Prosocial Vaccination*

Bakhtawar Ali†, Syngjoo Choi‡, Jun Hyung Kim§, and Avner Seror¶

January 9, 2024

Abstract

We study the role of loss aversion in behaviors with significant prosocial motives. We conduct a field experiment in a large developing country that uses gain-loss framing to deliver information about how vaccination reduces the spread of COVID-19, and confirm vaccine take-up using self reports and government certificates. For both vaccination intentions and vaccine take-up, we find positive impact of information treatment, with significantly greater impact for the messages focusing on the "loss" in prosocial motives when not vaccinated. We explain our findings in a theoretical model and calibrate the prosocial loss aversion parameter to be within range of private loss aversion parameter estimates in the literature.

Keywords: Altruism, Loss Aversion, Field experiment, Prosocial Behavior, COVID-19, Vaccination.

JEL codes: D8, D9, I18

*Acknowledgement. This paper was supported by National Natural Science Foundation of China (grant no. 72150410443) and by Creative-Pioneering Researchers Program from Seoul National University. We thank the participants at the Department of Economics seminar at Korea University and the Africa Meeting of the Econometric Society 2022 for helpful comments and suggestions. Yuexin Huang at Jinan University provided excellent research assistance. Declarations of interest: none.

†Aix-Marseille School of Economics; bakhtawar.ali@univ-amu.fr

‡Seoul National University; syngjooc@snu.ac.kr

§Jinan University, Tianhe-qu, Guangzhou, China; kimjunhyung@jnu.edu.cn; corresponding author

¶Aix-Marseille School of Economics; avner.seror@univ-amu.fr

1 Introduction

Vaccination is one of the most effective ways to improve public health, but is met with hesitancy and skepticism around the world (Dubé et al. 2013; Solís Arce et al. 2021). Because individuals may carry infectious diseases without serious symptoms and infect others who may be more vulnerable, it is important for public health to promote vaccination even to those who may not directly suffer from the diseases. During the COVID-19 pandemic, therefore, public vaccination campaigns often attempted to overcome vaccine hesitancy by emphasizing the benefits of vaccines both to the individuals and to the whole community.[1] Such campaigns act on the *prosocial* or social responsibility dimension of vaccination: getting vaccinated can protect others' health even when the individuals are not vulnerable to the disease.

Despite the widespread use of vaccination campaigns emphasizing prosocial benefits, only recently have researchers started assessing the effects of prosocial messages on vaccinations. The effects seem mixed and dependent on the context and the manner of delivering the messages (Böhm and Betsch 2022; Batteux et al. 2022). Therefore, our first question is whether prosocial messages increase vaccination intentions.

Furthermore, the effects of the messages may depend on how they are framed. Behavioral economics literature shows that framing decisions as losses tends to elicit stronger responses from decision makers than framing them as gains. This phenomenon is known as loss aversion and has large empirical support with respect to decisions affecting private utility (Brown et al. 2021). The effects of gain-loss framing on prosocial behaviors, such as vaccination,

[1] Although we could not find systematic review of public campaigns emphasizing prosocial benefits of vaccination, casual internet search revealed many examples of public campaigns emphasizing community benefits of vaccination by World Health Organization, government departments such as the US Department of Health and Human Services, and various news media. An example is Parsons (2021), titled "Getting a COVID-19 vaccine isn't just about you — it protects others in your life."

remain largely unknown in the literature.[2] As our second question, we ask whether prosocial behaviors, vaccination decisions in particular, depend on gain-loss framing.

To answer these two questions, we study the prosocial motives behind vaccination intentions in the context of the ongoing COVID-19 pandemic. We conduct a randomized field study in Pakistan where treatments consist of providing subjects with scientific information on the effects of receiving COVID-19 vaccines. Our main treatments contain the same informational content: vaccination has been shown in scientific studies to reduce the risk of transmitting COVID-19 to others. However, we frame the treatments as gains or losses of *the health of others*. Specifically, we inform individuals of the lower probability of infecting others when vaccinated (compared to when unvaccinated), or the higher probability of infecting others when unvaccinated (compared to when vaccinated). Therefore, the gain-loss framing primarily applies to prosocial utility benefits of vaccination.

Our approach is motivated by seminal works in linguistics arguing that language is more than a functional communication system (Chomsky 2009). A sentence has both the surface structure and the deep structure, an inner mental aspect, that conveys its semantic meaning (Arnauld and Lancelot 1810). The mind uncovers the deep structure by dividing a sentence into interrelated propositions. We hypothesize that the inner mental aspects of two sentences emphasizing either the loss or the gain associated with vaccination will be different, because these sentences set different mental reference points. A sentence emphasizing the gain associated with the vaccination *relative to a situation where the subject is not vaccinated* implicitly sets a mental reference point where the subject is not vaccinated. A sentence that emphasizes the loss associated with not getting vaccinated *relative to a situation where the subject is vaccinated* implicitly sets a mental reference point where the subject is vaccinated.

[2]One notable exception is Fatas and Restrepo-Plaza (2022), who test loss aversion in the context of prosocial preferences. The authors study several interventions aiming to reduce victims' reluctance to forgive their offenders. The authors tested whether emphasizing the gain (successful rehabilitation) or the loss (risk of recidivism) of a restorative justice program affect subjects' behavior, and find evidence of loss aversion.

Loss-averse subjects would react more strongly to the information treatment if the relevant decision were posed as a loss against the reference point set by the treatment message.[3]

Using probabilistic measures for vaccination intentions (Manski 2004), we find that both gain- and loss-framed prosocial messages increase vaccination intentions, but loss-framed messages tend to have a greater impact. Specifically, providing subjects with scientific information on the benefits of vaccination in reducing the probability of transmitting COVID-19 increases their vaccination intentions by 5 to 10 percentage points, regardless of framing. However, subjects' vaccination intentions increase further when they are informed of their higher chances of infecting others when not vaccinated compared to when they are informed of their lower chances of infecting others when vaccinated. The effects of loss-framed messages are greater by 2 to 4 percentage points than those of gain-framed messages. This evidence is consistent with the presence of loss aversion to prosocial benefits in the context of vaccination decisions.

The effects of the treatment are significant across different individual characteristics or hypothetical conditions that could discourage vaccination decisions. The effects are found for different hypothetical monetary costs of vaccination and hypothetical neighborhood environments; for subjects of different ages, genders, education levels; for subjects with low trust in government, science, medicine, and the press, who are typically predicted to have low vaccination intentions (Lazarus et al. 2021); and for those who showed low vaccination intentions before treatment.

To explain these findings, we build a theoretical model where subjects' utility is affected by the cost of infecting others and the gains or losses subjects feel when comparing their decision with a reference point. In the model, the treatments shift subjects' reference points. Gain-framed message treatment sets the reference point as being unvaccinated, against which

[3]The process of investigating the deep structure of a sentence can be interpreted as a linguistic foundation of the "editing" phase of the choice process in prospect theory. The "coding" operation in which the subject interprets an outcome as either a gain or a loss (Kahneman and Tversky 1979) may occur by the subject placing the sentences describing the outcomes in interrelated positions so that the emphasized outcome is compared to the outcome not emphasized, which becomes the reference point.

the subject evaluates the choice of whether to vaccinate. Loss-framed message treatment sets the reference point as being vaccinated, against which the subject evaluates the choice of whether to remain unvaccinated. The model implies that to the extent that subjects are loss averse, they will tend to report higher intention of getting vaccinated when the information treatment gives them a reference point where they get vaccinated. Indeed, in that treatment, subjects perceive the loss they feel when they do not get vaccinated and potentially transmit the disease relative to a reference point where they are vaccinated. Moreover, we show that the magnitude of the effect will depend on several key factors, including the monetary cost of the vaccine and the fraction of vaccinated individuals in the population.

Under standard assumptions on the utility specification, the model gives linear characterizations for subjects' vaccination intentions as a function of the reference point determined by the treatment. We can therefore identify subjects' loss aversion parameter in prosocial preferences using our empirical analysis. Our estimates of loss aversion parameter, based on prosocial preferences, are in line with the estimates in the literature based on non-prosocial preferences (Brown et al. 2021).

This paper relates to several strands of the literature. First, we contribute to the recent literature on the economics of vaccinations. Vaccine hesitancy remains a serious problem in both developing and developed countries, posing major threats to controlling the current pandemic and future health risks (Bloom, Kuhn and Prettner 2022; Dabla-Norris et al. 2021). Our contribution is that we show the effectiveness of providing scientific information combined with prosocial messages under gain-loss framing. Other studies examined the effectiveness of providing information on vaccines (Altay et al. 2021; Motta et al. 2021; Pfattheicher, Petersen and Böhm 2022) but they did not always lead to increases in vaccination intentions (Dai et al. 2021; Kachurka, Krawczyk and Rachubik 2021; Kerr et al. 2021; Batteux et al. 2022). A messaging intervention by Chen et al. (2021) using gain-loss framing to emphasize the selfish benefits of vaccines is found to be ineffective, whereas another intervention by Sasaki, Saito and Ohtake (2022), also using gain-loss framing on the impact

of own vaccination decisions on the decisions of others, is shown to impact only older adults and younger adults with high baseline vaccination intentions. Our intervention combines scientific information and gain- or loss-framed messages regarding the prosocial benefits of vaccination for the health of others, directly addressing the public health motivation for vaccination (Dubé et al. 2021). It is found to be effective across individuals with different individual preferences and observed characteristics that may determine the perceived benefits and costs of vaccination, such as demographic characteristics, initial vaccine hesitancy, and self-reported trust in government, press, science, and medicine.

Second, we investigate loss aversion in teh context of prosocial preferences, contributing to the vast loss aversion literature (Kahneman and Tversky 1979; Kőszegi and Rabin 2006; Fiedler and Hillenbrand 2020; Brown et al. 2021; Kimbrough, Porter and Schneider 2021). Loss aversion affects subjects' selfish preferences, but its effect on prosocial behavior remains poorly studied. A recent study by Fatas and Restrepo-Plaza (2022) investigate several interventions in Colombia aiming to reduce victims' reluctance to forgive their offenders. The authors show that emphasizing the loss implied by forgiving (risk of recidivism) has a significantly larger effect than emphasizing the gain (successful rehabilitation).[4] In this paper, we also find evidence of strong loss aversion in the context of prosocial preferences. Through our theoretical approach, we calibrate the loss aversion parameter of the model and compare it with existing parameter values in the economic literature following similar models.[5]

Third, we contribute to the experimental literature by showing that the framing of scientific information is critical in explaining both the magnitude and heterogeneity of information treatment effects.[6] The previous literature shows that the effects of information treatments

[4]Several studies in the psychology literature use lab experiments to examine the role of loss aversion in pro-social motives, but the results are mixed (Feng et al. 2021).

[5]Our theoretical approach draws on the seminal model of Kőszegi and Rabin (2006).

[6]Information treatments have been shown to increase donations (Shang and Croson 2009); change the behavior of men in patriarchal societies to be more gender egalitarian (Bursztyn, González and Yanagizawa-Drott 2020); promote prosocial behavior (Mehmood, Naseer and Chen 2021); improve prescribing behaviors of physicians (Dubois and Tunçel 2021); and change social-distancing behaviors during the COVID-19 pandemic (Allen IV et al. 2021).

on health-related behaviors, including social distancing and medicine prescription, may be heterogeneous based on perceived costs and benefits as well as individual preferences (Sasaki, Kurokawa and Ohtake 2021; Allen IV et al. 2021; Dubois and Tunçel 2021). Our results suggest that appropriate framing of messages focusing on prosocial benefits can be highly effective.

The rest of the paper is organized as follows. Section 2 presents the experimental design and the data. Section 3 show the empirical results. Section 4 develops the theoretical model, and Section 5 concludes. The theoretical proofs are relegated to the appendix.

2 Experimental Design and Empirical Strategy

2.1 Experimental Design and Background

The experiment is set in Pakistan, a large and populous developing country with a population of over 200 million. Despite the government's efforts to provide vaccines free of charge to large masses, the vaccination rate remained low and vaccine hesitancy remained high among Pakistanis at the time of the experiment (Siddiqui et al. 2021). For example, only 2.3% of the Pakistani population completed the initial vaccination protocol by July 2021 (Our World in Data 2022).

We conducted our information provision experiment and survey in June 2021 by contacting individuals in ages 25 to 65 via cellular phones. Cellular phone surveys allow surveyors to maintain social distancing. Although over 85% of Pakistanis have access to cellular phones (Pakistan Telecommunication Authority 2022), this approach likely under-samples unemployed or less educated individuals. The survey was restricted to those in the Punjab area. The Punjab province contains almost half of the population of Pakistan and has a literacy rate of around 60% (Rehman, Jingdong and Hussain 2015), close to the national average. The survey was conducted in English and Urdu, the two official languages of Pakistan. The survey length was approximately 10 to 15 minutes. Participants received a compensation of

155 Pakistani rupees (approximately 1 USD at the time of the survey) upon completing the survey.[7]

The experiment is restricted to those who did not get two doses of COVID-19 vaccines, the standard for full vaccination at the time of the survey. Out of 3,199 subjects, 531 completed vaccination, leaving 2,668 eligible for the experiment. The high proportion of fully vaccinated individuals in our sample likely reflects our sample restrictions, such as the restriction to working age (25–65) and the Punjab region, which is the most industrialized of the four Pakistani regions.

The survey includes information on basic demographic characteristics including age, income, education level, marital status, job characteristics, religiosity, and self-rated health. Participants are also asked about their beliefs about the chances of getting infected with COVID-19 with or without vaccination.

The experiment consists of five arms, including one control arm and four treatment arms. The following statements were read to the subjects in the respective arms:

Gain Treatment:

> Scientific studies have found that vaccinated people are much less likely to pass on the COVID-19 virus to others. For example, a study from the United Kingdom found that a single dose of coronavirus vaccine can reduce the chance of transmitting COVID inside a household by half. If you are vaccinated, you are less likely to transmit the virus to others.

Gain+ Treatment: The same statement as the Gain Treatment, with the following sentences added:

> Also, others may be more willing to meet you in person without fear of infection. Your social life could be protected if you are vaccinated.

[7] According to World Bank, GDP per capita of Pakistan is 1188.9 in US dollars as of 2020.

Loss Treatment:

> Scientific studies have found that unvaccinated people are much more likely to pass on the COVID-19 virus. For example, a study from the United Kingdom found that not getting vaccinated can double the chance of transmitting Covid inside a household. If you are not vaccinated, you are more likely to transmit the virus to others.

Loss+ Treatment: The same statement as the Loss Treatment, with the following sentences added:

> Also, others may be less willing to meet you in person given that they fear that you will transmit them the virus. Your social life could be interrupted if you are not vaccinated.

Gain Treatment and Gain+ Treatment emphasize the prosocial "gain" of getting vaccinated, in which the risk of infecting others is reduced when getting vaccinated. Loss Treatment and Loss+ Treatment emphasize the prosocial "loss" of not getting vaccinated, in which the risk of infecting others is heightened when not getting vaccinated.[8] Subjects are allocated to each treatment arms with equal probability.

In estimation, we do not find significant differences between Gain Treatment and Gain+ Treatment, and Loss and Loss+ Treatment. We therefore combine the Gain and Gain+ arms together (GAIN Treatment) and Loss and Loss+ arms together (LOSS Treatment) so that subjects are allocated to the control group with 20% probability and each of the treatment arms with 40% probability.[9]

[8] It is possible that subjects infer private gains to own health from hearing message emphasizing prosocial benefits. Even so, the difference between gain- and loss-framed messages comes entirely from the subjects' responses to different framing of prosocial benefits.

