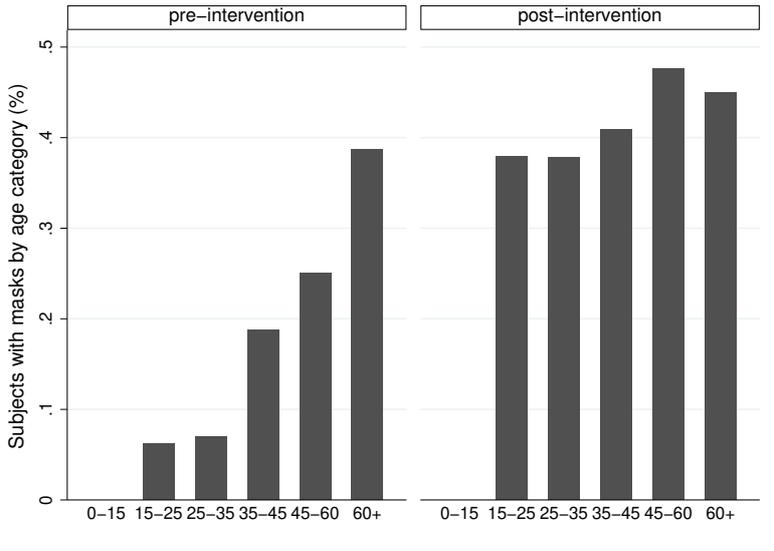
Figure 1: Mask usage by age.

in the data collected before the policy change. Figure 2 summarizes these observations. 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 shorter after the policy change in both treatments* MASK *and* NOMASK. *Hypotheses 1.A and 1.B are not supported.*

We then turn to regression analysis to better understand where these differences come from. We want to note that the mask mandate came alongside other measures that eased the restrictions put on social and business activities. Thus, we interpret the pre- and post-policy differences as the joint effect of the mandate and relaxations, naming it accordingly in our model and control for the relaxation measures.[19].

Pooling the entire sample, we estimate the following equation:

$$Distance_i = \beta_0 + \beta_1 MaskE_i + \beta_2 Policy_i + \beta_3 MaskE_i \times Policy_i$$
$$+ \beta_4 MaskS_i + \phi X_i + \varepsilon_i \qquad (5)$$

in which $Distance_i$ is the distance kept by subject i, $MaskE_i$ is an indicator for 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. The treatment effect from the field experiment is given by β_1, while β_2 captures any effect in distancing that results from the conjunction of policy changes between the first and the second data collection periods. The effect of the mask mandate jointly with other policy changes on the treatment effect is given by β_3. If we cannot reject $\beta_3 = 0$, this implies that the face mask effect on distancing is not significantly different between the two periods. X_i is a vector of all other covariates and controls used in different specifications of the model. Standard errors ε_i

---

[19]Please note that any interpretation of policy effects should be undertaken with caution as the time gap between observations might have come with other co-occurring changes not accounted for in the model: norm changes, supply change of masks, hospital capacity, change in infection rates. We revisit this in the discussion in Section 6
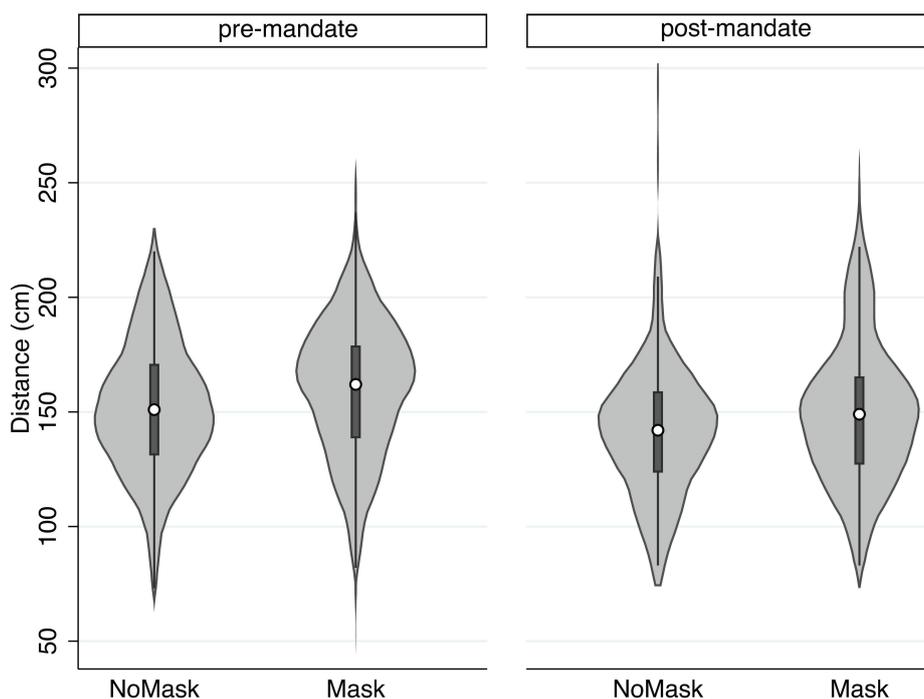
Figure 2: Summary of distance kept from the experimenter. The white dots represent the averages, the gray bars the interquartile ranges, and the light gray areas the kernel density of the distributions

are clustered according to store and date in order to mitigate concerns with respect to a potential correlation in error terms.[20]

Note first that results in Table 3, column (1) confirm Result 1: the coefficient on policy, β_2 is significantly negative while the coefficient on the interaction term between policy and the treatment dummy $MaskE$, β_3 remains insignificant. In line with the observed difference in distancing, column (1) shows that distances toward both the unmasked and the masked experimenter are about 9 cm shorter after the policy changes had taken effect. Further, the coefficient on $MaskE$, β_1, is positive and significant. As the interaction between treatment and the policy change is not significant, this shows that subjects keep significantly greater 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.

