

Unmasking the truth: experimental evidence of facemask compliance in Bangladesh, Kenya, and Nigeria during the COVID-19 pandemic

Karen Grépin (✉ kgrepin@hku.hk)

University of Hong Kong <https://orcid.org/0000-0003-4368-0045>

Valerie Mueller

Arizona State University <https://orcid.org/0000-0003-1246-2141>

Nicole Wu

University of Hong Kong

Atonu Rabbani

University of Dhaka

Article

Keywords: COVID-19, low and middle-income countries, facemask mandates

Posted Date: November 30th, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-1073440/v1>

License:   This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Unmasking the truth: experimental evidence of facemask compliance in Bangladesh, Kenya, and Nigeria during the COVID-19 pandemic

Abstract:

High levels of compliance with public health measures are critical to ensuring a successful response to the COVID-19 pandemic, especially in low and middle-income countries (LMICs). However, most data on compliance are self-reported. Tendency to overreport due to social desirability can yield biased estimates of compliance. We estimate rates of compliance with facemask mandates in Kenya, Nigeria, and Bangladesh using data from phone surveys conducted in March-April 2021. Data on compliance are collected from different survey modules: self-reported compliance (stated) and a list experiment (elicited). We find substantial gaps between stated and elicited rates of facemask wearing for different groups depending on specific country contexts and high levels of overreporting of facemask compliance in self-reported surveys. We observe differences in rates of self-reported facemask wearing among key groups but not using the elicited responses from the list experiment, which suggest that social desirability bias may vary by demographics. Data collected from self-reported surveys may not be reliable to monitor ongoing compliance with public health measures. Moreover, elicited compliance rates indicate levels of mask wearing are likely much lower than those estimated using self-reported data.

1 Introduction

High levels of compliance with public health measures are critical for a successful response to COVID-19.¹ Since the start of the pandemic, most countries have implemented measures to mitigate the spread and impact of the virus, many of which require substantial behavior change – an important public health challenge. Adding to this challenge, the prolonged nature of the pandemic will require that countries maintain high levels of compliance with these measures over extended periods, especially in low- and middle-income countries (LMICs) where limited access to vaccines will extend the effects of the pandemic in these settings.

Facemasks are among the most effective, low-cost measures that can be adopted to reduce the transmission of COVID-19. In June 2020, the World Health Organization (WHO) issued guidance to recommend their widespread use to prevent transmission² despite the lack of rigorous evidence of their effectiveness among the public at the time. Since then, a randomized study conducted in Bangladesh has confirmed that facemasks can reduce the number of symptomatic COVID-19 cases³ at the population level. As facemask wearing is relatively harmless, some have also argued that it should be adopted based on the precautionary

principle.⁴ In low and middle-income countries (LMICs), the mandated use of facemasks has also been argued as an alternative to much costlier control strategies.⁵ Due to these factors, as well as growing evidence supporting their effectiveness and the apparent success of jurisdictions with universal masking mandates, many countries adopted facemask mandates or recommended their use during the first few months of the pandemic, including in LMICs.^{6–10}

To be effective, however, facemasks need to be worn and worn properly. While studies generally find high levels of compliance with facemask mandates, given that it is socially desirable for people to wear a facemask, it is not clear if self-reported data are reliable. This may be especially true in LMICs where resources to enforce mask mandates may be more limited. While studies in many LMICs have generally found high rates of self-reported compliance,⁵ one study in Kenya found that self-reported compliance estimates were much higher than estimates made from direct observation in public spaces over the same period.¹¹ Achieving high rates of compliance is critical to reducing transmission, however, accurate measurement of actual compliance with public health directives is also essential to monitor policy effectiveness. In Kenya, Nigeria, and Bangladesh, the setting for this experiment, surveys conducted during the pandemic have generally found facemask compliance to be 80%-90% and public awareness of the importance of wearing a facemask to be nearly universal.^{12–14}

List experiments, also known as item count surveys, have frequently been used to overcome social desirability reporting bias in surveys.¹⁵ In such experiments, respondents are shown a list of behaviors with a potentially sensitive behavior randomly included or excluded and are then asked to report how many of the behaviors, but not which ones, they have engaged in over a fixed recall period. As it is not possible to directly infer which of the behaviors the respondent has engaged in, it is believed that respondents have less incentive to lie about sensitive behaviors. The difference-in-means estimator can then be used to generate unbiased estimates of the rates of the sensitive behaviors for the sample. List experiments have been used to measure intimate partner violence, condom use, and abortions.^{16–19} They have also been used to study compliance with some public health measures in high-income countries during the pandemic, for example, to measure social distancing in a set of high-income countries and handwashing in Ireland.^{20,21} But to our knowledge, they have not been used to estimate facemask wearing compliance during the pandemic nor in a LMIC setting.

In this paper, we use a list experiment to estimate actual compliance with facemask mandates among people living in three LMICs during the COVID-19 pandemic. At the time of our survey, all three countries had

mandatory facemask policies in place: in Kenya, the Ministry of Health mandated masks in public places starting in early April 2020,²² the Nigerian President mandated everyone to wear a mask in public starting in late April 2020, although similar policies had already been adopted in most states before this policy change,²³ and in late May a similar policy had been put into place in Bangladesh.³

2 Methodology

2.1 Data

Data for this paper come from longitudinal phone surveys that were conducted in Bangladesh, Kenya, and Nigeria during the COVID-19 pandemic. The surveys had been designed to study the gendered effects of the pandemic on a broad set of health, economic, and social outcomes. The first round of the survey was conducted between October-December 2020 in all three countries and the second round was conducted between March-April 2021. While the list experiment was only included in the second round of the survey, we use data on respondents that were collected during both survey rounds.

To construct our samples, in Kenya and Nigeria, we used a random digit dial (RDD) sampling technique. In each country, registered mobile phone numbers were obtained from a third-party vendor, which were then randomly called by the enumerators to recruit participants. In each country, we aimed to recruit approximately 2000 respondents in the first round and targeted to obtain at least 60% of the sample to be women. In Kenya, we used a referral experiment to oversample women, while in Nigeria we used quota sampling to increase the geographic, age, and gender representativeness of our samples.²⁴ More details of our sample are available in the following paper.²⁵ In Bangladesh, instead of RDD, we drew upon a recently completed household survey to construct the sample.²⁶ Individuals were randomly selected from the original survey, initially targeting a sample size of 1800. We oversampled women to ensure a gender balance. Innovations for Poverty Action (IPA) conducted the surveys in Kenya and Nigeria, while the BRAC James P Grant School of Public Health implemented the survey in Bangladesh. We used sampling weights to adjust for selection bias inherent in the RDD sampling approach, non-response, and attrition over the survey rounds in all countries (Supplementary Material).

In all countries, the inclusion criteria were that respondents, either men or women, must be at least 18 years of age, must be able to complete the survey in one of the commonly spoken languages in each country (2 in Kenya, 5 in Nigeria, and 2 in Bangladesh), must be willing to be contacted again in the future, and must be willing to provide contact information and a first name for future identification. In our first round, we randomly surveyed 1822 individuals in Bangladesh (914 men and 908 women), 2038 individuals in Kenya

(742 men and 1296 women), and 1969 individuals in Nigeria (823 men and 1146 women). In our second round, we completed follow-up surveys with 1722 people in Bangladesh (94.51% rate), 1647 people in Kenya (80.81%), and 1613 people in Nigeria (81.92%).