[9] Gain+ Treatment and Loss+ Treatment further emphasize the "strategic" or "social interaction" aspect of seemingly altruistic behavior. It is worth pointing out that Gain Treatments and Loss Treatment may elicit both purely altruistic and strategic motives. If Gain+ Treatment is more effective than Gain Treatment, then it may be concluded that Gain+ provides motivation on top of purely altruistic motives by reminding the subjects of strategic gains. We provide empirical results and our theoretical model related to Gain+ Treatment and Loss+ Treatment in the appendix.

Our primary outcome of interest is probabilistic intention (Manski 2004). The respondents are read the following statements:

> Suppose the government offers you an opportunity to be vaccinated some time in the next three months. Please answer your intention in 0 to 100%.[10]

The survey question for vaccination intention is:

> What is the chance you will choose to get vaccinated if it were free?

This measure is consistent with the policy of the Pakistani government at the time of the survey and many other governments around the world. We repeat the question four more times, under four different conditions: (1) if you had to pay 1,500 rupees; (2) if you had to pay 3,000 rupees; (3) if 30% of others near you got vaccinated; and (4) if 70% of others near you got vaccinated. The ordering of these questions is randomized into eight different groups to remove potential concerns that respondents' perception may be biased by immediately preceding questions (Stewart et al. 2002). We control for these groups in our analyses.

We also survey trust in institutions, which is an important predictor of vaccination decision (Larson 2016; Trent et al. 2022). We ask:

> How much confidence do you have in the government institutions?
>
> How much confidence do you have in the press?
>
> How much confidence do you have in the scientific community?
>
> How much confidence do you have in the medicine community?

The respondents are prompted to answer with a number between 0 and 100, 100 indicating full confidence and 0 indicating an entire lack of confidence.

[10] The respondents already answered several probabilistic intention questions earlier in the survey, so they would be familiar with the question's format.

2.2 Data

Subjects who completed their second dose of vaccination are included in the survey but excluded from the treatment. They comprise 16.6% of the overall sample. Table 1 presents the descriptive statistics of the analysis sample, including those who did not receive any dose and those who received only one dose of COVID-19 vaccines. The average age of the subjects is 34.75 and only 32.2% are female (3.3% describe themselves as "Other gender," an officially recognized gender category in Pakistan). Surprisingly few people in the sample responded that they were employed for pay, potentially reflecting the temporary setback caused by the COVID-19 pandemic. Average vaccination intention post-treatment is higher than average vaccination intention pre-treatment under all five hypothetical conditions of vaccination cost and neighborhood environment.

We compare the characteristics of the not fully vaccinated (in the analysis sample) and the fully vaccinated (up to two doses, included in the overall sample) individuals in Tables A1 and A2. Compared to fully vaccinated individuals, those who are not fully vaccinated (in the analysis sample) are younger, more likely to report being in good health, report lower belief in the chance of experiencing symptoms from contracting COVID-19, and show lower trust in the government and press. These differences are consistent with the characteristics of individuals reporting low vaccination intentions in international surveys (Lazarus et al. 2021; Sherman et al. 2021; Solís Arce et al. 2021). Interestingly, those not fully vaccinated are not different from the fully vaccinated in their trust in science or medicine and are more likely to have graduated high school.

Table 2 presents the balance test by regressing pre-treatment characteristics on the treatment indicators. We do not find evidence of consistent violation of balance across the treatment arms except for a single significant difference in the lower panel in column (2). Balance tests using all four arms of treatment also do not reveal a consistent imbalance of baseline characteristics across treatment arms (Table A3).

Figure 1 presents the distribution of vaccination intention outcomes. Pre-treatment intention is presented in panel (a) and post-treatment intentions are in panels (b) to (f). The figure shows that the stated intentions are widely spread between 0 and 100, although post-treatment intentions seems to have shifted upwards compared to pre-treatment intentions regardless of hypothetical conditions.

Table 1: Descriptive Statistics of the Analysis Sample

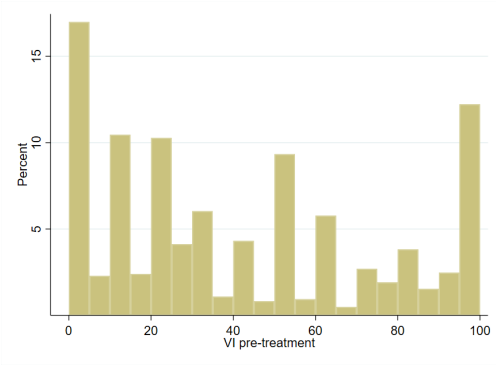
	Mean	SD	Min	Max	N
Age	34.750	10.161	25	65	2661
HS grad	0.407	0.491	0	1	2661
Married	0.511	0.500	0	1	2661
Working	0.427	0.495	0	1	2661
Self-reported health	0.877	0.328	0	1	2661
First dose	0.298	0.457	0	1	2661
Second dose	0	0	0	0	2661
Female	0.322	0.467	0	1	2522
Other gender	0.033	0.178	0	1	2522
Religiosity	2.757	2.081	0	5	2517
Chance of severe symptoms	26.831	24.525	0	100	2661
Chance of moderate symptoms	33.559	24.787	0	100	2661
Trust in government	63.053	25.215	0	100	2661
Trust in science	74.557	16.919	10	100	2661
Trust in medicine	78.600	19.202	0	100	2661
Trust in press	63.742	27.712	0	100	2661
Pre-trtment VI	39.986	33.777	0	100	2661
Post-trtment VI if free	54.938	29.689	0	100	2661
Post-trtment VI if 1500 rp	50.087	29.203	0	100	2661
Post-trtment VI if 3000 rp	51.005	30.012	0	100	2661
Post-trtment VI if 30 %	54.240	28.200	0	100	2661
Post-trtment VI if 70 %	55.612	29.316	0	100	2661

Note: Sample is restricted to those who did not receive the second dose of COVID-19 vaccine. *HS*: high school. *Self-reported health*: an indicator for the respondent being in good health. *Chance of severe/moderate symptom*: respondent's belief about the chance that contracting COVID-19 will lead to severe/moderate symptoms. *VI*: Vaccination Intention.

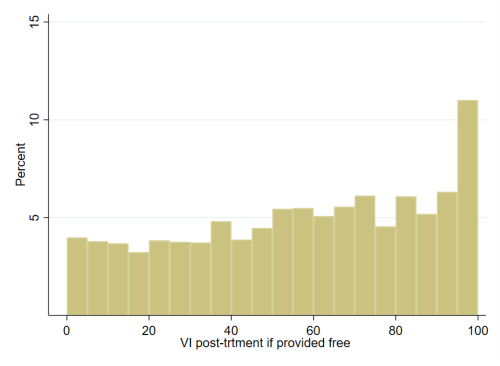
Table 2: Balance Tests

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Age	HS grad	Married	Working	Self-reported health	First dose	Female	Religiosity
Gain	-0.542	0.039	0.020	0.030	0.020	0.003	-0.003	0.104
	(0.629)	(0.028)	(0.024)	(0.022)	(0.022)	(0.021)	(0.025)	(0.118)
Loss	-1.000*	0.036	-0.014	0.024	0.022	0.011	-0.021	0.058
	(0.499)	(0.024)	(0.030)	(0.023)	(0.022)	(0.025)	(0.019)	(0.113)
Constant	35.367***	0.377***	0.509***	0.405***	0.860***	0.293***	0.331***	2.692***
	(0.392)	(0.026)	(0.024)	(0.018)	(0.020)	(0.020)	(0.018)	(0.072)
Observations	2,661	2,661	2,661	2,661	2,661	2,661	2,522	2,517
$H_0: \beta_G = \beta_L$	0.110	0.323	0.422	0.396	0.607	0.898	0.508	0.661
	Chance of severe symptoms	Chance of moderate symptoms	Trust in government	Trust in science	Trust in medicine	Trust in press	Pre-trtment VI	
Gain	-0.739	3.084***	-0.871	-1.174	0.329	0.167	2.029	
	(1.007)	(0.879)	(1.207)	(0.856)	(1.249)	(1.394)	(2.146)	
Loss	0.193	1.523	0.482	0.825	0.674	1.112	2.092	
	(1.468)	(1.108)	(1.444)	(0.916)	(1.350)	(1.717)	(1.700)	
Constant	27.055***	31.700***	63.216***	74.707***	78.199***	63.233***	38.331***	
	(1.182)	(0.911)	(1.242)	(0.811)	(1.045)	(1.310)	(1.647)	
Observations	2,661	2,661	2,661	2,661	2,661	2,661	2,661	
$H_0: \beta_G = \beta_L$	0.578	0.005	0.362	0.019	0.857	0.672	0.465	

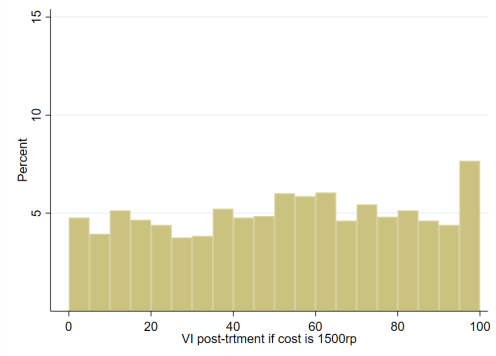
Figure 1: Vaccination Intentions



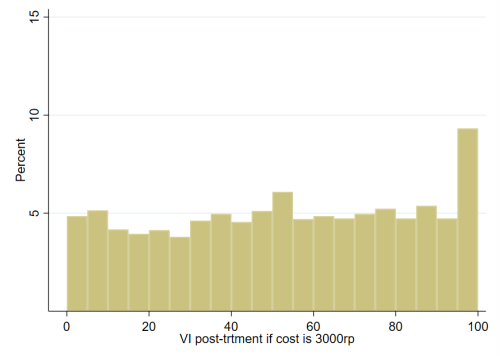
(a) Pre-Treatment



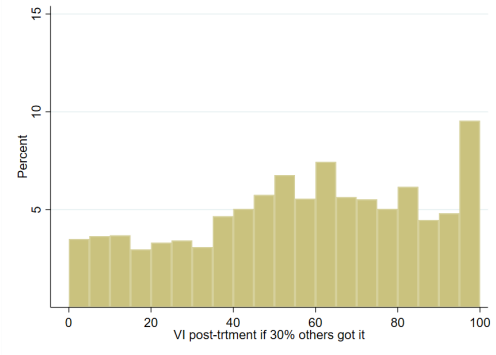
(b) Post-Treatment, if provided for free



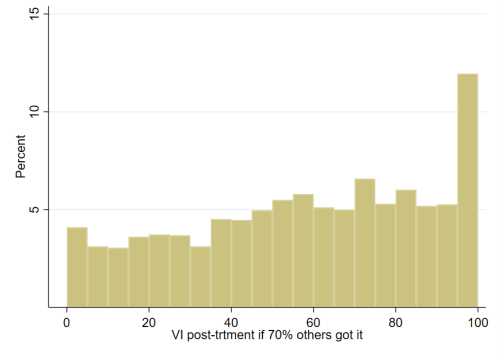
(c) Post-Treatment, if cost was 1500 rp



(d) Post-Treatment, if cost was 3000 rp



(e) Post-Treatment, if 30% got it



(f) Post-Treatment, if 70% got it

2.3 Empirical Model

First, we identify the effects of providing information with any framing using the following equation:

$$y_i = \beta_0 + T_{GL,i}\beta_{GL} + X_i\gamma + \epsilon_i,\ T_{GL,i} = 1 \text{ if } i \text{ is treated.} \tag{1}$$

β_{GL} identifies the effect of providing any message on the vaccination intention of the subject.

Second, we identify the effects of framing the information on the prosocial benefits of vaccination as either gain or loss using:

$$y_i = \beta_0 + \sum_{k\in\mathcal{K}} T_{k,i}\beta_k + X_i\gamma + \epsilon_i,\ \mathcal{K} = \{G,\ L\}, \tag{2}$$

in which T_G is an indicator for GAIN Treatment and T_L is an indicator for LOSS Treatment. β_G and β_L identify the effects of gain- and loss-framed messages, respectively. We conclude that framing effect exists if $\beta_G \neq \beta_L$.

In the appendix, we also estimate the following equation:

$$y_i = \beta_0 + \sum_{k\in\mathcal{K}} T_{k,i}\beta_k + X_i\gamma + \epsilon_i,\ \mathcal{K} = \{g, g+, l, l+\}, \tag{3}$$

where $T_{k,i}$ is an indicator for being in treatment group k for subject i. The difference between β_g and β_{g+} and the difference between β_l and β_{l+} identify the effects of informing subjects about strategic gains as well as the prosocial benefits of getting vaccinated.

In these equations, X_i is a vector of baseline control variables including age, high school graduation status, marital status, work status, self-reported health (indicator for being in "good health"), COVID vaccine first-dose status, belief in the chance of experiencing severe symptoms because of COVID-19, belief in the chance of experiencing moderate symptoms because of COVID-19, and questionnaire ordering group indicators. Trust measures are also not included because they are surveyed after the treatments. Table 2 shows that trust in

institutions are not affected by the treatments, and Section 3.2 examines treatment heterogeneity by trust level. Gender and religious intensity are not included in the main analysis because they have lower response rates. Gender, religiosity, and trust variables are included as controls in the robustness analysis in Section 3.3. Error term is clustered at the level of 36 districts.

3 Results

3.1 Main Results

Table 3 presents the effects of the treatments, showing that providing scientific information on the prosocial benefits of vaccination significantly increases vaccination intentions across different types of measures and interventions. Panel A shows that the overall effects of sending "any" prosocial message (T_G, T_L) are positive for all outcomes and significant for all outcomes except for the intention with 3,000 rupees hypothetical cost. Panel B shows that the loss framing is more effective than gain framing in promoting vaccination intention through prosocial messages. The differences of gain- and loss-framed effects are significant for all outcomes. Even for vaccination intention with 3,000 rupees cost, the effect is positive and significant for loss-framed messages although it is close to zero and insignificant for gain-framed messages.

The magnitude of treatment effect decreases as the hypothetical monetary cost of vaccination increases. The effect of loss-framed message on vaccination intention with 3,000 rupees cost is less than half of the effect on vaccination intention with no hypothetical cost. There is almost no difference between the effects on vaccination intention with 30% neighborhood vaccination rate and 70% neighborhood vaccination rate, although the effects are higher by about 0.5 percentage point for 70% rate in Panel A and Panel B.[11]

[11] A higher proportion of neighbors who got vaccinated may create peer pressure to get vaccinated, while lowering the private and prosocial inventive to get vaccinated. These two effects may cancel each other out in our sample.