In columns 2 to 4 of Table 3, we estimate modifications of equation 5 where we add additional explanatory variables individually and in combination. Using these additional specifications, we argue that the observed shift in behaviors may be explained by a combination of factors including the reopening of non-essential businesses. Specifically, we include the variable *Stores*, measuring the number of businesses within a 50-meter radius of the point of data collection that were legally closed in April but were open in May at the time of measurement. Further, changes in distancing may also be influenced by the general perception of the infection risk during the pandemic. As

---

[20]Specifically, the clustered standard errors are used to address any potential serial correlation in the error terms due to clustered sampling. As we are considering a relatively small number of clusters (55 in total), we also perform a wild cluster bootstrap method following Cameron et al. (2008) as a robustness check. Please see Appendix B for details. As the experimenter locations in our sample do not overlap, our approach also covers experimenter-related correlations.

Table 3: Ordinary least squares regression of distances kept by subjects on treatment and policy variables.

	(1) Pooled	(2) Pooled	(3) Pooled	(4) Pooled	(5) Pooled	(6) Pooled
MaskE	9.222*	9.449*	9.173*	9.392*	9.447*	9.677*
	(4.216)	(4.266)	(4.221)	(4.273)	(4.317)	(4.365)
MaskE × Policy	-1.614	-2.250	-1.481	-2.098	-1.121	-1.730
	(5.557)	(5.458)	(5.610)	(5.502)	(5.562)	(5.443)
Policy	-9.294*	-0.0833	-3.055	8.082	-8.213	0.866
	(4.610)	(6.407)	(9.891)	(10.88)	(4.616)	(6.537)
Newly Open Stores		-3.173**		-3.255**		-3.115**
		(1.177)		(1.198)		(1.149)
Online Search			0.170	0.216		
			(0.236)	(0.209)		
MaskS	7.376*	7.623*	7.499*	7.785*		
	(3.251)	(3.298)	(3.302)	(3.356)		
Population Density	-1.423***	-1.066**	-1.435***	-1.071**	-1.314***	-0.960**
	(0.309)	(0.330)	(0.300)	(0.315)	(0.301)	(0.329)
Accompanying Adult	-5.442	-5.788	-5.403	-5.748	-6.096	-6.458
	(4.999)	(4.944)	(4.992)	(4.933)	(5.028)	(4.969)
Accompanying Child	-4.444	-3.789	-4.482	-3.821	-4.271	-3.622
	(3.126)	(3.254)	(3.163)	(3.296)	(3.075)	(3.213)
People in Line	0.865**	0.938**	0.860**	0.934**	0.759*	0.827*
	(0.304)	(0.309)	(0.296)	(0.303)	(0.312)	(0.319)
Constant	165.2***	156.5***	151.5***	138.9***	167.0***	158.5***
	(5.932)	(6.701)	(20.15)	(18.32)	(5.802)	(6.699)
Observations	480	480	480	480	480	480
R^2	0.107	0.125	0.109	0.127	0.098	0.115

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 corrected for clustering at day and store level.

a measure of this factor, we gathered daily data from Google trends, the variable *Online Search*, representing the relative number of Berlin-specific hits on Google search for the novel coronavirus on the day of measurement.[21]

The results in columns 2 and 4 suggest that the decline in distancing after the policy change can be explained by reopening stores rather than the introduction of the mask mandate.[22] According to specification (4), an additional newly opened store near the location of measurement is related to a decrease in distancing of 3.26cm on average (p=0.002). To counter concerns that the variable *Stores* simply picks up an area-specific characteristic, we also reestimate the specification from column (4) in Table 3 on the pre-policy sample alone and replace the pre-policy values of *Stores* with the maximum value of *Stores* observed in the post-policy data at the respective location. If *Stores* is just a measure of location characteristics that determine the distance people keep, the estimated coefficient in this model should become significant. However, we find that the coefficient of the recorded variable *Stores* is not significantly related to distances kept in the pre-policy sample (z = .9702845, p = 0.206).[23] There is no evidence that online search for pandemic-related content

---

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

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

[23]Another possible channel influencing distancing could be the interaction between the length of the waiting line and the number of reopened stores. Removing the number of people in line from specification (4) does not change

as a proxy for concern with the virus predicts the difference between distances in the pre- and post-mandate samples well. Virus-related internet traffic (*Online Search*) positively correlates with greater distancing but the effect is not significant (p=0.257). This is consistent with evidence from a national survey reporting that the perceived risk did not substantially change over the time when we collected data (Betsch et al., 2020a). We conclude that the ease in business restrictions is likely to be the driving force behind the reduction in kept distances. A potential explanation is that more open businesses create more opportunity for people to be outside and the increase in pedestrian movement makes it harder to keep sufficiently long distances from others even when waiting in line.