2.2 *Measuring stated facemask compliance*

We asked respondents about compliance with a list of public health measures including facemask wearing. Specifically, we asked respondents to report how often in the past 7 days they wore a mask when in public and were given the choices of (1) all of the time, (2) most of the time, (3) about half of the time, (4) some of the time, (5) none of the time, and (6) not applicable, I have not been out in public in the past 7 days. The recall period was selected to be directly comparable to the list experiment. We dichotomized the responses by defining compliance with facemask mandates if the respondent reported wearing a facemask all the time (=1) or most of the time (=2). Responses were defined as not being compliant otherwise. In the Supplementary materials, we also examine the results by more conservatively defining facemask compliance by coding “yes” only when “all of the time” is reported. This leads to a considerable reduction in gaps between stated and elicited non-compliance with wearing masks (see Supplementary Material Figure A1).

2.3 *Eliciting facemask compliance using the list experiment*

In our experimental design, in the second round of the survey, male and female respondents were randomly assigned to be in either the treatment or the control arm of the experiment with equal probability. In each group, respondents were told that they would be provided a list of activities, and they were asked to report on how many of the activities they had done over the past 7 days in aggregates (see Supplementary Material Table A1). The control group received a list of activities that included: (1) called a friend or family member, (2) listened to the radio, (3) drove in a car or motorcycle, and (4) sent someone mobile money. The treatment received the same list, however, the option of “didn’t wear a mask when in public”, was included in the list between options (3) and (4) of the control group.

For a list experiment to be valid, that is to generate unbiased estimates of the rates of a sensitive behaviour, three conditions need to be met. First, the randomization between the treatment and control groups must have been carried out effectively. Second, there should not be any design effects, which means that the inclusion of the treatment list does not affect the responses to the non-sensitive items. In the Supplementary Appendix we present the results of a test for design effects in Table A3. While there are not measurable design effects in either Kenya or Nigeria, there were minor effects detected for Bangladesh in one of the tests. Finally, it is assumed that the respondents do not lie in their responses, which is not itself testable.

2.4 Estimation strategy

We followed a standard list experiment design in which respondents from each country were randomly assigned to either the control or treatment group. The control group respondents were read a list of four non-sensitive yes (= 1) or no (= 0) questions (see Supplementary Materials Table A1 for more details) and respondents were expected to report on the total number of behaviours among those presented to them that they did over the past 7 days. The treatment group respondents received the same list of non-sensitive items but with an additional question about facemask wearing inserted. We observed the total number of behaviours $Y_{ij}(T)$, the respondents agreed they had engaged with in the past week. Here, $T = 1$, if the respondents belonged to the treatment group, or $T = 0$, otherwise. This suggests $Y_i(T) = \sum_{j=1}^{J(T_i)} X_{ij}$, where, $J(0) = 4$, and $J(1) = 5$. Also, $X_{ij} = 1$, if the i -th respondent agrees to j -th question, = 0, otherwise. Hence, the sum over j gives us the total number of questions the respondent i agrees with. We can estimate the fraction of the population agreeing to the sensitive item using the following difference-in-means estimator:

	$\hat{\delta} = \frac{1}{N_1} \sum_{i=1}^{N_1} T_i Y_i(T_i) - \frac{1}{N_0} \sum_{i=1}^{N_0} (1 - T_i) Y_i(T_i)$	(1)
--	--	-----

Here, N_T is the size of the control (if $T = 0$) and treatment (if $T = 1$) group. We can also estimate the $\hat{\delta}$, using the following equation:

	$Y_i = \alpha + \delta T_i + \varepsilon_i$	(2)
--	---	-----

Moreover, we can use multivariable regressions to understand how responses to the sensitive items vary by respondents' characteristics by interacting the treatment assignment (i.e., T_i) with the individual characteristics (e.g., Z_i) by estimating the following equation:

	$Y_i = Z_i \gamma + T_i \cdot Z_i \delta + \varepsilon_i$	(3)
--	---	-----

Here, again, we are primarily interested in the vector δ . We can again estimate the parameters (γ, δ) using OLS models.

These estimators are valid if the list experiments comply with three desirable properties. The first one requires valid randomization or a balance between the treatment and control groups. In Supplementary Material Table A2, we present the results of a balance test to evaluate the effectiveness of the randomization in our sample. In all three countries, we observe a good balance between our treatment and control groups

on all variables used in our analysis. Although women were overrepresented in the overall sample, we observe no difference in the proportion who are assigned to either the treatment or control groups. The second one requires no design effects. We conduct tests for the presence of design affects in Appendix Table A3 and show that there were no issues for Kenya and Nigeria, but in Bangladesh we detected design effects in one of the 10 cases tested. The third one requires no liars or the respondent not changing their reports in presence of the sensitive items. This also calls for checking floor and ceiling effects.¹⁵ Our estimates are not sensitive to the inclusion of floor and ceiling effects in the estimation procedures.

3 Findings

3.1 Summary statistics

In Table 1, we provide the summary statistics for our experimental sample, weighted using the survey sampling weights. The final samples included 1647 respondents in Kenya, 1613 in Nigeria, and 1722 in Bangladesh. In all three countries, the samples were slightly more female (63% in Kenya, 58% in Nigeria, and 55% in Bangladesh). All data are weighted using the sampling weights, which are described in the Supplementary Materials. The weighted average age of the respondents was 32 in Kenya, 31 in Nigeria, and 40 in Bangladesh. About half of the respondents were married in Kenya (52%) and Nigeria (49%) but over 80% of respondents were married in Bangladesh (83%). Household size was largest in Bangladesh (5.0), followed by Nigeria (4.1), then Kenya (3.3). These differences were partially driven by the fact that there were more children per household in Bangladesh than elsewhere although most of the households in all three countries had children. The sample was slightly more urban in Nigeria (65%) than in the other countries (50% in Kenya and 46% in Bangladesh). Some of the differences between Bangladesh and the other two countries may partially be explained by the RDD sampling strategy used in Kenya and Nigeria.

More people had known people with COVID-19 in Kenya (30%) relative to the other countries (8% in Nigeria and 18% in Bangladesh) and Kenyan respondents were also more likely to report feeling vulnerable to COVID-19 (55% in Kenya but less than 20% in the other countries). Data on education was captured differently in each of the countries, reflecting differences in the national educational systems. The Bangladesh sample was overall less educated with a higher proportion of the sample completing less than secondary school than in the other two countries. We also qualitatively assessed time preference or forward lookingness using an 11-point Likert scale by answering how willing the respondent was to give up something to get benefits in the future²⁷. We find comparable scores from Kenya and Nigeria (about 8 in both) with a significant lower score for Bangladesh (3.2).

In Figure 1, we describe the self-reported data on compliance with mask-wearing. In all three countries, there is a high level of self-reported use of mask-wearing and the most common response to this question was to wear a mask “all of the time”. In Kenya, almost 90% of the sample reported using a mask all or most of the time compared to 73% in Nigeria and 66% in Bangladesh. Only 1% of respondents in Kenya, 5% in Nigeria, and 7% in Bangladesh reported wearing a “mask none of the time”.