Table 3: Effects of Prosocial Messages on Vaccination Intentions

$N = 2,661$	(1) VI free	(2) VI 1500 rp	(3) VI 3000 rp	(4) VI 30 p	(5) VI 70 p
Panel A					
Prosocial Treatment	8.939***	6.370***	2.490	5.780***	6.363***
(β_{GL})	(1.606)	(1.420)	(1.668)	(1.753)	(1.731)
R-squared	0.042	0.047	0.025	0.031	0.031
Panel B					
Prosocial Gain (β_G)	7.871***	5.172***	0.780	4.035**	4.502**
	(1.727)	(1.410)	(1.817)	(1.872)	(1.774)
Prosocial Loss (β_L)	10.015***	7.576***	4.213**	7.539***	8.237***
	(1.716)	(1.731)	(1.747)	(1.818)	(1.857)
R-squared	0.043	0.049	0.028	0.034	0.034
$H_0: \ \beta_G = \beta_L$	0.082	0.089	0.008	0.003	0.001

Note: *** p<0.01, ** p<0.05, * p<0.1. Sample is restricted to those who did not receive the second dose of COVID-19 vaccine. Control variables include age, education level, marital status, work status, self-reported health, pre-treatment vaccination intention, first-dose status, belief in the chance of experiencing severe symptoms because of COVID-19, belief in the chance of experiencing moderate symptoms because of COVID-19, and questionnaire ordering indicators. P-values of hypothesis tests are shown at the last row of Panel B.

3.2 Heterogeneity Analysis

3.2.1 Heterogeneity by Baseline Characteristics

We investigate heterogeneity by individual characteristics associated with vaccine hesitancy. The estimates are based on:

$$y_i = \beta_0 + \sum_{k \in \mathcal{K}} T_{k,i}\beta_{k,1} + \sum_{k \in \mathcal{K}} T_{k,i} \times I_i\{Subgroup\}\beta_{k,2} + X_i\gamma + \epsilon_i \tag{4}$$

where $\mathcal{K} = \{G, L\}$ and $I_i\{Subgroup\}$ is an indicator that equals 1 if subject i belongs to the subgroup of interest. We focus on measures that predict vaccine hesitancy or acceptance in various international surveys (Lazarus et al. 2021; Chaudhary et al. 2021) such as education, age, gender, self-reported health, and belief about the chance of experiencing severe or

moderate symptoms if infected with COVID-19. For continuous measures, the subgroups are divided by sample median.

Figure 2 shows estimates for $\beta_{k,1} + \beta_{k,2}$ and $\beta_{k,1}$ in equation (4) along with their 95% confidence intervals. Across all subgroups, both gain- and loss-framed messages have a significant positive impact on vaccination intentions. Additionally, with rare exceptions, the effects of loss-framed messages are greater than the effects of gain-framed messages. Other patterns are mixed. The effects of gain-framed messages are more likely to be different across subgroups than loss-framed messages.

Figures A1, A2, A3, and A4 in the appendix present heterogeneous results for other vaccination intention outcomes. Subgroup differences appear more pronounced, but they are not consistent across hypothetical conditions. One consistent pattern across all subgroups and hypothetical conditions is that loss-framed treatment is more effective than gain-framed treatment.

3.2.2 Heterogeneity by Trust in Institutions

We also consider measures of trust in four institutions, government, science, medicine, and the press.[12] Trust measures, especially trust in government, are highly predictive of vaccine hesitancy (e.g., Solís Arce et al. 2021). We estimate

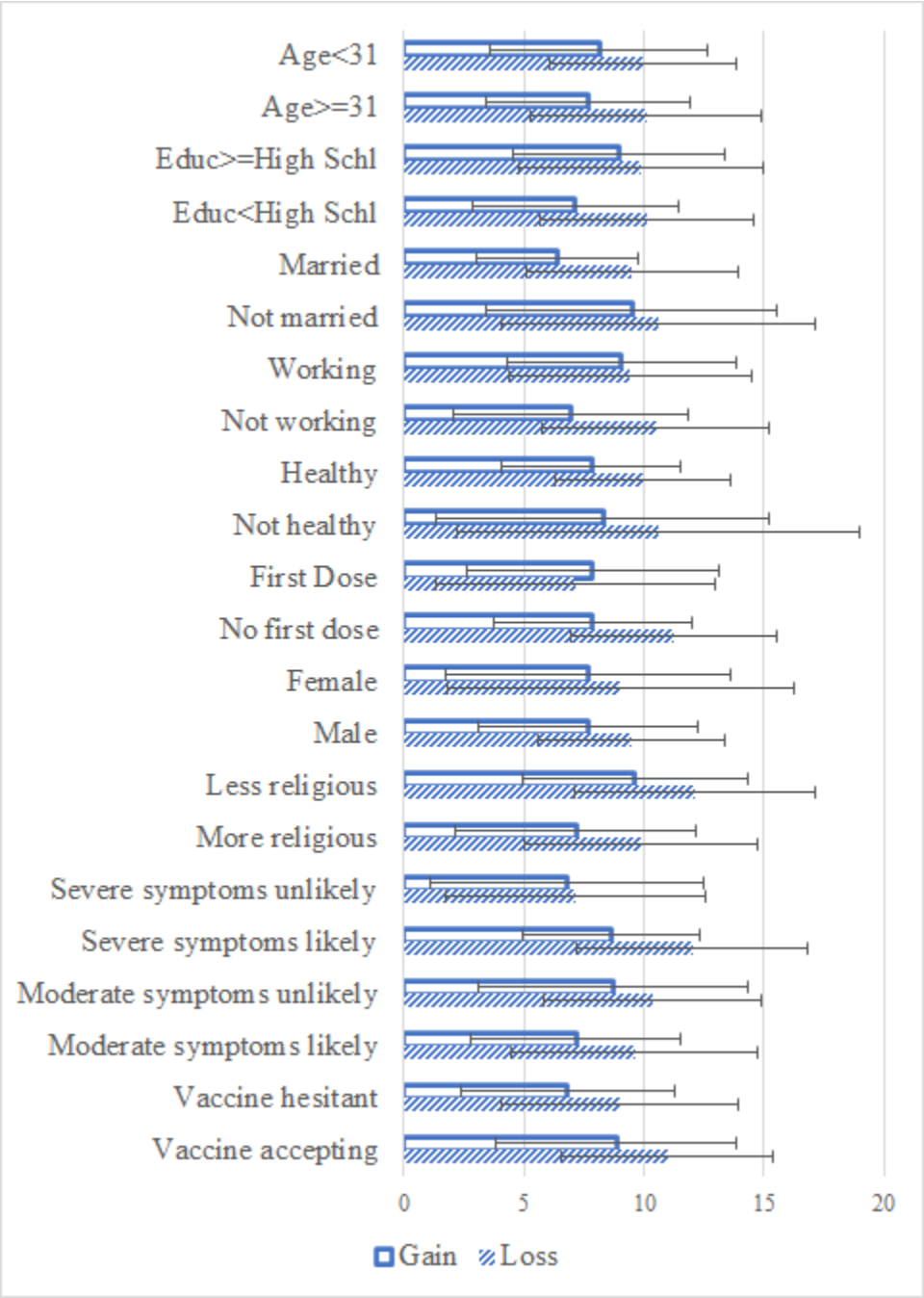
$$Y_i = \beta_0 + \sum_{k \in \mathcal{K}} T_{k,i}\beta_{k,1} + \sum_{k \in \mathcal{K}} T_{k,i} \cdot TrustH_i\beta_{k,2} + X_i\gamma + \epsilon_i, \tag{5}$$

where $\mathcal{K} = \{G, L\}$ and $TrustHi_i$ is a row vector consisting of indicators that subject i has above-sample-median trust in government, science, medicine, and the press, respectively. We also estimate

$$Y_i = \beta_0 + \sum_{k \in \mathcal{K}} T_{k,i}\beta_{k,1} + \sum_{j \in J} \sum_{k \in \mathcal{K}} T_{k,i} \cdot TrustH_j\beta_{k,j,2} + X_i\gamma + \epsilon_i, \tag{6}$$

[12] We examine whether trust in institutions predict vaccination status using the entire sample. Table A5 shows that trust in government and press are predictive of vaccination status.

Figure 2: Effects on VI (free) by Baseline Characteristics



Note: Sample is restricted to those who did not receive the second dose of COVID-19 vaccine. Control variables include age, education level, marital status, work status, self-reported health, pre-treatment vaccination intention, first-dose status, belief in the chance of experiencing severe symptoms because of COVID-19, belief in the chance of experiencing moderate symptoms because of COVID-19, and questionnaire ordering indicators. Bars represent treatment effect estimates based on interaction models. Capped lines represent 95% confidence intervals.

where $J \in \{Government,\ Science,\ Medicine,\ Press\}$. In this equation, $\beta_{k,1}$ represents the effect of treatment k for those who have below-median trust in all four institutions. $\beta_{k,1} + \sum_{j \in J} \beta_{k,j,2}$ represents the effect of treatment k for those with above-median trust in all four institutions.

Figure 3 shows that the effects of information treatments are significant even for those with low trust levels in institutions. Treatment effects are more effective for those with high trust levels, and loss-framed treatments are more effective than gain-framed treatments for all subgroups. An exception is trust in the press: treatment effects are greater for those with low trust in the press.

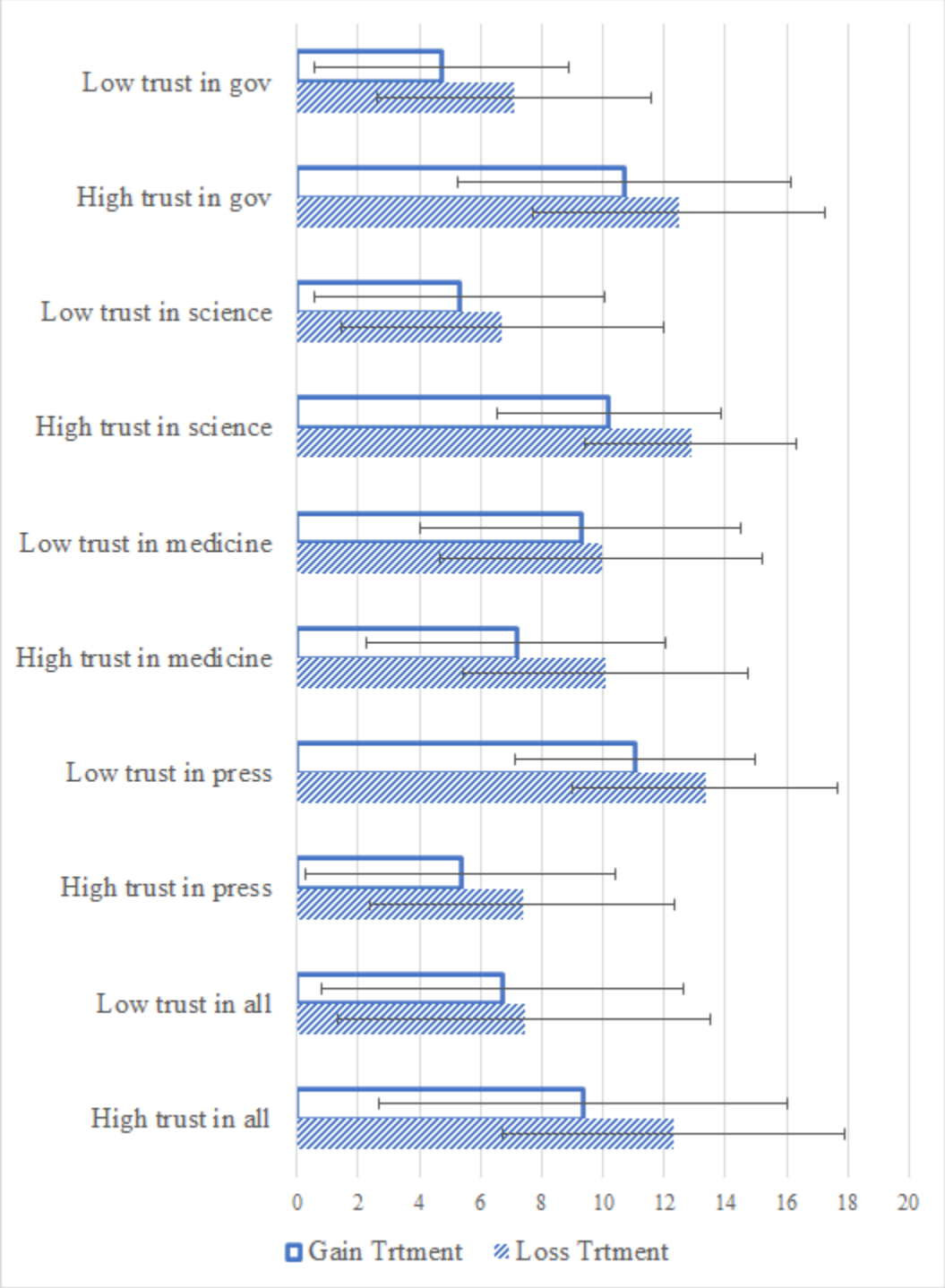
Figures A5, A6, A7, and A8 in the appendix present heterogeneous results for other vaccination intention outcomes. A common finding is that treatment effects tend to be greater for those with high trust in institutions except for the press. Loss-framed treatment effects are greater than gain-framed treatment effects, consistent with the results in Figure 3.

3.3 Robustness

We examine the robustness of the results to different specifications in Table 4. Panels A and B shows that the results are robust to removing all controls and adding additional controls (which are excluded in the main analysis because of item nonresponse). Panel C shows estimates based on a binary outcome variable that equals 1 if vaccination intention is 50 or greater. The estimates are therefore comparable to other studies in the literature that use binary intention variables as outcomes. The estimates, which are based on linear probability models, are even greater than the main results in Table 3, suggesting that the treatments are shifting the extensive margin of choice.

Section C in the Appendix shows quantile treatment effects for the deciles of vaccination intentions. Consistent with the findings in Table 4, the effects are concentrated between the

Figure 3: Effects on VI (free) by Trust Levels



Note: Sample is restricted to those who did not receive the second dose of COVID-19 vaccine. Control variables include age, education level, marital status, work status, self-reported health, pre-treatment vaccination intention, first-dose status, belief in the chance of experiencing severe symptoms because of COVID-19, belief in the chance of experiencing moderate symptoms because of COVID-19, and questionnaire ordering indicators. Bars represent treatment effect estimates based on interaction models. Capped lines represent 95% confidence intervals.

20th and 50th deciles, suggesting that the treatments are effective for those at the margin of vaccination decisions.

In Section B in the Appendix, we analyze the effects of emphasizing strategic concerns in addition to prosocial messages, using equation 3 that incorporates the four treatment arms. The results show that the magnitude of $l+$ Treatment that provides loss-framed messages with strategic concern is the greatest among all treatments, but the differences between $l+$ Treatment (with strategic concerns) and l Treatment (without strategic concerns) are not significant. The differences between $g+$ and g Treatments (gain-framed messages with or without strategic concerns, respectively) are likewise not significant. Emphasizing strategic concerns in addition to gain- or loss-framed messages does not increase the magnitude of the treatment effects.

An important concern regarding the survey experiment based on self-reported intention is the experimenter demand effect (EDE), in which the respondents bias their behaviors in response to what they perceive as the intention of the experimenter (Zizzo 2010; Haaland, Roth and Wohlfart 2021). If the respondents in our study believe that the appropriate behavior in the context of the study is to agree with the surveyor, then the treatment effect estimates would be biased away from zero.

Although we cannot rule out the effects of EDE, we believe that our results are robust to its presence. First, assuming that EDE applies similarly to both gain- and loss-framed messages, the difference between these two effects are not affected by EDE.