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 9cm in all specifications and significant. The effect is robust across time: the interaction with the policy change is negative, as predicted, but 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, we reject Hypothesis 2. Further, in the post-mandate sample, subjects kept greater distances from the masked than from the unmasked experimenters, on average. Qualitatively and quantitatively, Result 2 confirms the main conclusion of Seres et al. (2021). It can also be seen as a replication of the previous result after the introduction of the mandate.

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

We now turn to Hypothesis 3 that, in the post-mandate period, the subsample of subjects wearing a mask themselves do not react differently to the masked experimenter than the rest of the sample. The share of masked subjects is indeed much higher post-mandate than pre-mandate, even though masks were not mandatory in outside waiting lines (see Table 2).[24] The use of masks is significantly higher post-mandate: the share of masked subjects rises from 17.1% to 40.8%, (γ_2=0.9036 p=0.014). This surge in mask use is unlikely to be driven by preferences but rather is a consequence of the need to wear a mask upon entering the store.[25] As a consequence, we expect that mask wearing became less predictive of an individuals' distancing post-policy.

We find that subjects with a face mask keep a significantly greater distance from the experimenter than unmasked subjects in the pre-mandate sample (two-sample t=-2.3788, p=0.009). This is consistent with the idea that individuals self-select into mask wearing based on their intrinsic concern with respect to the pandemic. Those who are more concerned decide to wear a mask and also keep greater distances to others. But in the post-mandate sample, we cannot reject the null hypothesis that subjects with a mask keep the same distance on average than those without a mask (two-sample t=-0.6327, p=0.264). This is consistent with an encouragement effect from the mandate that makes people wear a mask in the waiting line even though their intrinsic concern with the pandemic alone, which also affects their distancing, would not make them wear a mask. As a consequence, the subsample of mask wearers post-mandate is less selective with respect to pandemic concerns and becomes more similar to the population of unmasked individuals than is the

---

the significance of the variable *Stores* (z=.-3.0223, p=0.015).

[24]The Senate of Berlin made masks mandatory in public waiting lines only in fall 2021, long after data collection was completed.

[25]This increase in mask use is in line with the development in the entire population: according to results from a national survey during our second phase of data collection, about 80% of respondents were wearing face masks in the public, up from 33% during our first phase of data collection (Betsch et al., 2020b,a).

subsample of mask wearers pre-mandate. When we run a regression on the post-mandate sample alone (using specification (2) from Table 3 without *Policy*) we confirm that masked subjects do not keep longer distances ($\beta_4 = 6.5871$, p=0.156). Thus, our data is consistent with Hypothesis 3.[26] However, it is worth noting that self-selection into mask-wearing does not allow for a causal interpretation of the estimated coefficient on *MaskS*.

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

Considering all factors that might be driving this result, in the following paragraphs we scrutinize the potentially relevant ones and explore further relationships between the observed variables. As the discussed models above control for mask-wearing by the subject, the coefficient of *MaskE* can be interpreted as the direct effect of observing others wearing a mask. However, the difference in distancing between the two periods might also be affected by the sharp increase of mask-wearing in the sample that occurred with the policy change. This indirect effect might be positive if those now adopting masks also pay more attention to distancing or the prevalent masks remind others to do so, but it could also be negative if individuals risk compensate. To account for the total effect, we estimate specifications (5)-(6) in Table 3 that are identical to (1)-(2) except that they omit *MaskS*. As evidenced by the similar coefficients, the difference between the pre- and post-mandate sample is driven by the direct effect rather than an indirect effect through the increased prevalence of masks.

It is also important to rule out the possibility that subjects might be deciding whether to wear their masks after seeing the experimenter being either masked or unmasked. Equation (5) is correctly specified only if the treatment dummy *MaskE* does not influence the subject's decision to put on a mask, *MaskS*. We argue that the exogeneity of *MaskS* is given as subjects decide on their use of a mask before seeing the experimenter. However, this decision may be reversed upon seeing the experimenter. We test the independence claim with the following logit binary choice model:

$$Pr(MaskS = 1) = \frac{\exp(\gamma_0 + MaskE + \gamma_2 Mandate + \gamma_3 MaskE \times Mandate + \phi X_i + \varepsilon_i)}{1 + \exp(\gamma_0 + MaskE + \gamma_2 Mandate + \gamma_3 MaskE \times Mandate + \phi X_i + \varepsilon_i)} \quad (6)$$

Using the same set of covariates as in specification (2) 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 whether to wear a mask independently of whether the experimenter is wearing a mask or not.

Another possible concern is that the presence of a masked experimenter makes subjects refrain from joining the line behind the experimenter. As argued in subsection 5, the data speaks against this being the case. We do not find evidence either for subjects' behavior being influenced by the experimenter putting on or removing the mask when they changed from one treatment to the other, which occurred once in a session. To test this, we add two variables to specification (3) in Table 3: A dummy variable that identifies these observations and an interaction term between this dummy and *MaskE*. We find that neither the dummy (p=0.061), giving the average change in distances when changing from MASK to NOMASK, nor the interaction term (p=0.772), giving the change in distances for subjects when switching to MASK, are significant.