3.2 *Comparing stated and elicited facemask compliance*

In Figure 2, we report the total counts or reported behaviors for the control (panel 2a) and treatment (2b) groups in each country. Although we cannot directly test for the presence of any ceiling or floor effects, we can analyze the distribution of responses to our list experiment questions. In both Kenya and Nigeria, we can see that the distribution to our answers was well distributed with many people answering in the 2-4 answers range and almost no one responded that they did not engage in any of the behaviors. However, in Bangladesh, between 30-40% of the sample in both groups reported having not done any of the behaviors in our sample (see previous discussion of our analysis of the design effects).

In Figure 3, we visually compare stated and elicited estimates of non-compliance with facemask mandates and see a notable difference in all three countries. In Kenya, while only 14% of the sample stated being in public without a facemask over the past week using the data from the self-reported module, 52% of the sample reported not wearing a mask out in public during the past 7 days using the estimates elicited from the list experiment. While not as large as it is in Kenya, important differences were also observed in the other two countries (about 30 and 24 percentage point differences in Nigeria and Bangladesh, respectively).

3.3 *Multivariable analyses*

In Table 2, we provide estimates of differences for subgroups in our sample, namely male vs female, urban vs urban, and other groups. The table reports the interaction terms between the coefficients listed and an indicator of whether the respondent was in the list experiment (see equation 3). In general, based on the practices elicited through the list experiment, we do not find major differences between groups with regards to their levels of compliance in any of the three countries. In Bangladesh, we find those who have autonomy over their mask-wearing, are also less likely to report noncompliance. We also find an education gradient, more educated respondents are more likely to be compliance with facemask mandates in Kenya and Bangladesh, with a statistically significant coefficient in the latter. While there are between-country differences in facemask compliance, the elicited practices generally do not systematically vary within countries.

However, there is more systematic within-country variation in stated facemask wearing when respondents are asked to self-report compliance (Table 3). In Kenya and Nigeria, female respondents are more likely to report wearing masks relative to males, so are urban respondents and respondents who knew someone with COVID-19 in Kenya and Bangladesh. In Bangladesh, respondents who are more forward-looking and more educated were also less likely to report not wearing a facemask when in public. The lack of measurable variation in elicited responses via the list experiment but measurable differences in variation in self-reported compliance suggests that social-desirability bias is likely to affect some groups more than others.

In Table 4, we present the difference of the means of the stated vs elicited estimates of facemask compliance for different sub-groups of the sample. A positive difference is an estimate of the rate of overstatement of facemask wearing in each group. In general, while we find that most groups tend to overreport, there are not many groups that consistently overreport across all the countries. Married people tend to overreport as compared to unmarried respondents and more educated people may be less likely to overreport than less educated people, although the pattern does not strictly hold in Nigeria.

4 Discussion

High rates of compliance with public health measures are essential to mounting effective responses during infectious disease outbreaks, however, given social pressures associated with the adoption of socially desirable preventive behaviors it is challenging to measure compliance using self-reported data alone. This study demonstrates that there were large and meaningful differences in stated and elicited rates of mask-wearing comparing data from a self-reported module to data from a list experiment in Kenya, Nigeria, and Bangladesh implying social desirability bias likely plagues estimates of this behavior collected in self-reported surveys. Indeed, there was an almost 40 percentage point gap between stated and elicited non-compliance in Kenya. Although smaller in the other countries, the gap was still almost 30 percentage points in Nigeria and 20 percentage points in Bangladesh. In Bangladesh, experimentally induced differences in facemask compliance of approximately 30 percentage points were associated with a reduction in symptomatic seroprevalence of COVID-19 by almost 10%.³ Non-randomized studies have also generally found that high levels of compliance are needed to reduce transmission at the community level.^{28–30} Thus, effectiveness facemask mandates to reduce population-level transmission of COVID-19 requires both high levels of mask-wearing – and importantly the ability accurately measure actual compliance. Commonly used self-reported surveys are likely overestimating facemask compliance.

Studies conducted in high-income countries during the pandemic have also shown that estimates of public health compliance may also depend on how survey questions are framed and worded. For example, an

online survey conducted in Ireland found lower levels of reported handwashing when questions were worded negatively.²¹ A guilt-free survey strategy, in which instead of only asking yes or no questions, respondents were also given the option to report such behaviors ‘occasionally’ or ‘only when necessary’, increased estimates of non-compliance with preventive public health measures by 9-16 percentage points in 12 high-income countries and, separately, in Canada.^{31,32} It is therefore not surprising that we have also identified important survey design effects in our estimates of facemask compliance in lower-income countries. However, the relevance of these findings may be more important from a public policy perspective.

Our study has several limitations which should be taken into consideration when interpreting our findings. First, although our study was conducted similarly in all three countries, there were some important differences in the actual data collection processes across the three countries, including in how the samples were constructed and small differences with regards to the wording of some of the questions across countries. Second, there appear to have been some design effects in the Bangladesh sample due to lower reported rates of the non-controversial behaviors (i.e., floor effects). While it is not possible to determine how much of an effect it may have had in the interpretation of our results or the validity of the study, it does suggest we should interpret the results from Bangladesh more cautiously.

Although the data collected from the compliance and list experiment were not perfectly comparable (one asks about never being out in public without a face mask and the other asked about the frequency of facemask wearing using coarser categories) our findings suggest substantial differences likely do exist between self-reported and actual compliance with facemask wearing in LMIC. Notably, social desirability bias may vary according to gender, location of residence, and education levels. All three countries had facemask mandates in place for many months at the time of our survey, and despite very high rates of compliance reported in self-reported surveys conducted in these countries around the time of our survey, our findings suggest that actual compliance with facemask wearing was far from optimal and likely lower than that estimated in other studies. It is therefore important to understand which factors, above and beyond mandates, are important to help increase actual rates of compliance. This may be especially true in LMIC settings where resources to enforce mandates may be more limited. A recent large scale randomized trial implemented in Bangladesh to test strategies to increase mask usage found that the free provision alone of masks had only a small effect on uptake but that periodic monitoring in public places led to large increases in mask-wearing and that these effects were sustained over long periods.³³ The monitoring intervention also led to increases in other preventive behaviors, namely physical distancing, which further points to the importance of social norms and pro-social learning in the promotion of public health measures. It is

therefore also important to consider these social norms in our ability to measure these behaviors and our study demonstrates the need to use measurement strategies to overcome these pro-social biases in surveys. Simple methodological innovations such as a list experiment help to reduce the effect of social desirability in the measurement of such behaviors.

Data availability

The datasets generated and analysed in this study are available from the corresponding author on reasonable request.

Ethical clearance

The survey and human subject participation were reviewed locally for ethics clearance as well as by respective Institutional Review Board committees at The University XX and XX University (intentionally removed for double blind peer review). It also obtained ethical clearance from National Health Research Ethics Committee of Nigeria (NHREC) in Nigeria (NHREC/01/01/2007), XX IRB in Bangladesh (IRB-13 October'20-043), and Maseno University Ethics Review Committee in Kenya (MSU/DRPI/MUERC/00906/20).

Acknowledgments:

The research team would like to thank all members of XX research project (blinded for peer review) as well as the many research assistants who assisted with the collection and analysis of data on this project. We also would like to thank the respondents in Kenya, Nigeria, and Bangladesh for taking the time to complete the survey.