Second, we point out that some of our heterogeneity analyses rely on measures of conformity to social norms. This is similar to the idea in Dhar, Jain and Jayachandran (2022) that accounts for the degree of social desirability bias of experiment subjects using measures of social desirability. In our case, trust in institutions, especially science, would be relevant because our messages specifically reference works of scientists. Those who wish to conform to the demands of the experimenter would express confidence in the messages provided. Religiosity would also be a relevant measure of social conformity for Pakistan, where religion is

an important factor in vaccination decisions (Chaudhary et al. 2021). Our analyses in Figures 2 and 3 show that the treatments are effective regardless of trust levels and religiosity, although treatment effects are greater among those with greater trust in science.

Finally, we examine treatment effects on mask-wearing intentions, for which the treatment provides no information. If the respondents, prompted by the treatment, interpret the experimenter's demand as promoting pandemic social guidelines, then we might expect positive treatment on mask wearing intentions as well. We observe no treatment effects on this outcome (Table 5).

Table 4: Robustness

	(1) VI free	(2) VI 1500 rp	(3) VI 3000 rp	(4) VI 30 p	(5) VI 70 p	N	Controls	Outcome
Panel A								
Gain	7.543***	4.685***	0.340	3.700*	4.005**	2,661	none	0–100
	(1.710)	(1.455)	(1.815)	(1.843)	(1.819)			
Loss	9.939***	7.524***	3.936**	7.446***	7.994***			
	(1.686)	(1.712)	(1.742)	(1.834)	(1.912)			
R-squared	0.015	0.009	0.004	0.010	0.010			
$H_0: \ \beta_G = \beta_L$	0.064	0.049	0.006	0.002	0.001			
Panel B								
Gain	7.949***	5.380***	1.238	3.363*	4.811***	2,381	baseline+	0–100
	(1.815)	(1.390)	(2.031)	(1.909)	(1.593)			
Loss	10.036***	7.272***	4.205**	7.359***	8.424***			
	(1.635)	(1.870)	(1.899)	(1.856)	(1.789)			
R-squared	0.045	0.054	0.032	0.036	0.040			
$H_0: \ \beta_G = \beta_L$	0.103	0.207	0.035	0.001	0.003			
Panel C								
Gain	0.130***	0.091***	0.030	0.032	0.083***	2,661	baseline	0,1
	(0.029)	(0.026)	(0.029)	(0.027)	(0.030)			
Loss	0.171***	0.132***	0.088***	0.088***	0.149***			
	(0.029)	(0.030)	(0.029)	(0.027)	(0.032)			
R-squared	0.043	0.037	0.026	0.028	0.031			
$H_0: \ \beta_G = \beta_L$	0.048	0.081	0.015	0.008	0.002			

Note: *** p<0.01, ** p<0.05, * p<0.1. Sample is restricted to those who did not receive the second dose of COVID-19 vaccine. Baseline control variables include age, education level, marital status, work status, self-reported health, pre-treatment vaccination intention, first-dose status, belief in the chance of experiencing severe symptoms because of COVID-19, belief in the chance of experiencing moderate symptoms because of COVID-19, and questionnaire ordering indicators. Baseline+ control variables additionally include indicators for female gender, "other gender", religiosity, and trust in government, science, medicine, and press. For Panel C, outcome variable is the binary indicator that equals 1 if $VI \geq 50$ and 0 otherwise. P-values of hypothesis tests are shown at the last row of each panel.

Table 5: Effects on Intention to Wear Mask

	Intention to wear masks
Prosocial Gain (β_G)	-0.986
	(1.747)
Prosocial Loss (β_L)	-1.707
	(1.541)
Age	-0.013
	(0.056)
HS grad	4.895***
	(1.232)
Married	2.778**
	(1.284)
Working	-0.856
	(1.078)
Self-reported health	1.938
	(1.883)
Pr-trtment VI	0.052***
	(0.016)
First dose	1.053
	(1.072)
Chance of moderate symptoms	0.063***
	(0.018)
Chance of severe symptoms	-0.013
	(0.025)
Constant	65.063***
	(3.786)
Observations	2,661
R-squared	0.016

Note: Sample is restricted to those who did not receive the second dose of COVID-19 vaccine. Control variables also include questionnaire ordering indicators.

4 The Model

In this section, we formalize the vaccination decisions occurring in the controlled settings of our experiment. There is a set of $\{1, \dots N\}$ subjects. We denote $v_i \in \{0, 1\}$ a dummy equal to 1 if subject i is vaccinated and 0 otherwise. Finally, we denote $p(v_i, v_j) \in [0, 1]$, the probability that subject i with vaccination status v_i infects subject j with vaccination status v_j.

We focus this section on a case where the agents are subject to one of two treatments.[13] We denote $t_i \in \{T_c, T_G, T_L\}$ the treatment group of subject i. Treatment T_c is the control group. Treatment T_G emphasizes the gains from getting vaccinated by comparing a situation where the subject is vaccinated to a situation where he is not. Treatment T_L emphasizes the loss from not getting vaccinated by comparing a situation where the subject is not vaccinated to a situation where he is. The treatments have two effects. First, they focus subjects' attention on the prosocial gains or losses associated with their vaccination decision. Second, the treatments make subjects envision a situation where they meet another person, whose vaccination status is unknown.

Utility: We specify a person i's utility function as $U_i(v_i \mid r(t_i))$, where $v_i \in \{0, 1\}$ indicates i's vaccination status, $t_i \in \{T_c, T_G, T_L\}$ her treatment group, and $r(t_i)$ the reference point suggested by treatment t_i. Indeed, treatment G emphasizes the gain associated with the vaccination compared to a situation where the subject is not vaccinated. Hence, treatment T_G suggests the reference point $r(t_i) = 0$ to subject i when she reports her intention to get vaccinated. Similarly, a subject i that belongs to treatment group T_L would feel a loss when envisioning the decision $v_i = 0$ compared to the decision $v_i = 1$. By emphasizing the loss associated with not getting vaccinated compared to getting vaccinated, treatment 2 suggests the reference point $r(t_i) = 1$ to subject i when she reports her vaccination intention.

In this model, we study how a subject's valuation for vaccination is endogenously determined by her prosocial preferences, standard health factors, and her treatment group.

[13] An extension to the four treatment arms is provided in Section B.1.

Suppose that subject i's utility takes the following form:

$$U_i(v_i \mid r(t_i)) = m_i(v_i) + n(v_i \mid r(t_i)), \tag{7}$$

where $m_i(v_i)$ is subject i's utility when her vaccination status is v_i and $n(v_i \mid r(t_i))$ is the "gain-loss" utility induced by the treatment. We assume that principally, an agent's utility m_i depends on three factors: (i) the health cost of getting infected $c > 0$, (ii) the prosocial cost of infecting someone $c_{ps} > 0$, and (iii) the idiosyncratic cost of getting vaccinated $\epsilon_i \in \mathbb{R}$. We can then express $m_i(v_i)$ as follows:

$$m_i(v_i) = -cp(v_j, v_i) - c_{ps}p(v_i, v_j) - v_i\epsilon_i. \tag{8}$$

With probability $p(v_j, v_i)$, j transmits the disease to i, who incurs a cost c. With probability $p(v_i, v_j)$, i transmits the disease to j and i incurs an prosocial cost c_{ps}. Finally, if $v_i = 1$, then subject i pays the vaccination cost ϵ_i.

We make the following assumption on the values of $p(v_i, v_j)$:

Assumption 1 *The vaccine protects against the virus* $p(v_j, 1) < p(v_j, 0)$ *for any* $v_j \in \{0, 1\}$, *and reduces the probability of transmitting,* $p(1, v_j) < p(0, v_j)$ *for any* $v_j \in \{0, 1\}$.

The first part of Assumption 1 states that the vaccine reduces the chance of getting sick. The second part of Assumption 1 states that the vaccine reduces the probability of transmitting the disease. This assumption accords with the existing evidence (e.g., Harris et al. (2021)), and is emphasized through our information treatments.

We assume that how a person feels about the prosocial gain or loss implied by vaccination depends on the changes in utility associated with such gains or losses. Because the treatments focus subjects' attention on the prosocial gains or losses associated with the vaccine, we express $n(v_i \mid r(t_i))$ as:

$$n(v_i \mid r(t_i)) = \mu[-c_{ps}p(v_i, v_j) + c_{ps}p(r(t_i), v_j)], \tag{9}$$

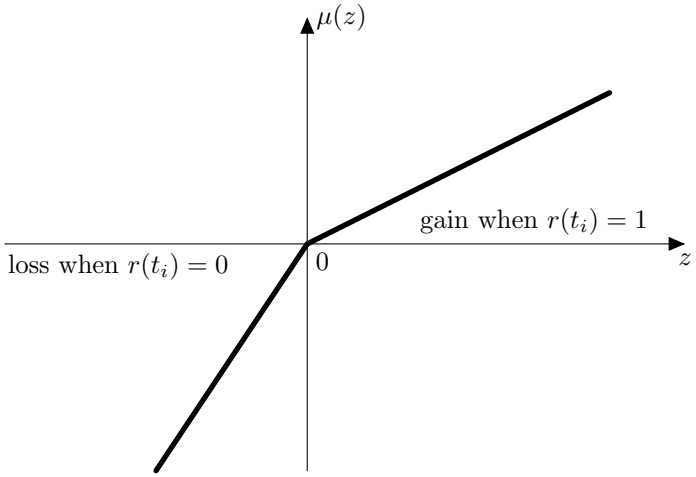


Figure 4: Gain-loss function $\mu(.)$

where $\mu(.)$ takes the following form:

$$\mu(z) = \begin{cases} \eta z \text{ for } z > 0, \text{ and} \\ \eta\lambda z \text{ for } z \leq 0, \end{cases} \tag{10}$$

with $\eta \geq 0$ a parameter that measures the degree to which the two treatments focus subjects' attention on the prosocial gain-loss dimension of their decision, while $\lambda \geq 1$ is a loss aversion coefficient.

The gain-loss function $\mu(.)$ is represented in Figure 4. Under gain framing, the gain-loss function $\mu(.)$ is positive. The subject evaluates the prosocial utility of getting vaccinated $(-c_{ps}p(v_i = 1, v_j))$ against the reference point of not getting vaccinated $(-c_{ps}p(r(t_i) = 0, v_j))$. The subject perceives a gain when getting vaccinated relative to not getting vaccinated, as $p(1, v_j) < p(0, v_j)$. Under loss framing, the gain-loss function $\mu(.)$ is negative. The subject compares the value of not getting vaccinated $(-c_{ps}p(v_i = 0, v_j))$ against that of getting vaccinated $(-c_{ps}p(r(t_i) = 1, v_j))$. The subject perceives a loss when not getting vaccinated, as $p(0, v_j) > p(1, v_j)$.

Through (9) and (10), we assume that the gains or losses suggested by the treatments affect subjects' utility. The specification of the gain-loss utility associated with the treatment

closely follows Kőszegi and Rabin (2006). Through parameter $\lambda \geq 1$, we posit that subjects might be loss averse. Feeling a loss from infecting someone else when not vaccinated might have a stronger effect on a subject's utility than feeling a gain from not infecting someone else when vaccinated.

When subject i belongs to treatment T_L, we can rewrite her utility as:

$$\begin{cases} U_i(1 \mid 1) = -\epsilon_i - cp(v_j, 1) - c_{ps}p(1, v_j) \\ U_i(0 \mid 1) = -cp(v_j, 0) - c_{ps}p(0, v_j) + \eta\lambda c_{ps}[-p(0, v_j) + p(1, v_j)] \end{cases} \tag{11}$$

The first line is the utility derived by subject i when she gets vaccinated. The second line is the utility derived by subject i when she does not get vaccinated. When the subject does not get vaccinated and belongs to treatment T_L, she incurs an additional loss given that the treatment focuses her attention on the higher likelihood of infecting someone in a situation where she does not get vaccinated compared to a situation where she gets vaccinated, given that $p(1, v_j) < p(0, v_j)$ under Assumption 1.

Similarly, when subject i belongs to treatment T_G, her utility can be expressed as:

$$\begin{cases} U_i(1 \mid 0) = -\epsilon_i - cp(v_j, 1) - c_{ps}p(1, v_j) + \eta c_{ps}[-p(1, v_j) + p(0, v_j)] \\ U_i(0 \mid 0) = -cp(v_j, 0) - c_{ps}p(0, v_j) \end{cases} \tag{12}$$

The first line is the utility from getting vaccinated. When the subject gets vaccinated and belongs to treatment T_G, she perceives an additional gain because the treatment focuses her attention on the lower likelihood of infecting someone in a situation where she gets vaccinated compared to a situation where she does not get vaccinated, given that $p(1, v_j) < p(0, v_j)$ under Assumption 1.

We assume that the cost of the vaccine ϵ_i can be divided into two components:

$$\epsilon_i = \pi + u_i, \tag{13}$$

with $\pi \geq 0$ the monetary cost of the vaccine and $u_i \in \mathbb{R}$ the psychological cost of the vaccine, which we assume is drawn from a well-behaved c.d.f $F(.)$.

Finally, given that subject i does not know the vaccination status of the agent j with whom she is matched, she expects j to be vaccinated with probability $x \in [0,1]$, with x the prevailing fraction of vaccinated individuals. Subject i gets vaccinated when

$$\mathbb{E}\, U_i(1 \mid r(t_i)) \geq \mathbb{E}\, U_i(0 \mid r(t_i)), \tag{14}$$

given that subject i is uncertain about the vaccination status of agent j, so $\mathbb{E}\, p(v_i, v_j) = xp(v_i, 1) + (1-x)p(v_i, 0)$ for any $v_i \in \{0, 1\}$. Substituting (11) and (12) in the last inequality, we deduce that subject i with $r(t_i) = 1$ gets vaccinated when

$$\epsilon_i \leq c\, \mathbb{E}[p(v_j, 0) - p(v_j, 1)] + c_{ps}(1 + \eta\lambda)\, \mathbb{E}[p(0, v_j) - p(1, v_j)], \tag{15}$$

while a subject i with $r(t_i) = 0$ gets vaccinated when

$$\epsilon_i \leq c\, \mathbb{E}[p(v_j, 0) - p(v_j, 1)] + c_{ps}(1 + \eta)\, \mathbb{E}[p(0, v_j) - p(1, v_j)], \tag{16}$$

with $\mathbb{E}\, p(v_i, v_j) = xp(v_i, 1) + (1 - x)p(v_i, 0)$ for any $v_i \in \{0, 1\}$.

Substituting ϵ_i with equation (13), we deduce the following result:

Proposition 1 *The likelihood $q_i(t_i) \in [0, 1]$ that subject i gets vaccinated is*

$$q_i(t_i) = F\left(-\pi + c\, \mathbb{E}[p(v_j, 0) - p(v_j, 1)] + c_{ps}(1 + \eta(1 + (\lambda - 1)r(t_i)))\, \mathbb{E}[p(0, v_j) - p(1, v_j)]\right). \tag{17}$$

From this result, we can deduce several testable predictions.

Prediction 1 *$q_i(T_L) > q_i(T_G)$ if and only if $\lambda > 1$.*

As treatment T_L suggests a reference point $r(t_i) = 1$ to subject i, it makes her feel a loss in a situation where she does not get vaccinated and potentially transmits the disease compared to a situation where she gets vaccinated. Hence, to the extent that subject i is loss averse and $\lambda > 1$, feeling this loss will increase the likelihood of her getting vaccinated compared to a situation where she belongs to treatment T_G.