In addition to the above-mentioned results, Table 3 reveals further interesting patterns. Population

---

[26]We acknowledge that this finding may eventually reverse if mask prevalence further increases. As masks became mandatory even in waiting lines, we would expect that only a few contrarians refused to wear masks in that setting, likely those who would also be less willing to adhere to the distancing recommendation.

density of the neighborhood decreases distancing; for an additional 1,000 inhabitants/km^2 average distances decrease by more than 1cm.[27] An explanation is that in areas with a large population density, more people will be outside at the same time, making distancing more difficult. Consistent with this idea, the effect of population density is smaller when we also include open stores, which also influence how many people are walking outside in the area where we are collecting data. Further, we note that 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.[28]

5.3   Compliance

German health authorities and official mandates to limit the spread of the coronavirus specify that individuals should keep a distance of at least 150cm to each other. In addition to our main analysis, we investigate how our treatment of masking the experimenter and the introduction of the indoor mask mandate in Berlin affect compliance with this required minimum distance. Instead of using mean distancing, we define compliance as a binary variable defined by whether the distance kept matches or exceeds the threshold of 150cm.

We find that compliance is higher toward the masked than toward the unmasked 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 they do. After the mandate is introduced, compliance is 40% if the experimenter does not wear a mask and 49.17% if they do.

The 150 cm rule may look arbitrary as the recommendations of safe distances vary substantially between countries.[29] Hence, we also consider whether compliance with alternative threshold values increases with masking the experimenter.

Figure 3 demonstrates that the choice of the critical value (within a relevant range) does not change our conclusion that a mask increases distancing from the experimenter. Three graphs on the left-hand side showcase the cumulative density function of *Distance* in both periods separately as well as in the combined sample. It is evident that subjects in the MASK condition are more likely to exceed any relevant threshold value, i.e., the distance effect not only changes the average, but compliance becomes generally higher in the MASK condition. Using non-parametrically estimated kernel density functions on the same sets of observations, we confirm a positive shift in distancing.[30]

5.4   External Validity

In a pandemic, a successful mitigation strategy needs to understand and take into account how policy affects behavior. The medical literature concludes that airborne contagion is a primary source of transmission of SARS-CoV-2. Hence, social distancing is important to prevent the spread

---

[27]Area is defined by postal code.

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

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

[30]Two-sided Kolmogorov–Smirnov tests yield $D=0.175$, $p=0.01$ in the combined sample. 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 ($\beta_1 =0.7145$, p=0.02). The interaction between *Mask Experimenter* and the policy change is insignificant as in the main model ($\beta_3 =-0.2893$, p=0.54). See the Appendix for details.
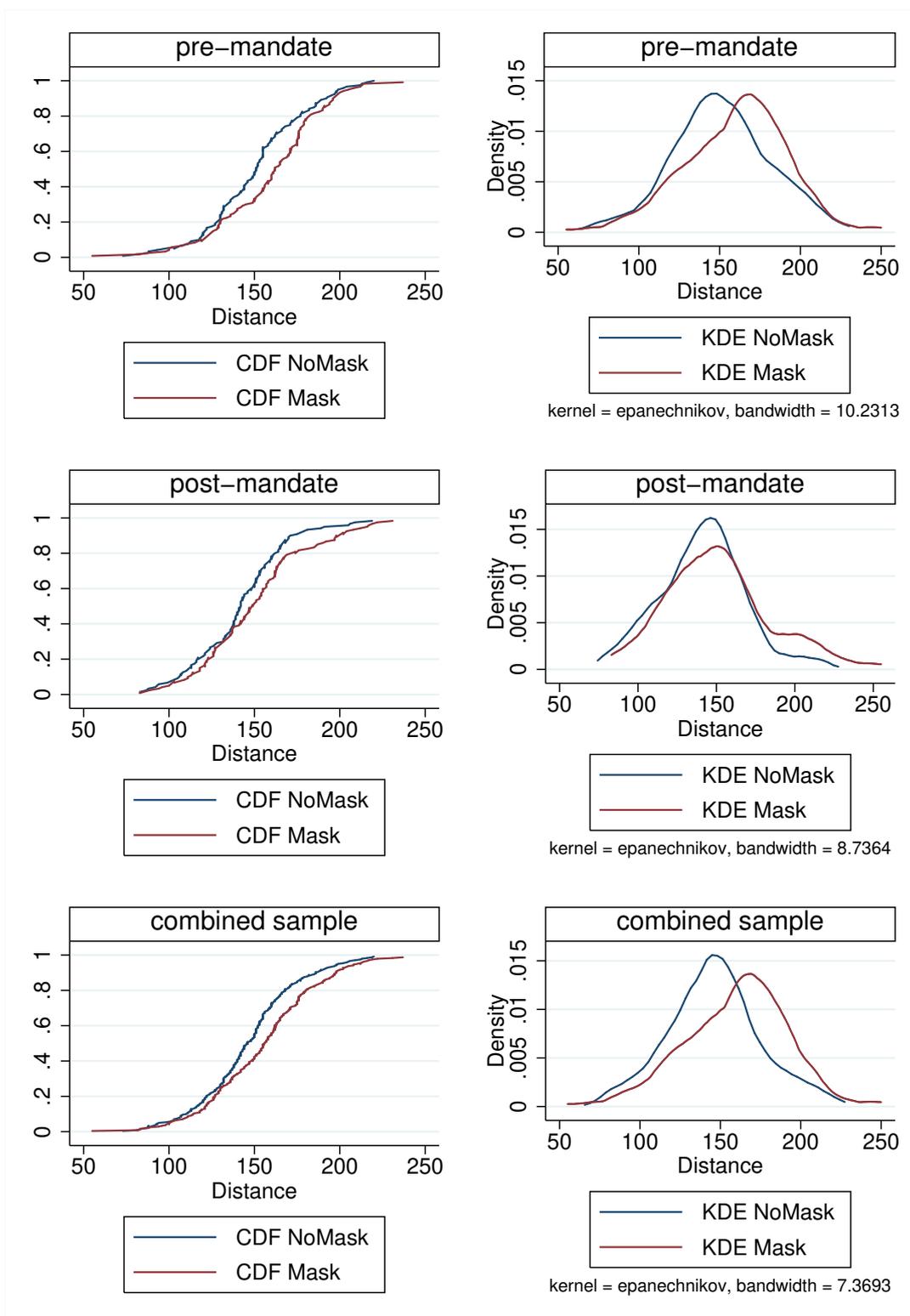
Figure 3: Cumulative distribution functions (CDF) of distances kept by the subject from the experimenter in NoMask (blue) and Mask (red) conditions (in centimeter), separately in the April (Policy= 0), the May (Policy= 1), and the full sample. Cumulative distributions are exact and densities are estimated univariate Epanechnikov kernel density estimation (KDE) functions.