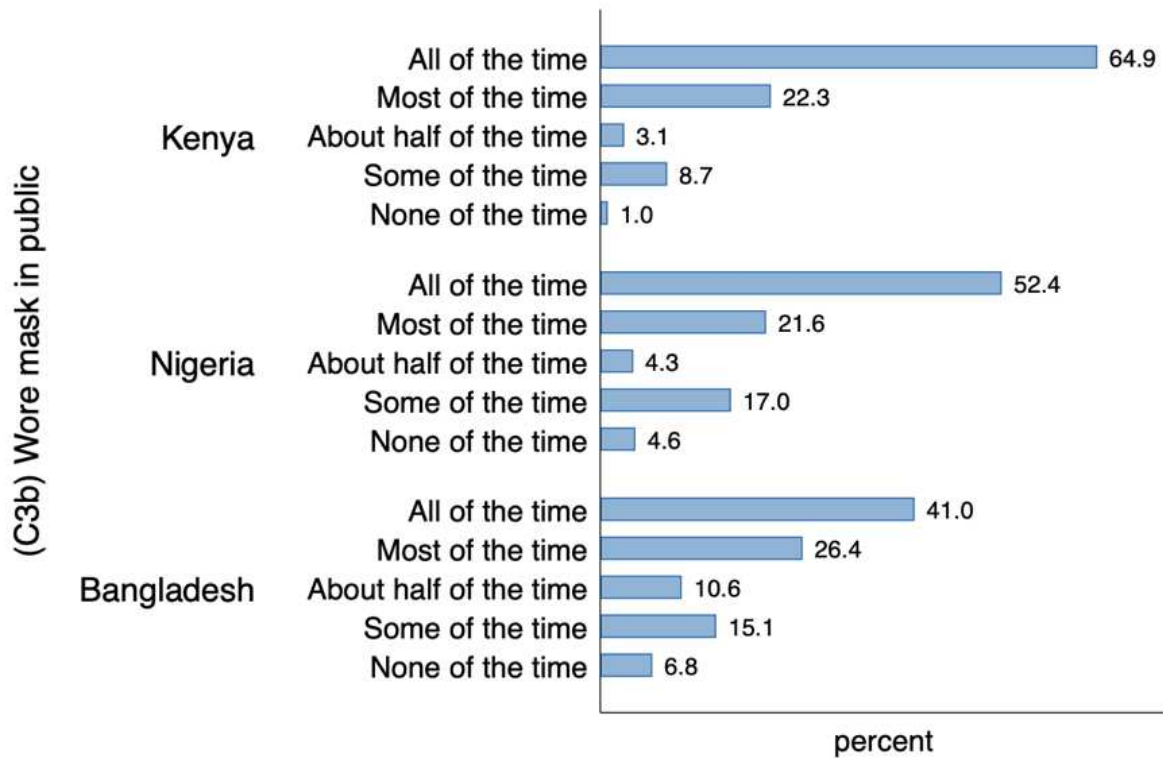
Author contributions:

Author1, Author2, and Author4 designed the study. Author4 and Author3 supervised the data collection. Author1 and Author4 led the analyses and drafting. All authors contributed to the revisions and finalization of the paper.

Competing interests:

The authors declare that they have no known competing interests.

Figure 1: Self-reported mask wearing



Notes: Findings from phones surveys carried out in Kenya (N = 1647), Nigeria (N = 1613), and Bangladesh (N = 1722). The respondents were asked “In the last 7 days, how often did you wear a mask when out in public?”. The tabulations reflect country specific weights.

Figure 2a: List Counts for Control

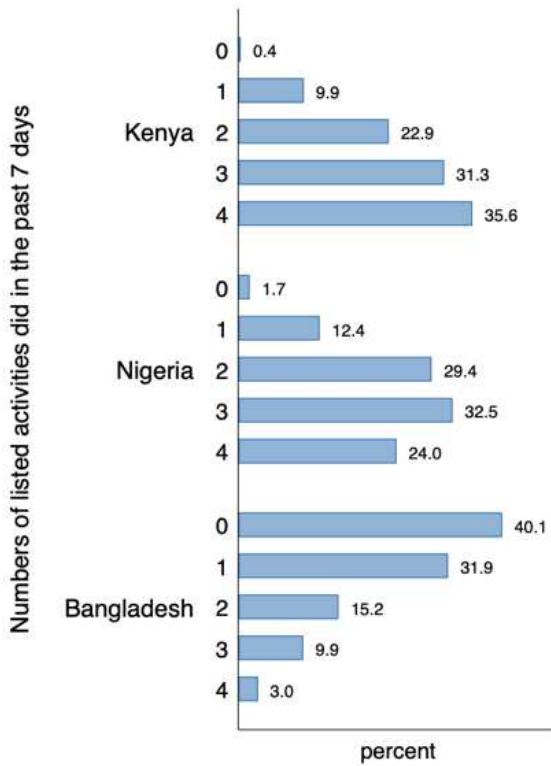
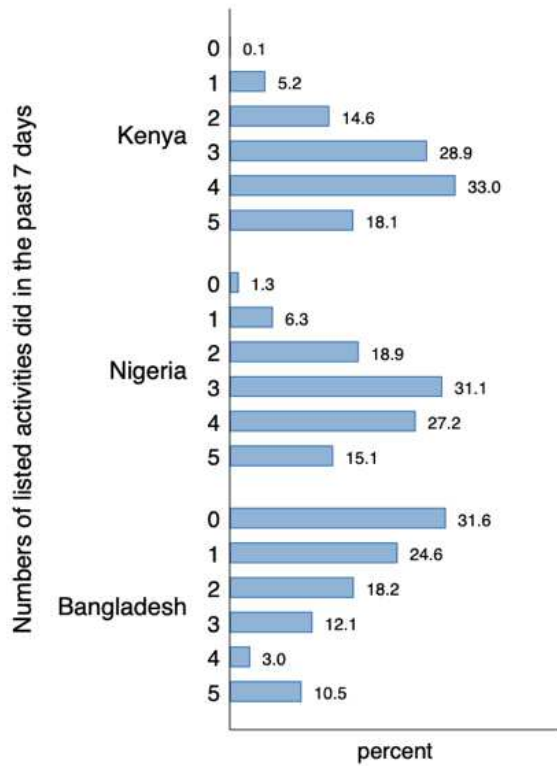
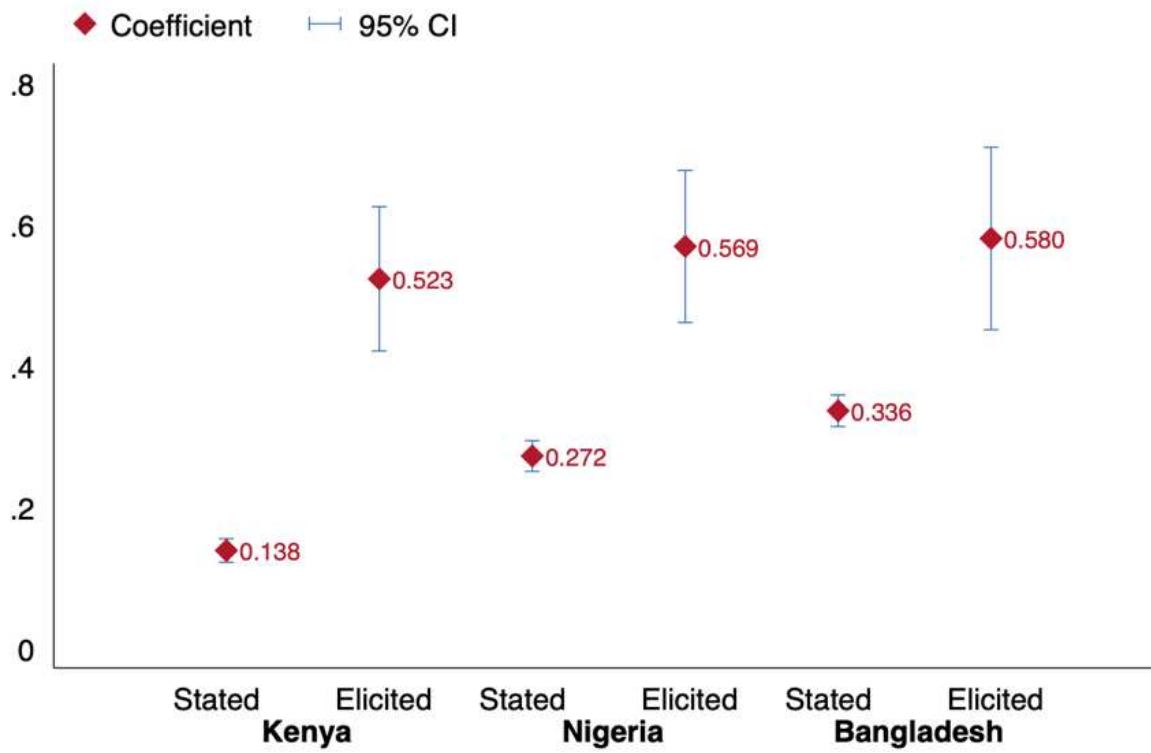