Subjects' attention in the control group is not focused on the prosocial dimension of their vaccination decision. Hence, these subjects neither feel an prosocial gain nor an prosocial loss when they answer the survey. For subject i in the control group, the "gain-loss" utility induced by the treatment is equal to zero, so she gets vaccinated when

$$m_i(1) > m_i(0), \tag{18}$$

which implies that

$$q_i(T_c) = F(-\pi + c\,\mathbb{E}[p(v_j, 0) - p(v_j, 1)] + c_{ps}\,\mathbb{E}[p(0, v_j) - p(1, v_j)]) \tag{19}$$

Prediction 2 *$q_i(T_k) > q_i(T_c)$ for any $k \in \{1, 2\}$.*

This prediction directly follows from the assumption that subjects' attention in the control group T_c is not focused on the prosocial dimension of their decision.

Prediction 3

- *$q_i(T_k)$ increases with η, c, c_{ps}, $p(0,0)$, and decreases with $p(1,1)$, π for any $k \in \{1, 2\}$. $q_i(T_L)$ increases with λ while $q_i(T_G)$ is not affected by λ.*
- *$q_i(T_k)$ decreases with x if and only if $p(0,1) - p(1,1) \leq p(0,0) - p(1,0)$ for any $k \in \{1, 2\}$.*

This prediction summarizes the comparative statics. When treatments focus more subjects' attention on the prosocial dimension of their decision (i.e., when η increases), subjects

are more likely to report a higher vaccination intention. Similarly, when either the cost of getting sick or the prosocial cost of infecting someone increases, then $q_i(T_k)$ increases too. This explains why $q_i(T_k)$ increases with c, c_{ps}. When the probability increases that an unvaccinated subject transmits to someone that is not vaccinated $p(0,0)$, then both the personal health cost of not getting vaccinated and the prosocial cost of not getting vaccinated increase. This explains why $q_i(T_k)$ increases with $p(0,0)$. In contrast, when $p(1,1)$ increases, then $q_i(T_k)$ decreases. Only $q_i(T_L)$ is affected by the loss aversion parameter $\lambda \geq 1$ because only treatment T_L suggests a reference point where subjects perceive a loss from not getting vaccinated. When the monetary cost of the vaccine π increases, subjects are less likely to get vaccinated.

How the fraction of vaccinated individuals $x \in [0,1]$ affects $q_i(T_k)$ depends on how the likelihood of catching the disease or of infecting others is affected by others' vaccination status. When the vaccine is more effective in reducing the probability of transmitting to someone who is not vaccinated, the inequality $p(0,1) - p(1,1) \leq p(0,0) - p(1,0)$ is satisfied. In that case, when there are more vaccinated individuals, the prosocial benefits of getting vaccinated are lower. Finally, observe that the the inequality $p(0,1) - p(1,1) \leq p(0,0) - p(1,0)$ is equivalent to the inequality $p(1,0) - p(1,1) \leq p(0,0) - p(0,1)$. This second inequality means that the vaccine is more effective in reducing the probability of getting sick when others are not vaccinated. Hence, when $p(1,0) - p(1,1) \leq p(0,0) - p(0,1)$ is satisfied and when there are more vaccinated individuals, the personal health benefits of getting vaccinated are lower too.

4.1 From the Model to the Data:

In this subsection, we map the model to the data. We focus on the case with two treatments T_G and T_L and a control group T_c for simplicity. The more complete case with the four treatment arms is discussed at the end of this section.

When subject i answers to the survey, she discloses $q_i(t_i)$, the likelihood of her getting vaccinated given her treatment group $t_i \in \{T_c, T_G, T_L\}$. That is, when answering the survey, subjects are aware of the distribution $F(.)$ of their own psychological cost u_i but they are unaware of the realization of u_i.

Let D_i be a dummy variable equal to one if subject i belongs to treatment T_G, and 0 if subject i belongs to the control group. Similarly, we denote B_i a dummy variable equal to 1 if subject i belongs to treatment T_L and zero if subject i belongs to treatment T_G.

We can express $q_i(T_G)$ and $q_i(T_L)$ as:

$$q_i(T_G) = q_i(T_c) + (q_i(T_G) - q_i(T_c))D_i + z_i, \tag{20}$$

and

$$q_i(T_L) = q_i(T_G) + (q_i(T_L) - q_i(T_G))B_i + e_i, \tag{21}$$

where z_i and e_i are independent random variables.

Let $q_i \in [0, 1]$ denote the vaccination intention reported by subject i. We can rewrite (20) and (21) as

$$\begin{cases} q_i = \alpha_1 + \beta_1 D_i + u_i \\ q_i = \alpha_2 + \beta_2 B_i + e_i \end{cases} \tag{22}$$

and given the random assignment of the subjects across the two treatments,

$$\beta_1 = E(q_i \mid D_i = 1) - E(q_i \mid D_i = 0) \tag{23}$$

gives an estimate of $q_i(T_G) - q_i(T_c)$, and

$$\beta_2 = E(q_i \mid B_i = 1) - E(q_i \mid B_i = 0) \tag{24}$$

gives an estimate of $q_i(T_L) - q_i(T_G)$.

We can now interpret the empirical results in light of the theoretical predictions.

- **Loss aversion in prosocial preferences:** The estimates of β_G and β_L are provided in the Panel B in Table 3. From the coefficient estimates, it is clear that $\beta_L > \beta_G$ across all specifications. According to Prediction 1, these empirical results imply that $\lambda > 1$ necessarily. Subjects are loss averse when it comes to their prosocial preferences.

- **Prosocial preferences affects vaccination intention:** The estimates of both β_G and β_L are positive and significant across all specifications. According to Prediction 2, this results imply that (i) the treatments are effective in focusing subjects' attention on the prosocial dimension of vaccination, and (ii) prosocial preferences matter in explaining vaccination decisions.

- **Comparative statics.** We find that the treatment effects are lower both when the cost of vaccination is higher, and when the fraction of vaccinated is higher. These results are consistent with Prediction 3. First, we find that $q_i(T_k)$ decreases with π, the monetary cost of the vaccine for any $k \in \{1, 2\}$. Second, we demonstrate that $q_i(T_k)$ decreases with $x \in [0, 1]$ the fraction of vaccinated individuals, if and only if $p(0, 1) - p(1, 1) \leq p(0, 0) - p(1, 0)$. Our empirical results are then consistent with Prediction 3 as long as subjects perceive that being vaccinated will primarily reduce the likelihood of transmitting to unvaccinated individuals.

- **Prosocial concerns vs social interactions concerns.** From Table A4, we find that the coefficient estimates for the effects of treatments T_g and T_{g+} are not statistically different. Similarly, the coefficient estimates for the effects of treatments T_l and T_{l+} are not statistically significant. Hence, focusing subjects' attention on their social interaction concerns does not change their vaccination intention. According to Proposition 2, these results imply the "social interaction" cost of infecting others c_{SI} is not statistically different from zero. Altruistic concerns dominate social interaction concerns in explaining vaccination intention.

4.2 Estimating the loss aversion parameter λ

We assume that $F(.)$ is the c.d.f of a uniform distribution on $[0,1]$. Provided that the parameter values are calibrated so that $q_i(T_k) \in (0,1)$, with $q_i(T_k)$ characterized in Proposition 1[14], the coefficient β_G of Panel B Table 3 is equal to the difference from Equation (17) (with $r(T_G) = 0$) in Proposition 1 to Equation (19). Similarly, β_L is equal to the difference from Equation (17) (with $r(T_L) = 1$) in Proposition 1 to Equation (19). We deduce that

$$\begin{cases} \beta_L = \eta\lambda c_{ps}\,\mathbb{E}[p(0,v_j) - p(1,v_j)] \\ \beta_G = \eta c_{ps}\,\mathbb{E}[p(0,v_j) - p(1,v_j)], \end{cases} \tag{25}$$

from which we get an estimate of λ without further assumptions on the parameter values:

$$\frac{\beta_L}{\beta_G} = \lambda \tag{26}$$

The coefficient estimates for β_G and β_L are in Panel B Table 3. We find that $\lambda \in [1.27, 1.87]$ across the specifications for which coefficient estimates for both β_L and β_G are statistically different from zero. Compared to the existing literature, our results suggest that loss aversion in altruistic preferences is in line with existing estimates of loss aversion in non-altruistic preferences (Brown et al. (2021)).

5 Conclusion

In this paper, we study the prosocial motives behind vaccination intentions to examine whether prosocial messages increase vaccination intentions, and whether gain-loss framing of prosocial benefits affects people's vaccination decisions. To this end, we conduct a randomized field study in Pakistan whereby treatments consisted in providing subjects with

[14] Solutions are interior when the parameter values are such that $-\pi + c\,\mathbb{E}[p(v_j,0) - p(v_j,1)] + c_{ps}(1 + \eta(1 + (\lambda - 1)r(t_i)))\,\mathbb{E}[p(0,v_j) - p(1,v_j)] \in (0,1)$.

scientific information about the effect of receiving COVID-19 vaccines on the probability of transmitting the disease to others. The messages are framed as the benefit to others if vaccinated or the loss to others if not vaccinated.

We find that informing subjects of prosocial benefits of vaccination increases vaccination intentions by 5 to 10 percentage point regardless of framing. Moreover, we show that subjects are loss averse to prosocial benefits: the treatment emphasizing the higher chance of transmitting COVID-19 to others when not vaccinated has significantly higher effects by 2 to 4 percentage point than the treatment emphasizing the lower chance of transmitting COVID-19 to others when vaccinated. This result is found across numerous baseline characteristics associated with vaccination intentions and trust in four types of major societal institutions, government, science, medicine, and the press.

Motivated by seminal works in linguistics, we explain these findings with a theoretical model where experimental treatments with the same informational content can have vastly different effects because they make subjects set different mental reference points. Accordingly, our two main treatments have the same information content: vaccination reduces the risk of transmitting COVID-19 to others. However, our first treatment focuses subjects' attention on a reference point where they are not vaccinated by emphasizing the gain associated with getting vaccinated relative to not getting vaccinated. Our second treatment focuses subjects' attention on a reference point where they are vaccinated by emphasizing the loss associated with not getting vaccinated relative to getting vaccinated.

We then demonstrated that our experimental results are consistent with the existence of loss aversion in prosocial preferences. Thanks to our model and experiment, we estimate the loss aversion parameter associated with prosocial motives in vaccination intentions. We find that subjects' loss aversion is within the range of estimates on loss aversion in the context of non-prosocial preferences found in the economic literature.

An interesting follow-up research would be to pair the messaging treatment with actual vaccine provision to examine the effects of different framing of messages.[15] Our messaging treatment can be interpreted as a single instance of an advertising campaign that takes place within a space of a 20 minutes-long phone survey. Actual implementation may be benefited by increasing the frequency of messages or placing them near vaccination sites while accounting for the potential negative side effects on mental well-being because of excessive negative messages (Sasaki, Saito and Ohtake 2022). Finally, prosocial motive is one of the fundamental drivers of human decision making. Whereas our study focuses on the COVID-19 pandemic setting, future work may examine loss aversion to prosocial motives in the context of a variety of other social behaviors including charity, bequest, family interactions, and political engagement.

References

Allen IV, James, Arlete Mahumane, James Riddell IV, Tanya Rosenblat, Dean Yang and Hang Yu. 2021. Correcting perceived social distancing norms to combat COVID-19. NBER Working Paper 28651 National Bureau of Economic Research.

Altay, Sacha, Anne-Sophie Hacquin, Coralie Chevallier and Hugo Mercier. 2021. "Information delivered by a chatbot has a positive impact on COVID-19 vaccines attitudes and intentions." *Journal of Experimental Psychology: Applied* .

Arnauld, A. and C. Lancelot. 1810. *Grammaire générale et raisonnée de Port-Royal*. Bossange et Masson.

Batteux, Eleonore, Freya Mills, Leah Ffion Jones, Charles Symons and Dale Weston. 2022. "The effectiveness of interventions for increasing COVID-19 vaccine uptake: A systematic review." *Vaccines* 10(3):386.

[15] We do not have data on the actual vaccination decision after the information treatment. This limitation is shared by the current literature on randomized interventions targeting COVID-19 vaccination intentions (Batteux et al. 2022).

Bénabou, Roland and Jean Tirole. 2003. "Intrinsic and Extrinsic Motivation." *Review of Economic Studies* 70(3):489–520.

Bloom, David E, Michael Kuhn and Klaus Prettner. 2022. "Modern infectious diseases: macroeconomic impacts and policy responses." *Journal of Economic Literature* 60(1):85–131.

Böhm, Robert and Cornelia Betsch. 2022. "Prosocial vaccination." *Current Opinion in Psychology* 43:307–311.

Brown, Alexander L., Taisuke Imai, Ferdinand M. Vieider and Colin Camerer. 2021. Meta-Analysis of Empirical Estimates of Loss-Aversion. CESifo Working Paper Series 8848 CESifo.

Bursztyn, Leonardo, Alessandra L. González and David Yanagizawa-Drott. 2020. "Misperceived Social Norms: Women Working Outside the Home in Saudi Arabia." *American Economic Review* 110(10):2997–3029.

Chaudhary, Farooq Ahmad, Basaruddin Ahmad, Muhammad Danial Khalid, Ayesha Fazal, Muhammad Mohsin Javaid and Danial Qasim Butt. 2021. "Factors influencing COVID-19 vaccine hesitancy and acceptance among the Pakistani population." *Human Vaccines & Immunotherapeutics* 17(10):3365–3370. PMID: 34236952.
URL: *https://doi.org/10.1080/21645515.2021.1944743*

Chen, Tianen, Minhao Dai, Shilin Xia and Yu Zhou. 2021. "Do messages matter? Investigating the combined effects of framing, outcome uncertainty, and number format on COVID-19 vaccination attitudes and intention." *Health Communication* pp. 1–8.

Chomsky, N. 2009. *Cartesian Linguistics: A Chapter in the History of Rationalist Thought*. Cambridge University Press.

Dabla-Norris, Era, Hibah Khan, Frederico Lima and Alexandre Sollaci. 2021. Who Doesn't Want to be Vaccinated? Determinants of Vaccine Hesitancy During COVID-19. IMF Working Paper 130.

Dai, Hengchen, Silvia Saccardo, Maria A Han, Lily Roh, Naveen Raja, Sitaram Vangala, Hardikkumar Modi, Shital Pandya and Daniel M Croymans. 2021. "Behavioral nudges increase COVID-19 vaccinations: Two randomized controlled trials." *medRxiv* .

Dhar, Diva, Tarun Jain and Seema Jayachandran. 2022. "Reshaping Adolescents' Gender Attitudes: Evidence from a School-Based Experiment in India." *American Economic Review* 112(3):899–927.

Dubé, Eve, Caroline Laberge, Maryse Guay, Paul Bramadat, Réal Roy and Julie A Bettinger. 2013. "Vaccine hesitancy: An overview." *Human Vaccines & Immunotherapeutics* 9(8):1763–1773.

Dubé, Ève, Jeremy K Ward, Pierre Verger and Noni E MacDonald. 2021. "Vaccine hesitancy, acceptance, and anti-vaccination: trends and future prospects for public health." *Annual Review of Public Health* 42(1):175–91.

Dubois, Pierre and Tuba Tunçel. 2021. "Identifying the effects of scientific information and recommendations on physicians' prescribing behavior." *Journal of Health Economics* 78:102461.