of the virus (Anderson et al., 2020; Liu et al., 2020).[31]

In this study, we analyze how the introduction of a mask mandate in shops and public transport affected distancing in outside waiting lines. An overall evaluation of policies on distancing must take into account as many facets of individual behavior as possible. Restrictions and re-openings alter the choice set of customers, resulting in changes in behavior that affect the exposure to infectious particles. Thus, in order to put our results in perspective and discuss their external validity and relevance, we provide more details on the choice set available to subjects, the recommendation of public authorities, as well as publicly available knowledge about contagion risk at the time of the field experiment.

The data was collected in an environment where transmission is possible (Qian et al., 2021) even though unlikely if subjects adhere to the distancing guidelines.[32] At the same time, wearing a mask remained optional outdoors even during the mandate under study. Distancing, however, was recommended by authorities in the places where we collected data already at the time and the suggested minimum distance was 150cm. In addition to the recommended distancing, local authorities advocated diligent hand hygiene and respiratory etiquette to slow the spread of the virus.[33]

The success of social distancing is measured in the number and psychical distance of close encounters between people. Here, we focus on closeness. The effect of the mask mandate on behavior might be different in stores or high-risk areas where the mandate made wearing a mask indeed mandatory. However, our results are fully in line with evidence from mobility patterns in Germany, which have not changed negatively with the introduction of the mask mandate (Kovacs et al., 2020).

Our study provides results on behavioral effects in individuals distancing from a generic and frequent outdoor setting where contagion is possible and considered relevant by the authorities. Our main results cast doubt on the idea that individuals generally engage in risk compensating behavior. In contrast, we find no evidence that the distances people keep from others are shorter when the other person is masked rather than unmasked or when a general mask mandate is in place as compared to an unregulated setting. We interpret this as evidence that the risk compensation theory and its predictions, which proved helpful in understanding driver behavior and traffic deaths, does not necessarily generalize. In particular, it cannot simply be applied to the setting of a viral pandemic.

In terms of policy interventions, this means that preventive measures need not thwart one another. In contrast, some measures may even trigger complementary efforts (e.g., through signaling or reminder effects) instead of lulling people at risk into a false sense of security. This insight notwithstanding, the field experiment also shows that compliance with policies is highly context-dependent and to design the most effective policies we need to better understand the effects of context. We believe that randomized controlled trials are an extremely helpful tool to gather the necessary data and to evaluate the impact of new or potential alternative measures.

---

[31]In 2020, some jurisdictions imposed a mask mandate even in outdoor spaces, e.g. Berlin introduced it in some areas on October 20, 2020 (Berlin Senate, 2020).

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

[33]At the time of recording the data, the possibility of outdoor transmission was already known (Qian et al., 2021), but estimates on the risk of airborne infection were unavailable (Robert Koch-Institute, 2020). This was acknowledged as an infection channel by the World Health Organization only in July 2020 (WHO, 2020b).

6   Discussion and conclusion

This study utilizes a repeated field experiment on how mask-wearing affects distancing to assess how the introduction of an indoor mask mandate further affected distancing. The experimenters positioned themselves at the end of outdoor waiting lines during the COVID-19 pandemic, varying systematically whether or not they would wear a face mask. Data was collected before and after masking became mandatory inside stores. Our study design allows us to investigate whether risk compensation behavior occurred in response to two exogenous safety conditions: an indoor mask mandate and someone else wearing a mask. This question is highly relevant as risk compensation, ironically, could thwart the goals of a mask mandate by making the public reduce their efforts in social distancing, which is considered one of the most important measures against the spread of the virus. Using this reasoning, the World Health Organization and other health bodies initially did not recommend the community use of face masks as a preventive measure.

Our results speak against such a direct backlash against the use of face masks but also suggest that distancing is sensitive to contextual changes such as increased shop openings. The main result of this study is that the effect of masks worn by the experimenters on distancing is significantly positive and does not differ significantly between the pre- and post-mandate phases. However, behavior seems to react to policy. Our analyses suggest that individuals became less careful in their distancing behavior as relaxation measures became effective. In particular, average distancing declined in areas that witnessed store re-openings in the post-policy sample, possibly making it harder to keep safe distancing because sidewalks became more crowded.