Figure 2b: List Counts for Treatment Groups



Notes: Authors' calculations from phone survey data from Kenya (N = 1647), Nigeria (N = 1613), and Bangladesh (N = 1722). The questions and designs of the list experiments are described in Appendix Table 1. The tabulations of the number of "yes" answers reflect country specific weights.

Figure 3: Not wearing masks in public - Stated and Elicited



Notes: Means and 95% confidence intervals Stated mask non-compliance includes respondents who did not report either “All of the time” or “Most of the time” to the direct question on mask wearing practices over the past seven days. The elicited mask non-compliances are measured using list experiments. All estimates include country specific weights to ensure representability.

Table 1: Summary Statistics¹

	Kenya	Nigeria	Bangladesh
	(N = 1647)	(N = 1613)	(N = 1722)
= 1 if female	0.63 (0.48)	0.58 (0.49)	0.55 (0.50)
Age (in years)	31.75 (10.48)	31.20 (10.01)	39.88 (13.56)
= 1 if currently married	0.52 (0.50)	0.49 (0.50)	0.83 (0.37)
= 1 if employed	0.73 (0.44)	0.64 (0.48)	0.27 (0.45)
= 1 if experienced an income shock ²	0.33 (0.47)	0.10 (0.30)	0.03 (0.18)
= 1 if living in urban areas	0.50 (0.50)	0.65 (0.48)	0.46 (0.50)
= 1 if respondent has children	0.72 (0.45)	0.70 (0.46)	0.86 (0.34)
Total number of children	0.73 (0.90)	0.83 (1.29)	1.76 (1.28)
Household size	3.27 (1.70)	4.10 (2.74)	5.03 (2.38)
= 1 if knows people with COVID-19 ³	0.30 (0.46)	0.08 (0.27)	0.18 (0.38)
= 1 if considers vulnerable to COVID-19 ⁴	0.55 (0.50)	0.18 (0.39)	0.17 (0.38)
= 1 if decide her/himself to wear mask ⁵	0.67 (0.47)	0.73 (0.44)	0.70 (0.46)
= 1 if food insecure over last 7 days	0.49 (0.50)	0.51 (0.50)	0.14 (0.35)
Score for forward lookingness ⁶	7.84 (3.27)	8.08 (2.82)	3.23 (3.32)
Respondent's education categories Kenya			
Primary or below	16.7%		
Secondary	38.7%		
Tertiary	44.6%		
Respondent's education categories Nigeria			
Secondary		44.0%	
Higher		56.0%	
Respondent's education categories Bangladesh			
Pre-primary			31.7%
Primary			36.7%
SSC or higher ⁷			31.6%

Notes:

Standard errors are presented in parentheses.

¹ All statistics presented are weighted to adjust for the representativeness of individuals by gender, age range, and rural/urban location using national representative surveys; and weights are applied using inverse probability weighting.

² An income shock is defined as lived in a household where at least one member experienced either 1) job loss, 2) nonfarm business closure, or 3) disruption of farming, livestock, fishing activities in the past 12 months.

³ Self-reported responses to whether they personally know anyone that has, or has had, COVID-19, including self, other family members living in the same household, family members living outside of the household, friends/neighbours in the same community, friends living outside of the same community, people from work, including colleagues, boss, clients, etc., and other.

⁴ Perceived COVID-19 risk of self or any other household member contracting COVID-19.

⁵ Defined as being the only person to decide wearing face masks to protect again COVID-19 in the household.

⁶ The forward lookingness is measured through the question “How willing are you to give up something that is beneficial for you today in order to benefit more from that in the future?”. The self-reported scores are on a scale from 0 to 10, where 0 means “completely unwilling to do so” and a 10 means “very willing to do so”.

⁷ SSC stands for Secondary School Certificate, which means Class 10th in Bangladesh.

Table 2: Multivariable analyses for the list experiments

VARIABLES	(1) Kenya	(2) Nigeria	(3) Bangladesh
= 1 if female	-0.017 (0.107)	0.133 (0.111)	-0.08 (0.192)
Age (in years)	0.003 (0.005)	-0.007 (0.006)	-0.004 (0.006)
= 1 if currently married	0.115 (0.118)	0.136 (0.146)	0.337 (0.198)
= 1 if employed	0.153 (0.119)	0.007 (0.115)	0.214 (0.183)
= 1 if experienced an income shock	0.008 (0.108)	0.029 (0.174)	-0.05 (0.290)
= 1 if living in urban areas	-0.078 (0.103)	-0.027 (0.112)	-0.015 (0.149)
= 1 if respondent has children	0.066 (0.129)	-0.102 (0.136)	0.359 (0.236)
Total number of children	0.011 (0.084)	-0.072 (0.059)	-0.087 (0.088)
Household size	-0.043 (0.045)	0.032 (0.026)	0.001 (0.041)
= 1 if knows people with COVID-19	-0.007 (0.111)	-0.065 (0.203)	-0.066 (0.203)
= 1 if considers vulnerable to COVID-19	0.006 (0.105)	0.054 (0.134)	0.106 (0.215)
= 1 if decide her/himself to wear mask	0.044 (0.109)	0.056 (0.127)	-0.502** (0.185)
= 1 if food insecure over last 7 days	-0.029 (0.103)	-0.155 (0.108)	-0.13 (0.200)
Score for forward lookingness	-0.011 (0.016)	0.013 (0.019)	0.028 (0.023)
Respondent's education categories Kenya			
Primary			
Secondary	Base		
	-0.241		
Tertiary	(0.156)		
	-0.218		
Respondent's education categories Nigeria			
Secondary		Base	
Higher		0.049	
		(0.110)	
Respondent's education categories Bangladesh			
Pre-primary			Base
Primary			-0.330
			(0.173)
SSC or higher			-0.411*
			(0.174)
Observations	1,647	1,608	1,722

Notes: Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. All statistics presented are weighted to adjust for the representation of individuals by gender, age range and rural/urban region using national representative surveys; and weights are applied using inverse probability weighting.