Fatas, Enrique and Lina Restrepo-Plaza. 2022. "When losses can be a gain. A large lab-in-the-field experiment on reference dependent forgiveness in Colombia." *Journal of Economic Psychology* 88:102463.

Feng, Chunliang, Yijie Zhang, Zhixin Zhang and Jie Yuan. 2021. "Prosocial Gains and Losses: Modulations of Human Social Decision-Making by Loss-Gain Context." *Frontiers in Psychology* 12.

Fiedler, Susann and Adrian Hillenbrand. 2020. "Gain-loss framing in interdependent choice." *Games and Economic Behavior* 121:232–251.

Haaland, Ingar, Christopher Roth and Johannes Wohlfart. 2021. Designing Information Provision Experiments. CESifo Working Paper 8406.

Harris, Ross J., Jennifer A. Hall, Asad Zaidi, Nick J. Andrews, J. Kevin Dunbar and Gavin Dabrera. 2021. "Effect of Vaccination on Household Transmission of SARS-CoV-2 in England." *New England Journal of Medicine* 385(8):759–760.

Kachurka, Raman, Michał Krawczyk and Joanna Rachubik. 2021. "Persuasive Messages Will Not Increase COVID-19 Vaccine Acceptance: Evidence from a Nationwide Online Experiment." *Vaccines* 9(10):1113.

Kahneman, Daniel and Amos Tversky. 1979. "Prospect Theory: An Analysis of Decision under Risk." *Econometrica* 47(2):263–291.

Kerr, John R, Alexandra LJ Freeman, Theresa M Marteau and Sander van der Linden. 2021. "Effect of information about COVID-19 vaccine effectiveness and side effects on behavioural intentions: two online experiments." *Vaccines* 9(4):379.

Kimbrough, Erik O., David Porter and Mark Schneider. 2021. "Reference dependent prices in bargaining: An experimental examination of precise first offers." *Journal of Economic Psychology* 86(C).

Kőszegi, Botond and Matthew Rabin. 2006. "A Model of Reference-Dependent Preferences." *The Quarterly Journal of Economics* 121(4):1133–1165.

Larson, Heidi J. 2016. "Vaccine trust and the limits of information." *Science* 353(6305):1207–1208.

Lazarus, Jeffrey V, Scott C Ratzan, Adam Palayew, Lawrence O Gostin, Heidi J Larson, Kenneth Rabin, Spencer Kimball and Ayman El-Mohandes. 2021. "A global survey of potential acceptance of a COVID-19 vaccine." *Nature Medicine* 27(2):225–228.

Manski, Charles. 2004. "Measuring Expectations." *Econometrica* pp. 1329–1376.

Mehmood, Sultan, Shaheen Naseer and Daniel Chen. 2021. Training Effective Altruism. Unpublished manuscript.

Motta, Matt, Steven Sylvester, Timothy Callaghan and Kristin Lunz-Trujillo. 2021. "Encouraging COVID-19 vaccine uptake through effective health communication." *Frontiers in Political Science* 3:1.

Our World in Data. 2022. "Share of people who completed the initial COVID-19 vaccination protocol.". Published online at OurWorldInData.org. Retrieved from: 'https://ourworldindata.org/explorers/coronavirus-data-explorer' [Online Resource].

Pakistan Telecommunication Authority. 2022. "Telecom Indicators.". Published online at pta.gov.pk. Retrieved from: 'https://www.pta.gov.pk/en/telecom-indicators/2#annual-teledensity-subscribers' [Online Resource].

Parsons, Jonathan. 2021. "Getting a COVID-19 vaccine isn't just about you — it protects others in your life.". Published online at the Health News section of the Ohio State Wexner Medical Center website. Retrieved from: 'https://wexnermedical.osu.edu/blog/covid-19-vaccine-protects-others-too' [Online Resource].

Pfattheicher, Stefan, Michael Bang Petersen and Robert Böhm. 2022. "Information About Herd Immunity Through Vaccination and Empathy Promote COVID-19 Vaccination Intentions." *Health Psychology* 41(2):85–93.

Rehman, Abdul, Luan Jingdong and Imran Hussain. 2015. "The province-wise literacy rate in Pakistan and its impact on the economy." *Pacific Science Review B: Humanities and Social Sciences* 1(3):140–144.

Sasaki, Shusaku, Hirofumi Kurokawa and Fumio Ohtake. 2021. "Effective but fragile? Responses to repeated nudge-based messages for preventing the spread of COVID-19 infection." *The Japanese Economic Review* 72(3):371–408.

Sasaki, Shusaku, Tomoya Saito and Fumio Ohtake. 2022. "Nudges for COVID-19 voluntary vaccination: How to explain peer information?" *Social Science & Medicine* 292:114561.

Shang, Jen and Rachel Croson. 2009. "A field experiment in charitable contribution: The impact of social information on the voluntary provision of public goods." *The Economic Journal* 119(540):1422–1439.

Sherman, Susan M, Louise E Smith, Julius Sim, Richard Amlôt, Megan Cutts, Hannah Dasch, G James Rubin and Nick Sevdalis. 2021. "COVID-19 vaccination intention in the UK: results from the COVID-19 vaccination acceptability study (CoVAccS), a nationally representative cross-sectional survey." *Human vaccines & immunotherapeutics* 17(6):1612–1621.

Siddiqui, Adeel, Ali Ahmed, Maria Tanveer, Muhammad Saqlain, Chia Siang Kow and Syed Shahzad Hasan. 2021. "An overview of procurement, pricing, and uptake of COVID-19 vaccines in Pakistan." *Vaccine* 39(37):5251–5253.

Solís Arce, Julio S., Shana S. Warren, Niccolò F. Meriggi, Alexandra Scacco, Nina McMurry, Maarten Voors, Georgiy Syunyaev, Amyn Abdul Malik, Samya Aboutajdine, Opeyemi Adeojo, Deborah Anigo, Alex Armand, Saher Asad, Martin Atyera, Britta Augsburg, Manisha Awasthi, Gloria Eden Ayesiga, Antonella Bancalari, Martina Björkman Nyqvist, Ekaterina Borisova, Constantin Manuel Bosancianu, Magarita Rosa Cabra García, Ali Cheema, Elliott Collins, Filippo Cuccaro, Ahsan Zia Farooqi, Tatheer Fatima, Mattia

Fracchia, Mery Len Galindo Soria, Andrea Guariso, Ali Hasanain, Sofía Jaramillo, Sellu Kallon, Anthony Kamwesigye, Arjun Kharel, Sarah Kreps, Madison Levine, Rebecca Littman, Mohammad Malik, Gisele Manirabaruta, Jean Léodomir Habarimana Mfura, Fatoma Momoh, Alberto Mucauque, Imamo Mussa, Jean Aime Nsabimana, Isaac Obara, María Juliana Otálora, Béchir Wendemi Ouédraogo, Touba Bakary Pare, Melina R. Platas, Laura Polanco, Javaeria Ashraf Qureshi, Mariam Raheem, Vasudha Ramakrishna, Ismail Rendrá, Taimur Shah, Sarene Eyla Shaked, Jacob N. Shapiro, Jakob Svensson, Ahsan Tariq, Achille Mignondo Tchibozo, Hamid Ali Tiwana, Bhartendu Trivedi, Corey Vernot, Pedro C. Vicente, Laurin B. Weissinger, Basit Zafar, Baobao Zhang, Dean Karlan, Michael Callen, Matthieu Teachout, Macartan Humphreys, Ahmed Mushfiq Mobarak and Saad B. Omer. 2021. "COVID-19 vaccine acceptance and hesitancy in low-and middle-income countries." *Nature Medicine* 27(8):1385–1394.

Stewart, Jennifer M, Eamon O'Shea, Cam Donaldson and Phil Shackley. 2002. "Do ordering effects matter in willingness-to-pay studies of health care?" *Journal of Health Economics* 21(4):585–599.

Trent, Mallory, Holly Seale, Abrar Ahmad Chughtai, Daniel Salmon and C Raina MacIntyre. 2022. "Trust in government, intention to vaccinate and COVID-19 vaccine hesitancy: A comparative survey of five large cities in the United States, United Kingdom, and Australia." *Vaccine* 40(17):2498–2505.

Zizzo, Daniel John. 2010. "Experimenter demand effects in economic experiments." *Experimental Economics* 13(1):75–98.

Appendices

A Additional Results

Table A1: Descriptive Statistics of the Entire Sample

	Mean	SD	Min	Max	N
Age	35.057	10.329	25	65	3199
HS grad	0.396	0.489	0	1	3199
Married	0.511	0.500	0	1	3198
Working	0.425	0.494	0	1	3199
Self-reported health	0.871	0.335	0	1	3198
First dose	0.414	0.493	0	1	3199
Second dose	0.166	0.372	0	1	3199
Female	0.324	0.468	0	1	3036
Other gender	0.032	0.177	0	1	3036
Religiosity	2.798	2.076	0	5	3027
Chance of severe symptoms	27.312	24.940	0	100	3197
Chance of moderate symptoms	34.208	25.142	0	100	3197
Trust in government	63.533	24.908	0	100	3198
Trust in science	74.513	17.047	10	100	3198
Trust in medicine	78.593	19.392	0	100	3198
Trust in press	64.246	27.497	0	100	3198

Note: *Self-reported health*: an indicator for the respondent being in good health. *Chance of severe/moderate symptom*: respondent's belief of the chance that contracting COVID-19 will lead to severe/moderate symptoms. *VI*: Vaccination Intention.

Table A2: Comparison of the Entire Sample and the Analysis Sample

	(1) Age	(2) HS grad	(3) Married	(4) Working	(5) Self-reported health	(6) First dose	(7) Female	(8) Other gender
Analysis Sample	-1.822***	0.061***	0.001	0.012	0.035**	-0.691***	-0.011	0.004
	(0.487)	(0.023)	(0.024)	(0.023)	(0.016)	(0.020)	(0.023)	(0.009)
Constant	36.572***	0.346***	0.510***	0.414***	0.842***	0.989***	0.333***	0.029***
	(0.444)	(0.021)	(0.022)	(0.021)	(0.014)	(0.018)	(0.021)	(0.008)
Observations	3,199	3,199	3,198	3,199	3,198	3,199	3,036	3,036
R-squared	0.004	0.002	0.000	0.000	0.002	0.275	0.000	0.000
	Religiosity	Chance of severe symptoms	Chance of moderate symptoms	Trust in government	Trust in science	Trust in medicine	Trust in press	
Analysis Sample	-0.247**	-2.866**	-3.870***	-2.864**	0.259	0.045	-2.999**	
	(0.101)	(1.180)	(1.189)	(1.177)	(0.807)	(0.918)	(1.300)	
Constant	3.004***	29.698***	37.429***	65.916***	74.298***	78.555***	66.741***	
	(0.092)	(1.076)	(1.084)	(1.074)	(0.736)	(0.837)	(1.186)	
Observations	3,027	3,197	3,197	3,198	3,198	3,198	3,198	
R-squared	0.002	0.002	0.003	0.002	0.000	0.000	0.002	

Note: *** p<0.01, ** p<0.05, * p<0.1. *Analysis sample* is an indicator that equals 1 if an observation belongs in the analysis sample, i.e., without the second dose of COVID-19 vaccine. *Self-reported health*: Indicator for being in good health based on self report. *VI*: Vaccination Intention.

Table A3: Balance Tests (four arm treatment)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
N = 2661	Age	HS grad	Married	Working	Self-reported health	First dose	Female	Religiosity
gain	-0.524	0.028	0.007	0.025	0.013	0.013	-0.002	0.037
	(0.614)	(0.032)	(0.027)	(0.025)	(0.025)	(0.026)	(0.030)	(0.131)
gain+	-0.560	0.050	0.033	0.034	0.026	-0.008	-0.003	0.176
	(0.734)	(0.030)	(0.028)	(0.029)	(0.024)	(0.025)	(0.026)	(0.130)
loss	-0.709	0.044	-0.004	0.019	0.022	0.025	-0.019	-0.004
	(0.595)	(0.030)	(0.029)	(0.029)	(0.028)	(0.027)	(0.022)	(0.139)
loss+	-1.300**	0.027	-0.024	0.029	0.022	-0.004	-0.023	0.123
	(0.526)	(0.025)	(0.037)	(0.026)	(0.021)	(0.032)	(0.025)	(0.115)
F test	0.179	0.548	0.543	0.728	0.829	0.810	0.833	0.484

	Chance of severe symptoms	Chance of moderate symptoms	Turst in govern-ment	Trust in science	Trust in medicine	Trust in press	Pre-trtment VI
gain	-0.160	3.792***	-1.046	-1.152	0.826	0.501	3.738*
	(1.272)	(1.059)	(1.329)	(1.080)	(1.006)	(1.597)	(2.120)
gain+	-1.342	2.345*	-0.688	-1.196	-0.189	-0.180	0.247
	(1.276)	(1.172)	(1.617)	(0.999)	(1.610)	(1.555)	(2.680)
loss	-0.164	1.281	0.991	0.917	1.168	1.141	3.761*
	(1.874)	(1.315)	(1.639)	(1.151)	(1.089)	(2.181)	(1.900)
loss+	0.560	1.771	-0.043	0.731	0.167	1.082	0.375
	(1.576)	(1.334)	(1.533)	(0.973)	(1.887)	(1.685)	(1.998)
F test	0.395	0.011	0.502	0.087	0.304	0.834	0.163

Table A4: Effects of Prosocial Messages with Strategic Concerns on Vaccination Intentions

$N = 2,661$	(1) VI free	(2) VI 1500 rp	(3) VI 3000 rp	(4) VI 30 p	(5) VI 70 p
Prosocial Gain (β_g)	7.963***	6.452***	-1.489	4.665**	5.550**
	(1.824)	(1.730)	(1.992)	(2.193)	(2.096)
Prosocial Gain+ (β_{g+})	7.776***	3.850**	3.122	3.383*	3.419*
	(2.000)	(1.476)	(2.036)	(1.884)	(1.816)
Prosocial Loss (β_l)	9.315***	6.657***	4.114*	6.404***	7.697***
	(1.720)	(2.154)	(2.115)	(1.786)	(1.954)
Prosocial Loss+ (β_{l+})	10.736***	8.526***	4.307**	8.710***	8.797***
	(2.093)	(1.663)	(2.000)	(2.165)	(2.153)
Constant	39.227***	37.848***	42.735***	42.112***	43.867***
	(2.715)	(3.474)	(4.652)	(3.375)	(3.823)
R-squared	0.043	0.050	0.030	0.035	0.035
$H_0 : \beta_g = \beta_{g+} = \beta_l = \beta_{l+} = 0$	0.000	0.000	0.004	0.001	0.000
$H_0 : \beta_g = \beta_{g+}$	0.910	0.113	0.012	0.446	0.214
$H_0 : \beta_l = \beta_{l+}$	0.403	0.279	0.930	0.156	0.533
$H_0 : \beta_g = \beta_l$	0.426	0.909	0.014	0.213	0.214
$H_0 : \beta_{g+} = \beta_{l+}$	0.053	0.003	0.454	0.003	0.000