More specifically, we do not find evidence of risk compensation in any of three cases where it might be expected: a) Average distances kept from the experimenter are not shorter post-mandate than pre-mandate; b) distances kept from a masked experimenter were not shorter (but larger) than distances kept from an unmasked experimenter; and c) masked subjects did not maintain a shorter distance than unmasked subjects. We argue that these results, which are inconsistent with risk compensation, may be driven by the fact that risk compensation, theoretically, occurs only under assumptions that are invalid in the pandemic context: specifically, masks may serve as cues, reducing instead of increasing the effort cost of physical distancing.

While we do not find evidence of risk compensation, we do not find convincing evidence that masks act as a trigger for maximally cautious behavior either. If masks triggered maximal caution, we would expect that the treatment becomes ineffective after the mandate and that the mandate increases distancing if the experimenter is unmasked, thereby also pushing up average distancing. To the contrary, not only does the treatment effect of the experimenter being masked persist before and after the introduction of a mask mandate, we also find no evidence that the mask mandate positively affected distancing. Rather, distancing decreases post-mandate even if we control for confounding factors like the partial revocation of shop closures. In addition, survey evidence from the first phase of data collection speaks against a trigger mechanism. Using that data, Seres et al. (2021) argue that individuals who wear a mask signal to others that they want them to keep a large enough distance, which, in turn, increases distancing. This reasoning, even though not tested for directly and possibly weaker, is likely still at work post-mandate.

Countervailing effects may influence the average mask effect. Hence, we must exercise caution in how we interpret it, depending on who is wearing the mask. In particular, although the ex-

perimenter randomizes mask use, masking by subjects is endogenous and may be influenced by risk preferences or health status. If these factors are associated with their beliefs about the experimenter, these can trigger a behavioral response in either direction. While we cannot rule out this possibility, the online survey experiment by (Seres et al., 2021) reveals that, on average, face masks do not change beliefs about the health status of the mask wearer, suggesting that the average change in distancing toward a masked person is unlikely to be driven by beliefs about him or her being more likely infectious or sick.[34]

Since the introduction of the first mandate in April 2020, the situation has changed multiple times. It remains unknown how the changes and the long-term experience with face masks affect behavior, thus understanding the underlying mental models would help predict how individuals react to changes in the environment. For instance, subjects in our field experiment may have taken policy relaxations in May 2020 to imply that the severity of the situation had decreased. Such a decrease in perceived risk would, according to the simple risk compensation model laid out above or any other model with an optimizing individual, lead to more risk taking. However, in principle, a mask mandate could also be taken as a signal that the situation is severe, making such a mandate necessary to bring down the level of transmission (Seegert et al., 2020). While we can say little as to the absolute strength of these two effects, the relaxations seem to have outweighed the signaling effect in April/May 2020 in Berlin. Assuming that the mask mandate is of informational value, we would expect that distancing would have decreased further without the mandate. Further, the persistent positive treatment effect from the mask suggests that individuals remain aware of the mask as a signal of individual preferences and adjusted their distancing toward it, which – due to the increased prevalence of masked individuals – also counteracted the tendency to shorter distances.

An interesting question for further research would be whether and to what extent the signaling effect of masks persists in places where masking is mandatory and whether political communication can affect the signaling value of policies like a mask mandate. These questions go beyond this study, which we see as a first step in understanding the behavioral reactions to mask wearing and mask mandates. To better understand the behavioral effects of various health interventions and non-pharmaceutical interventions to curb the pandemic, we encourage further randomized controlled trials that are better able to capture causal effects and less speculative in nature.

Another interesting avenue to follow up on this work would be to study heterogeneity. As discussed in Section 4, individual behavior does not depend directly on exogenous risk and safety but individual decisions are driven by the individual's perceptions of those variables. In particular, individuals' reactions to mask-wearing will be driven by their perception of the effect of masks which may differ from the actual effect of masking. In line with our modeling approach, many people seem to believe that masks mostly reduce their probability of infection (and the probability of infecting others) but a fraction of individuals also believe that mask-wearing reduces the severity of the disease when infected (Betsch et al., 2020c). The model can be enriched to allow for these beliefs without qualitatively changing its predictions. But, according to the model, heterogeneity in beliefs about the effects of mask-wearing (and the effects of other policies), can explain why individuals react differently to mask-wearing and a policy. No data on beliefs is available to follow up on this idea within our study.

---

[34]Note that this online survey used pictures of the same experimenters who carried out this field experiment.

References

Abaluck, J., L. H. Kwong, A. Styczynski, A. Haque, M. A. Kabir, E. Bates-Jefferys, E. Crawford, J. Benjamin-Chung, S. Benhachmi, S. Raihan, et al. (2021). Normalizing community mask-wearing: A cluster randomized trial in bangladesh. Technical report, National Bureau of Economic Research.

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. Yang (2020). Polarization and public health: Partisan differences in social distancing during the coronavirus pandemic. *Journal of Public Economics 191*, 104254.

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.

Bakhit, M., N. Krzyzaniak, A. M. Scott, J. Clark, P. Glasziou, and C. Del Mar (2021). Downsides of face masks and possible mitigation strategies: a systematic review and meta-analysis. *BMJ open 11*(2), e044364.