Table 3: Multivariable analyses for stated noncompliance with wearing facemask

VARIABLES	(1) Kenya	(2) Nigeria	(3) Bangladesh
= 1 if female	-0.049** (0.018)	-0.077** (0.024)	-0.01 (0.031)
Age (in years)	-0.002* (0.001)	-0.003 (0.001)	-0.002 (0.001)
= 1 if currently married	-0.023 (0.019)	-0.005 (0.030)	0.007 (0.038)
= 1 if employed	0.021 (0.017)	-0.008 (0.024)	0.015 (0.033)
= 1 if experienced an income shock	-0.012 (0.017)	0.024 (0.039)	-0.147** (0.052)
= 1 if living in urban areas	-0.070*** (0.016)	0.006 (0.023)	-0.130*** (0.027)
= 1 if respondent has children	0.007 (0.020)	-0.003 (0.028)	0.035 (0.043)
Total number of children	0.013 (0.012)	-0.002 (0.011)	-0.011 (0.017)
Household size	-0.003 (0.006)	0 (0.005)	0.001 (0.007)
= 1 if knows people with COVID-19	-0.055** (0.017)	-0.068 (0.038)	-0.072* (0.035)
= 1 if considers vulnerable to COVID-19	-0.031 (0.017)	-0.085** (0.027)	-0.01 (0.038)
= 1 if decide her/himself to wear mask	-0.005 (0.017)	-0.021 (0.027)	0.025 (0.030)
= 1 if food insecure over last 7 days	0.016 (0.017)	-0.01 (0.023)	-0.039 (0.039)
Score for forward lookingness	-0.004 (0.002)	0.003 (0.004)	-0.016*** (0.004)
Respondent's education categories Kenya			
Primary	Base		
Secondary	0.029 (0.026)		
Tertiary	0.033 (0.026)		
Respondent's education categories Nigeria			
Secondary		Base	
Higher		-0.019 (0.023)	
Respondent's education categories Bangladesh			
Pre-primary			Base
Primary			-0.063* (0.031)
SSC or higher			-0.136*** (0.031)
Observations	1,647	1,608	1,722

Note: The dependent variable is a binary outcome indicating not regularly wearing mask in public. We are reporting the marginal effects from logit regressions. *** p<0.01, ** p<0.05, * p<0.1. All statistics presented is weighted to adjust for the representation of individuals by gender, age range and rural/urban region using national representative surveys; and weights are applied using inverse probability weighting.

Table 4: Comparing Stated and Elicited Facemask Compliance

		Kenya		Nigeria		Bangladesh	
		Difference	p-value	Difference	p-value	Difference	p-value
Gender	Male	0.380	0.000	0.170	0.035	0.343	0.004
	Female	0.383	0.000	0.395	0.000	0.233	0.014
Age^a	Below median	0.256	0.001	0.339	0.000	0.297	0.008
	Above median	0.511	0.000	0.249	0.001	0.186	0.054
Marital status	Unmarried	0.267	0.000	0.261	0.001	-0.204	0.198
	Married	0.489	0.000	0.347	0.000	0.340	0.000
Job status	Unemployed	0.230	0.024	0.345	0.000	0.224	0.011
	Employed	0.438	0.000	0.253	0.000	0.423	0.004
Employment shock	No	0.357	0.000	0.304	0.000	0.263	0.001
	Yes	0.441	0.000	0.275	0.112	-0.079	0.789
Location	Rural	0.390	0.000	0.319	0.001	0.220	0.038
	Urban	0.373	0.000	0.294	0.000	0.308	0.007
Have a child?	No	0.333	0.001	0.367	0.000	-0.092	0.595
	Yes	0.402	0.000	0.275	0.000	0.316	0.000
Household Size	1-2	0.313	0.001	0.221	0.038	-0.208	0.462
	3-5	0.385	0.000	0.301	0.000	0.254	0.001
	6+	0.151	0.463	0.385	0.003	0.285	0.038
Knows somebody with Covid	No	0.365	0.000	0.302	0.000	0.277	0.001
	Yes	0.422	0.000	0.297	0.138	0.199	0.311
Considers him/herself vulnerable to Covid	No	0.346	0.000	0.272	0.000	0.237	0.004
	Yes	0.436	0.000	0.455	0.000	0.365	0.083
Decides him/herself whether to wear a mask	No	0.328	0.000	0.288	0.008	0.604	0.000
	Yes	0.412	0.000	0.310	0.000	0.112	0.189
Food insecure over past 7 days	No	0.368	0.000	0.369	0.000	0.288	0.001
	Yes	0.395	0.000	0.212	0.008	0.103	0.629
Score for forward lookingness^a	Below median	0.407	0.000	0.308	0.000	0.129	0.190
	Above median	0.357	0.000	0.270	0.005	0.330	0.005
Kenya education categories	Primary	0.662	0.000				
	Secondary	0.303	0.000				
	Tertiary	0.356	0.000				
Nigeria education categories	Secondary			0.259	0.002		
	Higher			0.321	0.000		
Bangladesh education categories	Pre-primary					0.312	0.009
	Primary					0.280	0.051
	SSC or higher					0.173	0.205

^a Median values are calculated within each country.

Unmasking the truth: experimental evidence of facemask compliance in Bangladesh, Kenya, and Nigeria during the COVID-19 pandemic

Supplementary Materials

Survey Weights

In our analysis, we use sampling weights to adjust for effects of the RDD sampling methodology as well as attrition bias due to the challenges associated with recontacting all baseline respondents. The sampling weights were constructed by Innovations for Poverty Action (IPA) for the Bangladesh, Kenya, and Nigeria surveys using nationally representative surveys. Specifically, we used the 2016 Bangladesh Household, Income, and Expenditure Survey, the 2016 Kenya Integrated Household Budget Survey and the 2019 Kenya Population and Housing Census in Kenya, and the 2018-9 Nigeria General Household Panel Survey in Nigeria. The weights were designed to adjust for the proportion of individuals in our sample relative to what would be expected in a nationally representative survey according to gender, age, region, and phone ownership following the approach adopted by the World Bank.³⁴ Within-cell post-stratification weights are constructed for the baseline surveys in all countries and then top-coded at the 99th percentile.³⁵ In Kenya and Nigeria, an attrition correction factor was derived based on the response rate within demographic group bins in the second round.³⁵

Table A1: List Experiment Design

Considering your behaviour over the past 7 days, how many of the following statements are true statements?		
	Version 1: Control Group	Version 2: Treatment Group
Called a friend/family member	1 = Yes 0 = No	1 = Yes 0 = No
Listened to the radio		
Drove in a car or on a motorcycle		
Sent someone mobile money		
<u>Left the house without a face mask</u>	Not included	
<u>Total possible score</u>	0-4	0-5