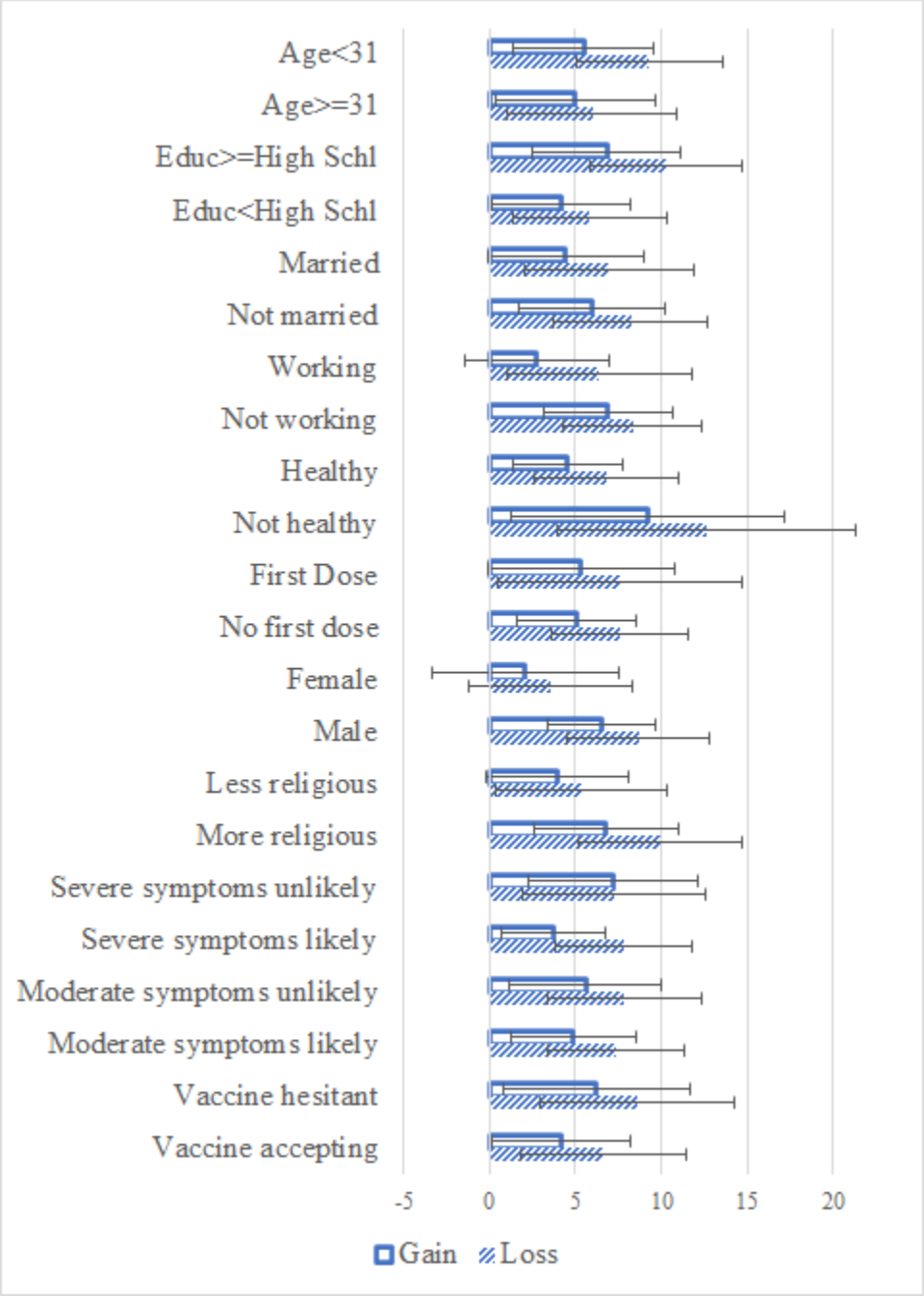
Note: *** p<0.01, ** p<0.05, * p<0.1. Sample is restricted to those who did not receive the second dose of COVID-19 vaccine. Control variables include age, education level, marital status, work status, self-reported health, pre-treatment vaccination intention, first-dose status, belief in the chance of experiencing severe symptoms because of COVID-19, belief in the chance of experiencing moderate symptoms because of COVID-19, and questionnaire ordering indicators. P-values of hypothesis tests are shown.

Table A5: Predictors of Vaccination Status

	(1) Second Dose	(2) First Dose
IHS Trust in Government	0.011**	0.003
	(0.004)	(0.005)
IHS Trust in Press	0.002	0.045***
	(0.005)	(0.008)
IHS Trust in Science	-0.017	0.014
	(0.025)	(0.034)
IHS Trust in Medicine	-0.000	-0.004
	(0.016)	(0.025)
IHS Severe Symptoms Likely	0.000	0.006
	(0.003)	(0.005)
IHS Moderate Symptoms Likely	0.008**	0.010**
	(0.003)	(0.005)
Age	0.001**	0.003***
	(0.001)	(0.001)
High School Graduation	-0.022**	-0.019
	(0.009)	(0.017)
Married	-0.009	-0.007
	(0.009)	(0.023)
Working	-0.001	-0.015
	(0.013)	(0.015)
Self-reported health	-0.042***	0.016
	(0.014)	(0.020)
First dose	0.397***	
	(0.011)	
Constant	0.013	-0.027
	(0.151)	(0.222)
Observations	3,197	3,197
R-squared	0.287	0.014

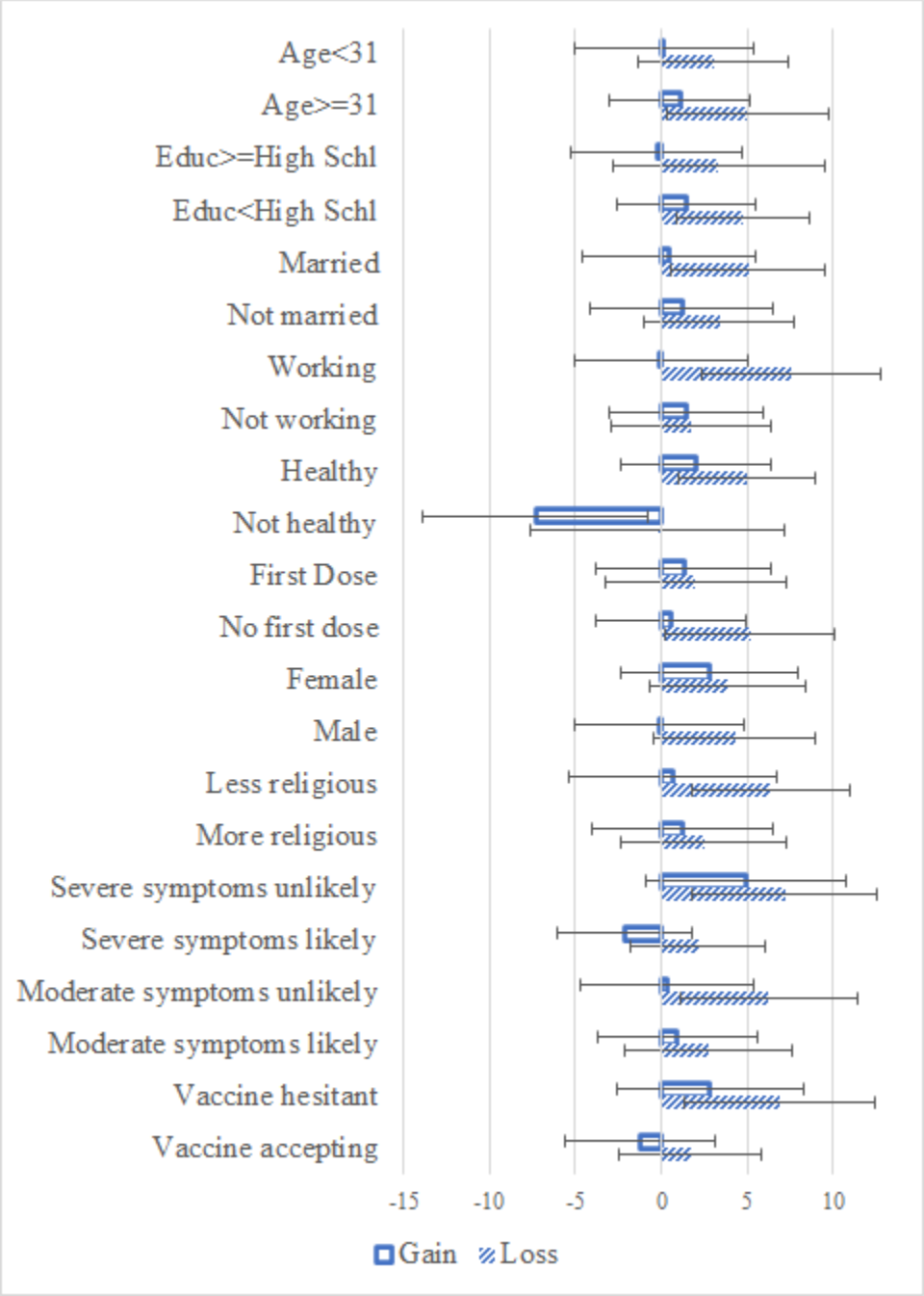
Note: IHS refers to inverse hyperbolic sine transformation that approximates log transformation while retaining zero values.

Figure A1: Effects on VI (1500rp cost) by Baseline Characteristics



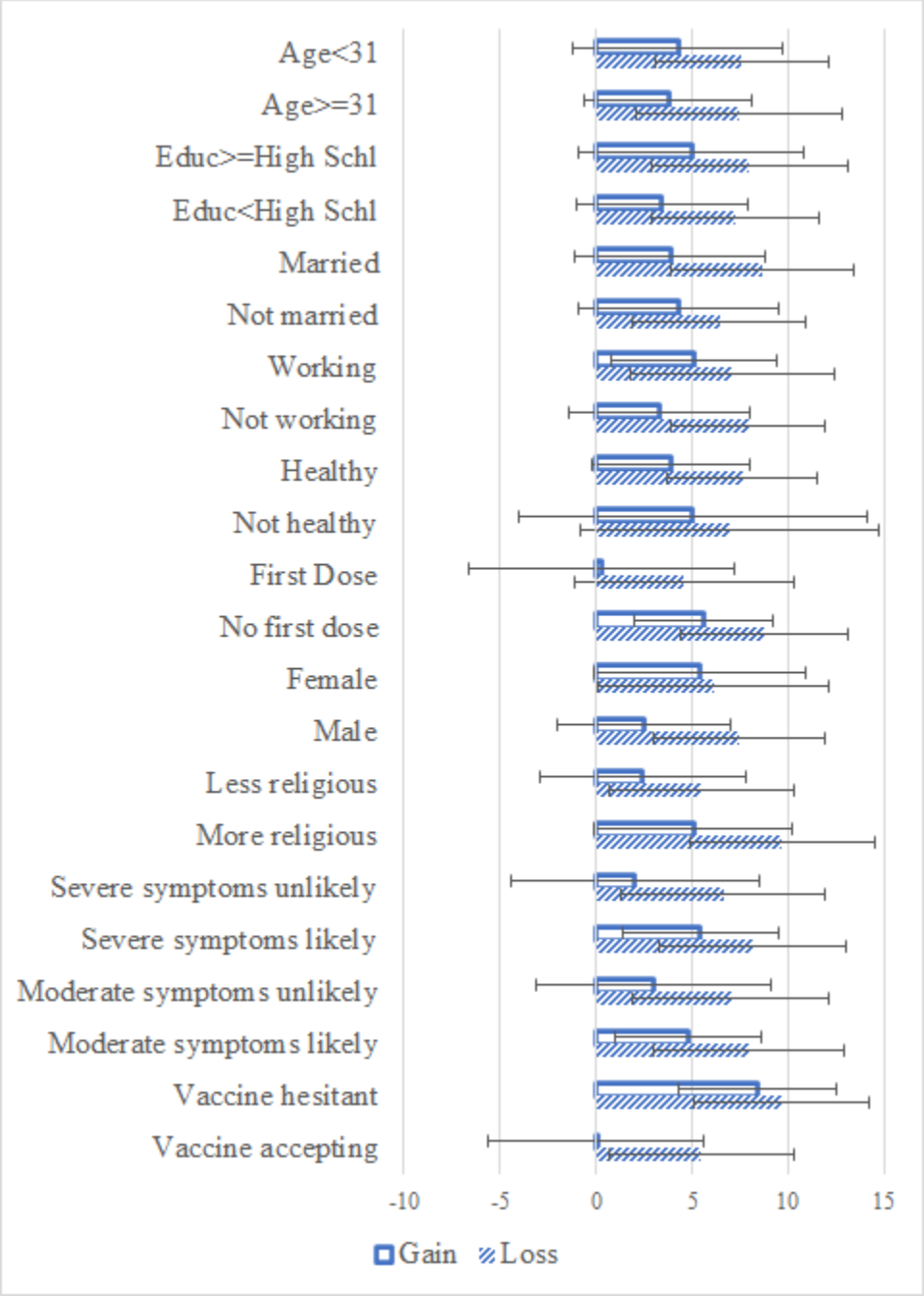
Note: Sample is restricted to those who did not receive the second dose of COVID-19 vaccine. Control variables include age, education level, marital status, work status, self-reported health, pre-treatment vaccination intention, first-dose status, belief in the chance of experiencing severe symptoms because of COVID-19, belief in the chance of experiencing moderate symptoms because of COVID-19, and questionnaire ordering indicators. Bars represent treatment effect estimates based on interaction models. Capped lines represent 95% confidence intervals.