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

Berlin LAGeSo (2021). COVID-19 in Berlin, Fallzahlen und Indikatoren - Gesamtübersicht . https://daten.berlin.de/datensaetze/covid-19-berlin-fallzahlen-und-indikatoren-gesamtübersicht-0 (accessed: August 27, 2021).

Berlin Senate (2020). Maskenpflicht fur belebte Strasen, Markte und Shoppingmalls . https://www.berlin.de/aktuelles/berlin/6330152-958092-maskenpflicht-fuer-belebte-strassen-maer.html.

Betsch, C., L. Korn, L. Felgendreff, S. Eitze, P. Schmid, P. Sprengholz, L. Wieler, P. Schmich, V. Stollorz, M. Ramharter, et al. (2020a). German COVID-19 Snapshot Monitoring (COSMO)-Welle 15 (15.05. 2020).

Betsch, C., L. Korn, L. Felgendreff, S. Eitze, P. Schmid, P. Sprengholz, L. Wieler, P. Schmich, V. Stollorz, M. Ramharter, et al. (2020b). German COVID-19 Snapshot Monitoring (COSMO)-Welle 15 (18.04. 2020).

Betsch, C., L. Korn, L. Felgendreff, S. Eitze, P. Schmid, P. Sprengholz, L. Wieler, P. Schmich, V. Stollorz, M. Ramharter, et al. (2020c). Masketragen: 4.3 Wissen über die Effekte des Masketragens (Stand: 30.09.20). https://projekte.uni-erfurt.de/cosmo2020/web/topic/wissen-verhalten/40-masketragen/#wissen-über-die-effekte-des-masketragens-stand-30.09.20.

Blomquist, G. (1986). A utility maximization model of driver traffic safety behavior. *Accident Analysis & Prevention 18*(5), 371–375.

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.

Chernozhukov, V., H. Kasahara, and P. Schrimpf (2021). Causal impact of masks, policies, behavior on early COVID-19 pandemic in the US. *Journal of Econometrics 220*(1), 23–62.

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.

Courtney, J. M. and A. Bax (2021). Hydrating the respiratory tract: An alternative explanation why masks lower severity of covid-19. *Biophysical Journal 120*(6), 994–1000.

CSSE (2021). COVID-19 Dashboard by the Center for Systems Science and Engineering (CSSE) at Johns Hopkins University (JHU).

DIVI (2021). DIVI Intensivregister – Gesamtzahl gemeldeter Intensivbetten (Betreibbare Betten und Notfallreserve), Berlin (Stand: 1. September 2021) . https://www.intensivregister.de/#/aktuelle-lage/zeitreihen.

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 5*, 293–308.

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.

Gandhi, M. and G. W. Rutherford (2020). Facial masking for covid-19—potential for "variolation" as we await a vaccine. *New England Journal of Medicine 383*(18), e101.

Grossman, G., S. Kim, J. M. Rexer, and H. Thirumurthy (2020). Political partisanship influences behavioral responses to governors' recommendations for COVID-19 prevention in the United States. *Proceedings of the National Academy of Sciences 117*(39), 24144–24153.

Guenther, B., M. M. Galizzi, and J. G. Sanders (2021). Heterogeneity in risk-taking during the COVID-19 pandemic: evidence from the UK lockdown. *Frontiers in psychology 12*, 852.

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*, 1–14.

He, X., E. H. Lau, P. Wu, X. Deng, J. Wang, X. Hao, Y. C. Lau, J. Y. Wong, Y. Guan, X. Tan, et al. (2020). Temporal dynamics in viral shedding and transmissibility of COVID-19. *Nature Medicine 26*(5), 672–675.

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.

Howard, J., A. Huang, Z. Li, Z. Tufekci, V. Zdimal, H.-M. van der Westhuizen, A. von Delft, A. Price, L. Fridman, L.-H. Tang, et al. (2021). An evidence review of face masks against COVID-19. *Proceedings of the National Academy of Sciences 118*(4).

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., J. Nyrup, and M. B. Petersen (2020). Do survey estimates of the public's compliance with covid-19 regulations suffer from social desirability bias? *Journal of Behavioral Public Administration 3*(2).

Liebst, L. S., P. Ejbye-Ernst, M. Bruin, J. Thomas, and M. R. Lindegaard (2021). Mask-wearing and social distancing: Evidence from a video-observational and natural-experimental study of public space behavior during the COVID-19 pandemic.

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.

Liu, Y., L.-M. Yan, L. Wan, T.-X. Xiang, A. Le, J.-M. Liu, M. Peiris, L. L. Poon, and W. Zhang (2020). Viral dynamics in mild and severe cases of COVID-19. *The Lancet Infectious Diseases 20*(6), 656–657.

Mantzari, E., G. J. Rubin, and T. M. Marteau (2020). Is risk compensation threatening public health in the COVID-19 pandemic? *BMJ 370*.

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).

Marks, M., P. Millat-Martinez, D. Ouchi, C. h Roberts, A. Alemany, M. Corbacho-Monné, M. Ubals, A. Tobias, C. Tebé, E. Ballana, et al. (2021). Transmission of COVID-19 in 282 clusters in Catalonia, Spain: a cohort study. *The Lancet Infectious Diseases*.

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/.