Table A2: Balance Test

	Kenya			Nigeria			Bangladesh		
	Control	Treatme nt	p-value	Control	Treatme nt	p-value	Control	Treatme nt	p-value
	(N = 824)	(N = 824)		(N = 765)	(N = 848)		(N = 862)	(N = 860)	
= 1 if female	0.64 (0.48)	0.63 (0.48)	0.784	0.58 (0.49)	0.57 (0.49)	0.894	0.53 (0.50)	0.57 (0.50)	0.093
Age (in years)	31.42 (10.06)	32.09 (10.88)	0.194	31.05 (10.22)	31.34 (9.82)	0.557	39.94 (13.54)	39.83 (13.58)	0.859
= 1 if currently married	0.51 (0.50)	0.53 (0.50)	0.329	0.50 (0.50)	0.48 (0.50)	0.449	0.82 (0.38)	0.84 (0.37)	0.387
= 1 if employed	0.72 (0.45)	0.73 (0.44)	0.711	0.62 (0.49)	0.65 (0.48)	0.163	0.29 (0.45)	0.25 (0.43)	0.064
= 1 if experienced an income shock	0.33 (0.47)	0.33 (0.47)	0.869	0.10 (0.31)	0.09 (0.29)	0.422	0.03 (0.18)	0.03 (0.17)	0.546
= 1 if living in urban areas	0.49 (0.50)	0.52 (0.50)	0.353	0.65 (0.48)	0.64 (0.48)	0.686	0.48 (0.50)	0.44 (0.50)	0.130
= 1 if respondent has children	0.73 (0.45)	0.71 (0.45)	0.492	0.69 (0.46)	0.70 (0.46)	0.420	0.88 (0.33)	0.85 (0.36)	0.073
Total number of children	0.73 (0.89)	0.73 (0.90)	0.959	0.79 (1.34)	0.85 (1.24)	0.368	1.79 (1.25)	1.73 (1.31)	0.329
Household size	3.27 (1.74)	3.28 (1.65)	0.850	4.08 (2.80)	4.12 (2.69)	0.745	5.08 (2.46)	4.99 (2.29)	0.437
= 1 if knows people with COVID-19	0.30 (0.46)	0.31 (0.46)	0.757	0.08 (0.28)	0.08 (0.27)	0.880	0.19 (0.39)	0.16 (0.37)	0.175
= 1 if considers vulnerable to COVID-19	0.57 (0.49)	0.53 (0.50)	0.078	0.19 (0.40)	0.17 (0.38)	0.253	0.18 (0.39)	0.17 (0.37)	0.461
= 1 if decide her/himself to wear weak mask	0.67 (0.47)	0.67 (0.47)	0.961	0.74 (0.44)	0.73 (0.44)	0.590	0.72 (0.45)	0.68 (0.47)	0.096
= 1 if food insecure over last 7 days	0.50 (0.50)	0.49 (0.50)	0.541	0.54 (0.50)	0.49 (0.50)	0.025	0.13 (0.33)	0.15 (0.36)	0.145
Score for forward lookingness	7.91 (3.27)	7.76 (3.28)	0.327	8.16 (2.80)	8.01 (2.84)	0.290	3.14 (3.37)	3.33 (3.27)	0.227
Respondent's education categories Kenya									
Primary	16.8%	16.6%	0.728						
Secondary	45.4%	43.8%							
Tertiary	16.8%	16.6%							
Respondent's education categories Nigeria									
Secondary				43.2%	42.2%	0.662			
Higher				56.8%	57.8%				
Respondent's education categories Bangladesh									
Pre-primary							31.1%	32.3%	0.849
Primary							37.5%	35.9%	
SSC or higher							31.4%	31.8%	

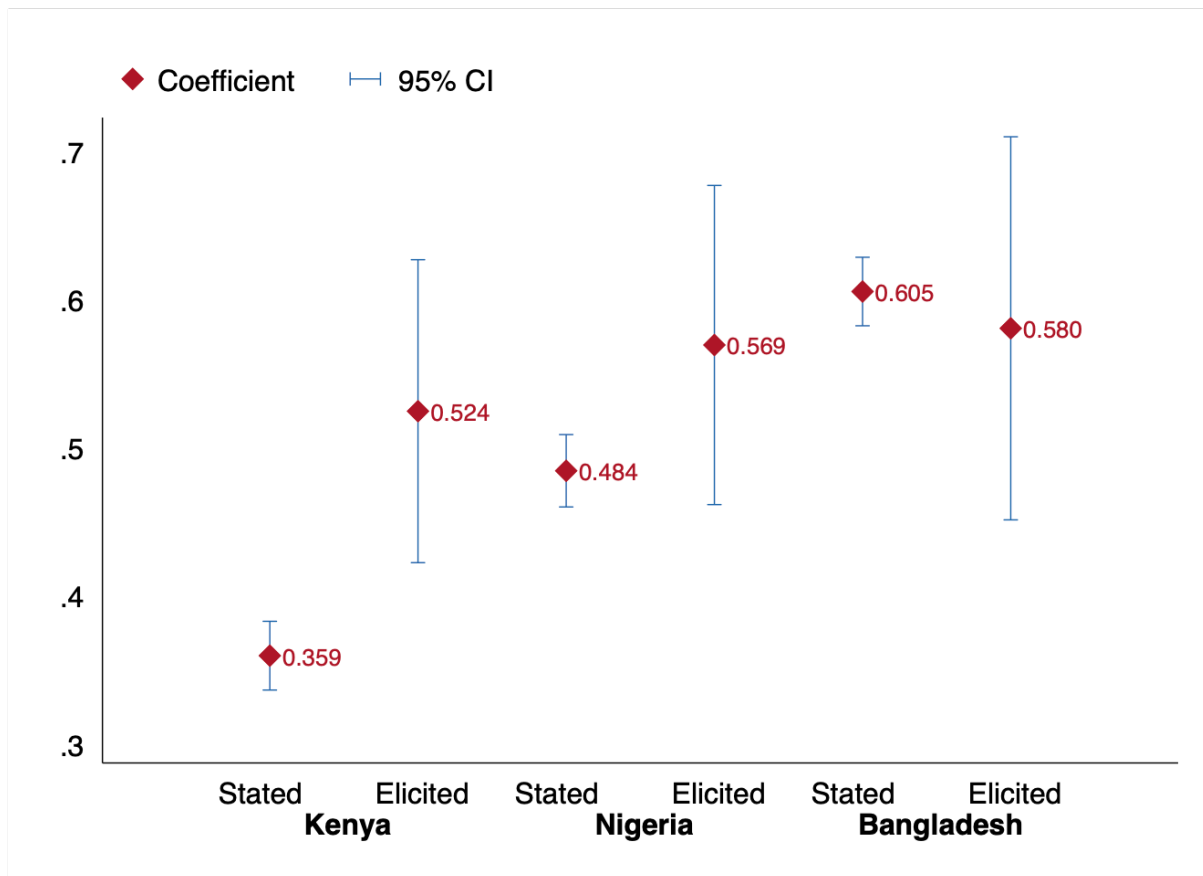
Table A3: Design Test

	Kenya		Nigeria		Bangladesh	
	$\pi_{R,S}$	p-value	$\pi_{R,S}$	p-value	$\pi_{R,S}$	p-value
Pr(R=0,S=1)	0.002	0.842	0.004	0.746	0.092	1.000
Pr(R=0,S=0)	0.001	0.842	0.013	1.000	0.320	1.000
Pr(R=1,S=1)	0.050	1.000	0.063	1.000	0.160	1.000
Pr(R=1,S=0)	0.050	1.000	0.058	1.000	0.152	1.000
Pr(R=2,S=1)	0.133	1.000	0.167	1.000	0.121	1.000
Pr(R=2,S=0)	0.096	1.000	0.126	1.000	0.026	0.891
Pr(R=3,S=1)	0.157	1.000	0.183	1.000	0.102	1.000
Pr(R=3,S=0)	0.155	1.000	0.143	1.000	-0.001 ^a	0.464
Pr(R=4,S=1)	0.182	1.000	0.152	1.000	0.105	1.000
Pr(R=4,S=0)	0.174	1.000	0.091	1.000	-0.077 ^b	0.000