O'Driscoll, M., G. R. Dos Santos, L. Wang, D. A. Cummings, A. S. Azman, J. Paireau, A. Fontanet, S. Cauchemez, and H. Salje (2021). Age-specific mortality and immunity patterns of SARS-CoV-2. *Nature 590*(7844), 140–145.

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

Pless, B. (2016). Risk compensation: Revisited and rebutted. *Safety 2*(3), 16.

Qian, H., T. Miao, L. Liu, X. Zheng, D. Luo, and Y. Li (2021). Indoor transmission of SARS-CoV-2. *Indoor Air 31*(3), 639–645.

Radun, I., J. Radun, M. Esmaeilikia, and T. Lajunen (2018). Risk compensation and bicycle helmets: A false conclusion and uncritical citations. *Transportation Research Part F: Traffic Psychology and Behaviour 58*, 548–555.

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.

Robert Koch-Institute (2020). SARS-CoV-2 Steckbrief zur Coronavirus-Krankheit-2019 (COVID-19). https://www.rki.de/DE/Content/InfAZ/N/.

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.

Seegert, N., M. Gaulin, M.-J. Yang, and F. Navarro-Sanchez (2020). Information revelation of decentralized crisis management: Evidence from natural experiments on mask mandates. *Available at SSRN 3736407*.

Seres, G., A. Balleyer, N. Cerutti, A. Danilov, J. Friedrichsen, Y. Liu, and M. Süer (2021). Face Masks Increase Compliance with Physical Distancing Recommendations During the COVID-19 Pandemic. *Journal of the Economic Science Association.* forthcoming.

Spinelli, M. A., D. V. Glidden, E. D. Gennatas, M. Bielecki, C. Beyrer, G. Rutherford, H. Chambers, E. Goosby, and M. Gandhi (2021). Importance of non-pharmaceutical interventions in lowering the viral inoculum to reduce susceptibility to infection by SARS-CoV-2 and potentially disease severity. *The Lancet Infectious Diseases.*

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.

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.

Tull, M. T., K. A. Edmonds, K. M. Scamaldo, J. R. Richmond, J. P. Rose, and K. L. Gratz (2020). Psychological outcomes associated with stay-at-home orders and the perceived impact of COVID-19 on daily life. *Psychiatry Research 289*, 113098.

Van Damme, W., R. Dahake, R. van de Pas, G. Vanham, and Y. Assefa (2021). COVID-19: Does the infectious inoculum dose-response relationship contribute to understanding heterogeneity in disease severity and transmission dynamics? *Medical Hypotheses 146*, 110431.

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 20*(6), 669–677.

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

WHO (2020b). Transmission of SARS-CoV-2: implications for infection prevention precautions . https://www.who.int/news-room/commentaries/detail/transmission-of-sars-cov-2-implications-for-infection-prevention-precautions.

Wilde, G. J. (1982). The theory of risk homeostasis: implications for safety and health. *Risk Analysis 2*(4), 209–225.

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 three appendices. The first appendix contains the pre-registered protocol for the field experiment. (Appendix A). The second consists of robustness tests (Appendix B). The third includes the modified risk compensation model (Appendix 4.1).

A   Experimental Protocol

[35] ***Disclaimer:*** *Experimenters signed up to this experiment on voluntary basis under the condition that they do not belong to any risk groups. To prevent posing any further risk on any of the parties, the Robert Koch Institute's health recommendations are strictly followed.*[36]

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.[37] 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

---

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

[36]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/.

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

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.

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.[38] 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.

---

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

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.

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   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 5.2, 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 × Policy	-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.
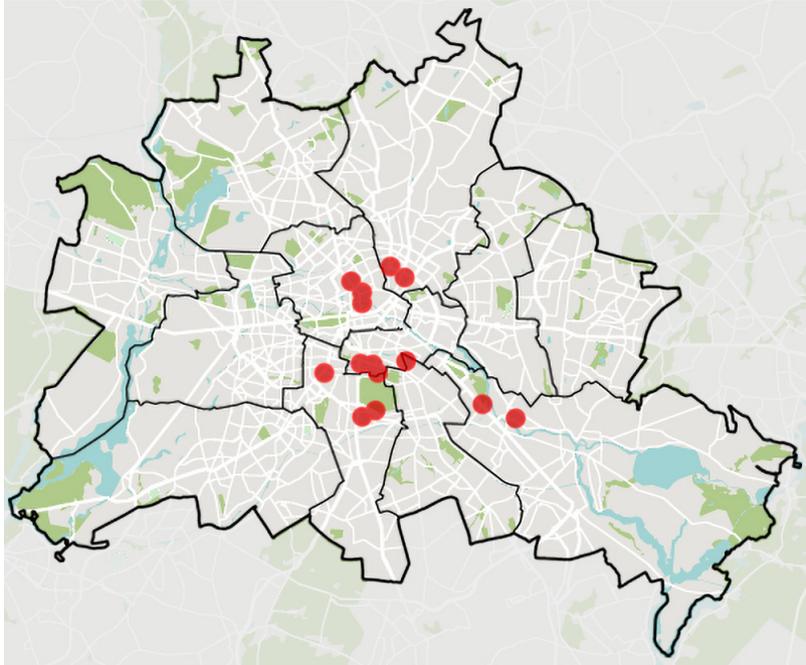
C Map Of Observation Sites

Figure 4: Map indicating the observation sites in Berlin, Germany. The lines indicate district borders within city limits.