Note: $\pi_{R,S}$ sa

Basic findings: No problem for Kenya and Nigeria. Somewhat minor problem for Bangladesh

Figure A1: Differences between stated and elicited non-compliance with facemask wearing



References

1. Lee, J.-K. *et al.* Institutional and behaviour-change interventions to support COVID-19 public health measures: a review by the Lancet Commission Task Force on public health measures to suppress the pandemic. *Int Health* ihab022- (2021) doi:10.1093/inthealth/ihab022.
2. World Health Organization. *Advice on the use of masks in the context of COVID-19: interim guidance, 5 June 2020*. <https://apps.who.int/iris/handle/10665/332293> (2020).
3. Abaluck, J. *et al.* The Impact of Community Masking on COVID-19: A Cluster-Randomized Trial in Bangladesh. *IPA Working Paper* (2021).
4. Greenhalgh, T., Schmid, M. B., Czypionka, T., Bassler, D. & Gruer, L. Face masks for the public during the covid-19 crisis. *Bmj* 369, m1435 (2020).
5. Fodjo, J. N. S. *et al.* Mass masking as a way to contain COVID-19 and exit lockdown in low- and middle-income countries. *J Infection* 81, e1–e5 (2020).
6. Feng, S. *et al.* Rational use of face masks in the COVID-19 pandemic. *Lancet Respir Medicine* 8, 434–436 (2020).
7. Leung, N. H. L. *et al.* Respiratory virus shedding in exhaled breath and efficacy of face masks. *Nat Med* 26, 676–680 (2020).
8. Leffler, C. T. *et al.* Association of Country-wide Coronavirus Mortality with Demographics, Testing, Lockdowns, and Public Wearing of Masks. *Am J Tropical Medicine Hyg* 103, 2400–2411 (2020).
9. Cheng, Y. *et al.* Face masks effectively limit the probability of SARS-CoV-2 transmission. *Science* 372, 1439–1443 (2021).
10. Chu, D. K. *et al.* Physical distancing, face masks, and eye protection to prevent person-to-person transmission of SARS-CoV-2 and COVID-19: a systematic review and meta-analysis. *Lancet* 395, 1973–1987 (2020).
11. Jakubowski, A. *et al.* Self-reported vs Directly Observed Face Mask Use in Kenya. *Jama Netw Open* 4, e2118830 (2021).
12. Reuben, R. C., Danladi, M. M. A., Saleh, D. A. & Ejembi, P. E. Knowledge, Attitudes and Practices Towards COVID-19: An Epidemiological Survey in North-Central Nigeria. *J Commun Health* 46, 457–470 (2021).
13. Ferdous, Most. Z. *et al.* Knowledge, attitude, and practice regarding COVID-19 outbreak in Bangladesh: An online-based cross-sectional study. *Plos One* 15, e0239254 (2020).
14. Islam, S. *et al.* Knowledge, attitudes and practices associated with the COVID-19 among slum dwellers resided in Dhaka City: a Bangladeshi interview-based survey. *J Public Health* 43, fdad182- (2020).
15. Blair, G. & Imai, K. Statistical Analysis of List Experiments. *Polit Anal* 20, 47–77 (2012).

16. Bell, S. O. & Bishai, D. Can a List Experiment Improve Validity of Abortion Measurement? *Stud Family Plann* 50, 43–61 (2019).
17. Porter, C., Favara, M., Sánchez, A. & Scott, D. The impact of COVID-19 lockdowns on physical domestic violence: Evidence from a list randomization experiment. *Ssm - Popul Heal* 14, 100792 (2021).
18. Lépine, A., Treibich, C., Ndour, C. T., Gueye, K. & Vickerman, P. HIV infection risk and condom use among sex workers in Senegal: evidence from the list experiment method. *Health Policy Plann* 35, 408–415 (2020).
19. Treibich, C. & Lépine, A. Estimating misreporting in condom use and its determinants among sex workers: Evidence from the list randomisation method. *Health Econ* 28, 144–160 (2018).
20. Becher, M., Stegmueller, D., Brouard, S. & Kerrouche, E. Comparative experimental evidence on compliance with social distancing during the COVID-19 pandemic. *Medrxiv* 2020.07.29.20164806 (2020) doi:10.1101/2020.07.29.20164806.
21. Timmons, S., McGinnity, F., Belton, C., Barjaková, M. & Lunn, P. It depends on how you ask: measuring bias in population surveys of compliance with COVID-19 public health guidance. *J Epidemiol Commun H* 75, 387–389 (2021).
22. Muraya, J. Kenya: Masks Now Mandatory in Public Places, Kenya Declares. *AllAfrica.com* (2020).
23. Ogoina, D. COVID-19: The Need for Rational Use of Face Masks in Nigeria. *Am J Tropical Medicine Hyg* 103, 33–34 (2020).
24. Glazerman, S., Grépin, K. A., Mueller, V., Rosenbaum, M. & Wu, N. Do referrals improve the representation of women in mobile phone surveys?
25. Mueller, V. *et al.* Food insecurity and COVID-19 risk in low- and middle-income countries. *Appl Econ Perspect P* (2021) doi:10.1002/aepp.13200.
26. BRAC James P Grant School of Public Health & National Nutrition Services. *State of food security and nutrition in Bangladesh 2018-2019*. (2019).
27. Falk, A. *et al.* Global Evidence on Economic Preferences. *Q J Econ* 133, 1645–1692 (2018).
28. Ollila, H. M. *et al.* Face masks to prevent transmission of respiratory diseases: Systematic review and meta-analysis of randomized controlled trials*. *Medrxiv* 2020.07.31.20166116 (2021) doi:10.1101/2020.07.31.20166116.
29. Leech, G. *et al.* Mass mask-wearing notably reduces COVID-19 transmission. *Medrxiv* 2021.06.16.21258817 (2021) doi:10.1101/2021.06.16.21258817.
30. Eikenberry, S. E. *et al.* To mask or not to mask: Modeling the potential for face mask use by the general public to curtail the COVID-19 pandemic. *Infect Dis Model* 5, 293–308 (2020).
31. Daoust, J.-F. *et al.* A guilt-free strategy increases self-reported non-compliance with COVID-19 preventive measures: Experimental evidence from 12 countries. *Plos One* 16, e0249914 (2021).
32. Daoust, J.-F. *et al.* How to Survey Citizens' Compliance with COVID-19 Public Health Measures: Evidence from Three Survey Experiments. *J Exp Political Sci* 1–8 (2020) doi:10.1017/xps.2020.25.

33. Abaluck, J. *et al.* Normalizing Community Mask-Wearing: A Cluster Randomized Trial in Bangladesh. *NBER Working Paper 28734* (2021) doi:10.3386/w28734.
34. Pape, U. J., Sinha, N. & Delius, A. J. S. *COVID-19 Rapid Response Phone Survey Households 2020-2021, Panel*. <https://microdata.worldbank.org/index.php/catalog/3774> (2020).
35. Kastelic, K. H. *et al.* *High Frequency Mobile Phone Surveys of Households to Assess the Impacts of COVID-19 (Vol. 2) : Guidelines on Sampling Design*. <https://documents.worldbank.org/en/publication/documents-reports/documentdetail/742581588695955271/guidelines-on-sampling-design> (2020).