Figure A2: Effects on VI (3000rp cost) by Baseline Characteristics



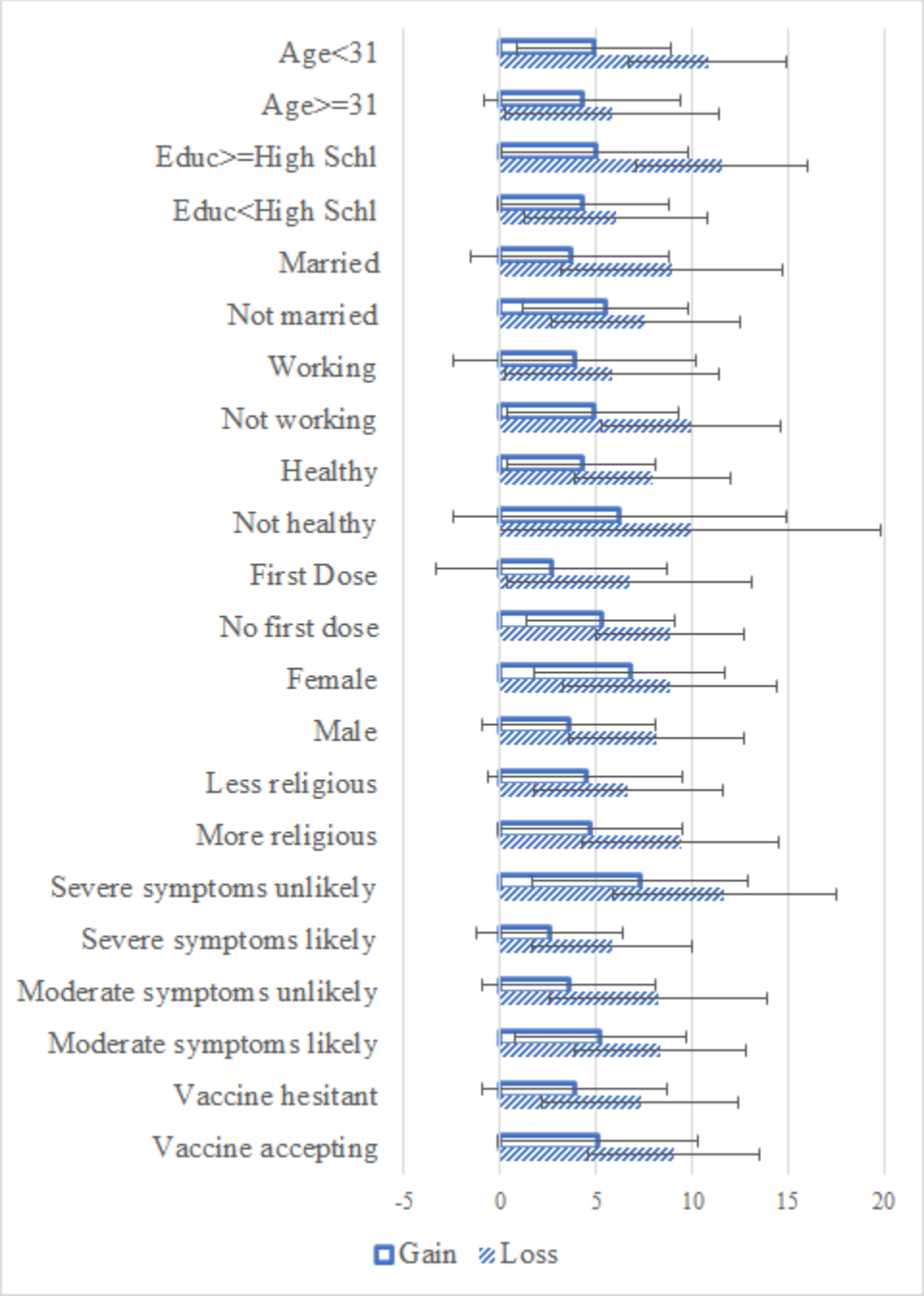
Note: Sample is restricted to those who did not receive the second dose of COVID-19 vaccine. Control variables include age, education level, marital status, work status, self-reported health, pre-treatment vaccination intention, first-dose status, belief in the chance of experiencing severe symptoms because of COVID-19, belief in the chance of experiencing moderate symptoms because of COVID-19, and questionnaire ordering indicators. Bars represent treatment effect estimates based on interaction models. Capped lines represent 95% confidence intervals.

Figure A3: Effects on VI (30% neighbors) by Baseline Characteristics



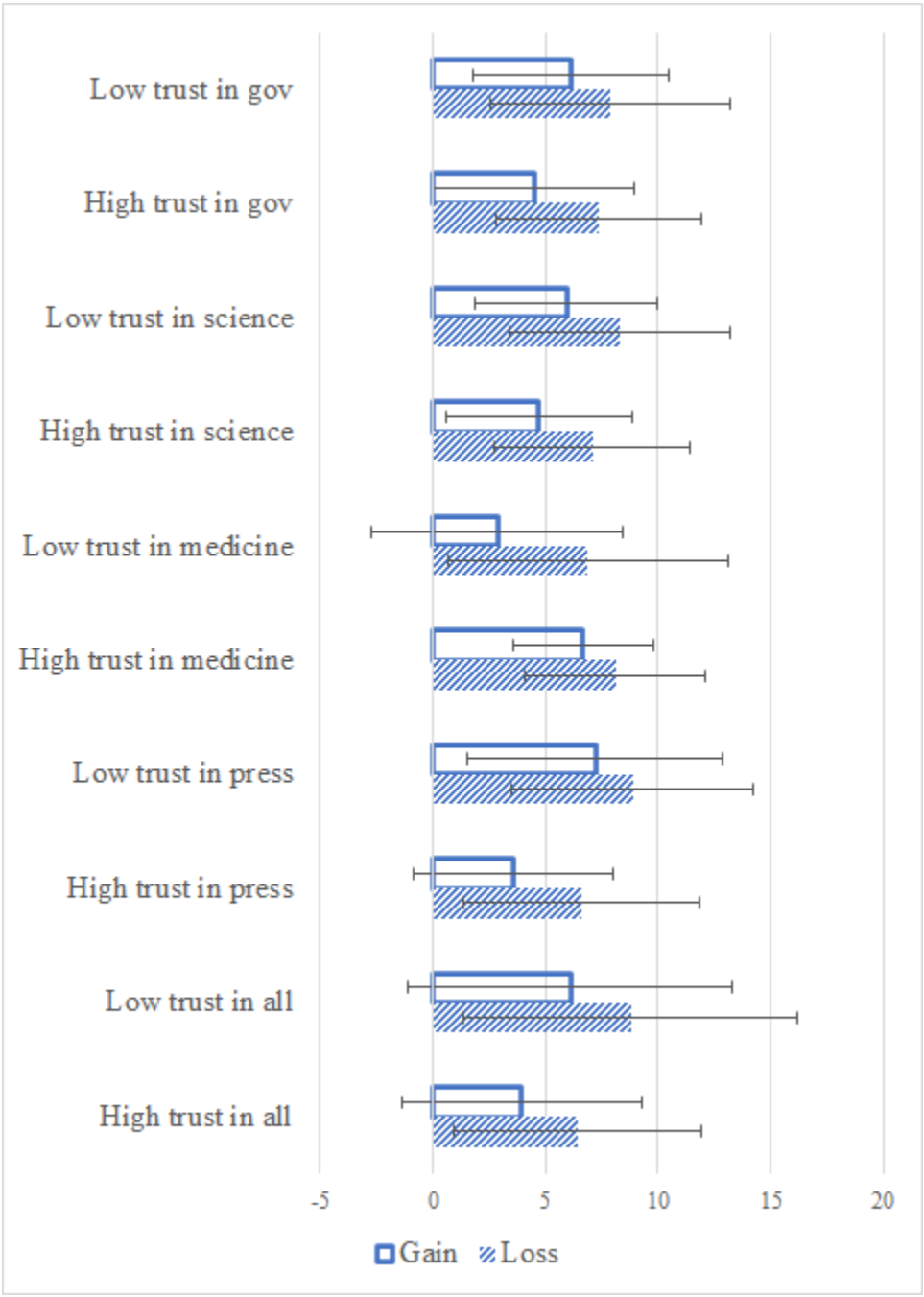
Note: Sample is restricted to those who did not receive the second dose of COVID-19 vaccine. Control variables include age, education level, marital status, work status, self-reported health, pre-treatment vaccination intention, first-dose status, belief in the chance of experiencing severe symptoms because of COVID-19, belief in the chance of experiencing moderate symptoms because of COVID-19, and questionnaire ordering indicators. Bars represent treatment effect estimates based on interaction models. Capped lines represent 95% confidence intervals.

Figure A4: Effects on VI (70% neighbors) by Baseline Characteristics



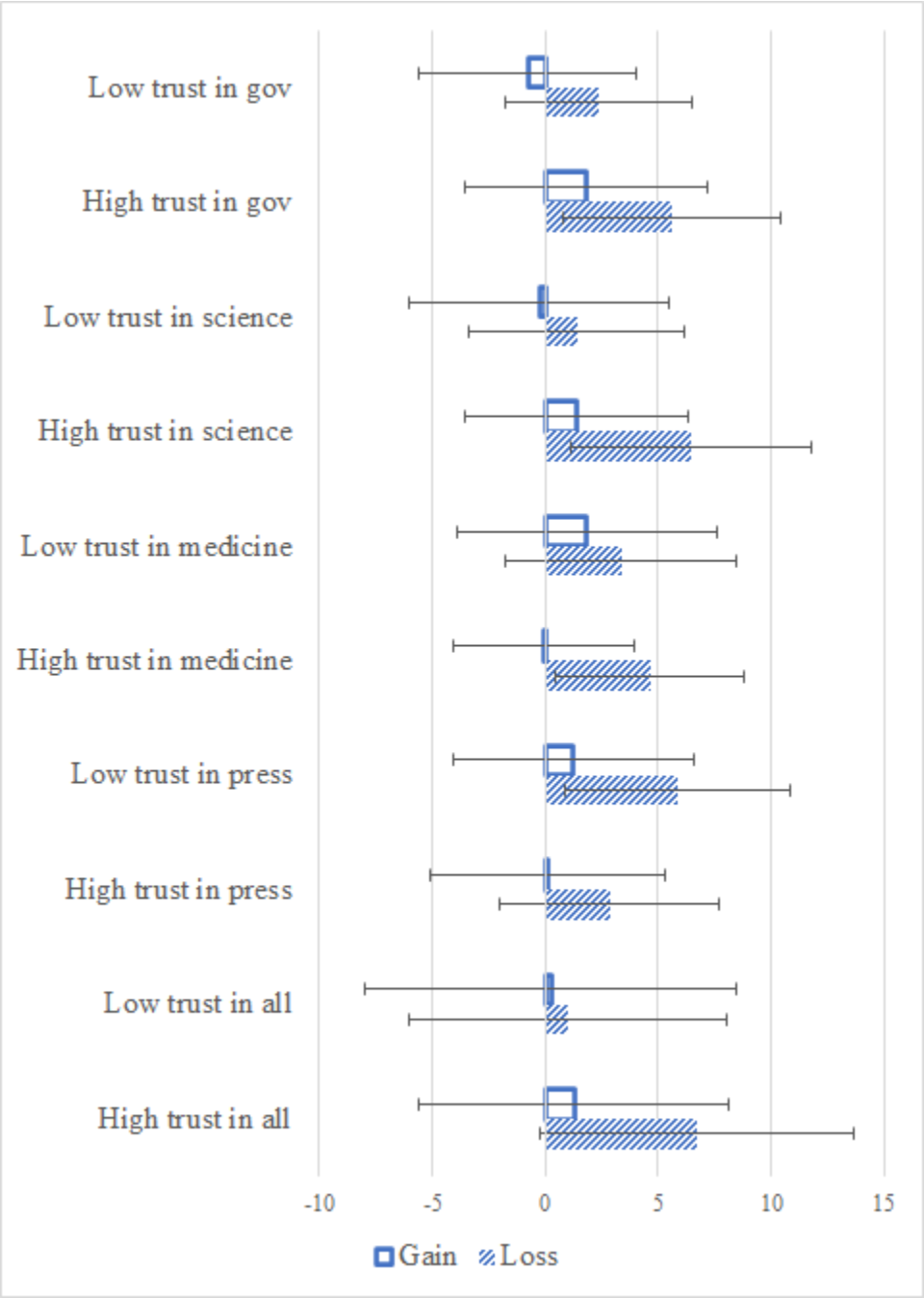
Note: Sample is restricted to those who did not receive the second dose of COVID-19 vaccine. Control variables include age, education level, marital status, work status, self-reported health, pre-treatment vaccination intention, first-dose status, belief in the chance of experiencing severe symptoms because of COVID-19, belief in the chance of experiencing moderate symptoms because of COVID-19, and questionnaire ordering indicators. Bars represent treatment effect estimates based on interaction models. Capped lines represent 95% confidence intervals.

Figure A5: Effects on VI (1500rp cost) by Trust Levels



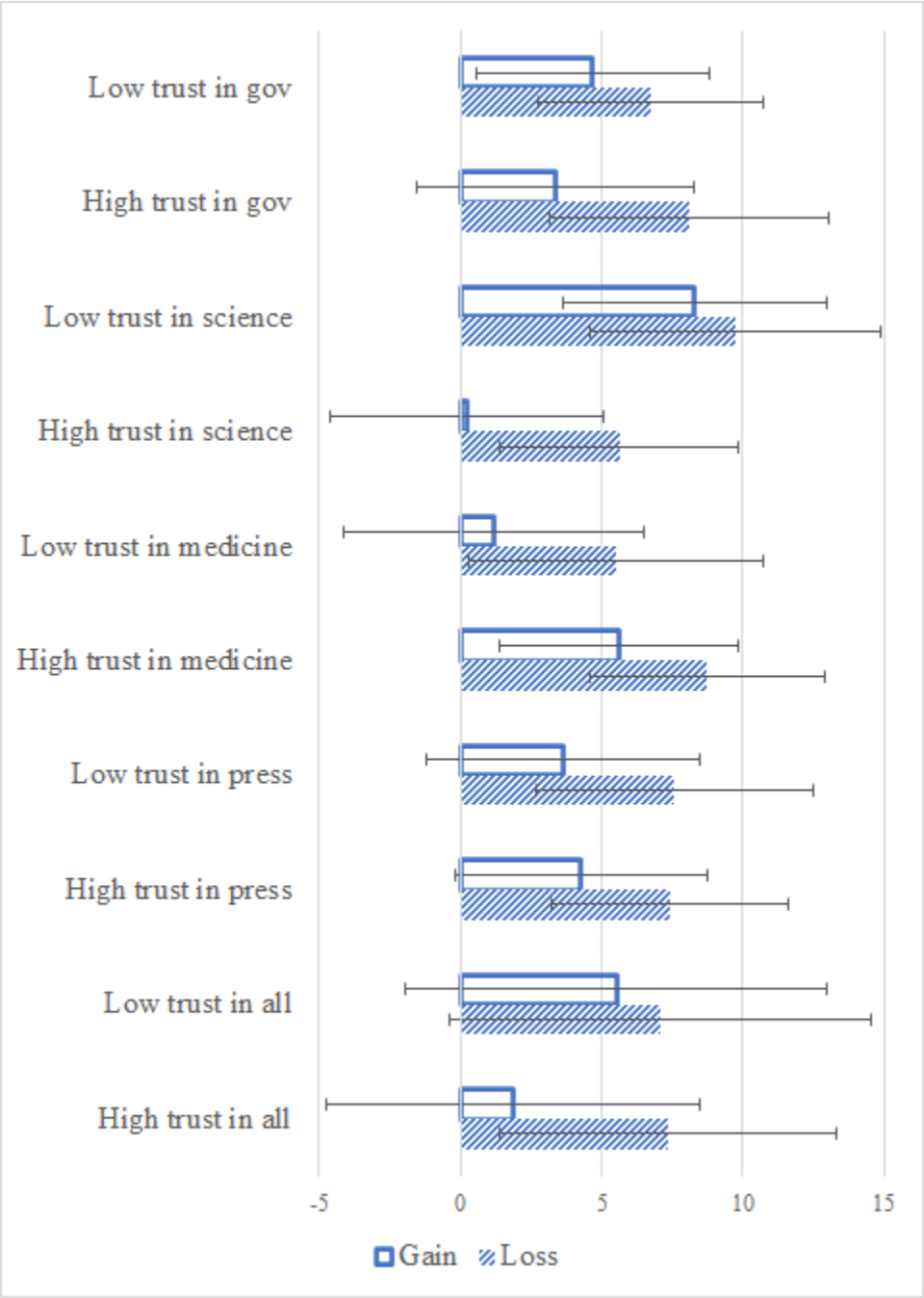
Note: Sample is restricted to those who did not receive the second dose of COVID-19 vaccine. Control variables include age, education level, marital status, work status, self-reported health, pre-treatment vaccination intention, first-dose status, belief in the chance of experiencing severe symptoms because of COVID-19, belief in the chance of experiencing moderate symptoms because of COVID-19, and questionnaire ordering indicators. Bars represent treatment effect estimates based on interaction models. Capped lines represent 95% confidence intervals.

Figure A6: Effects on VI (3000rp cost) by Trust Levels



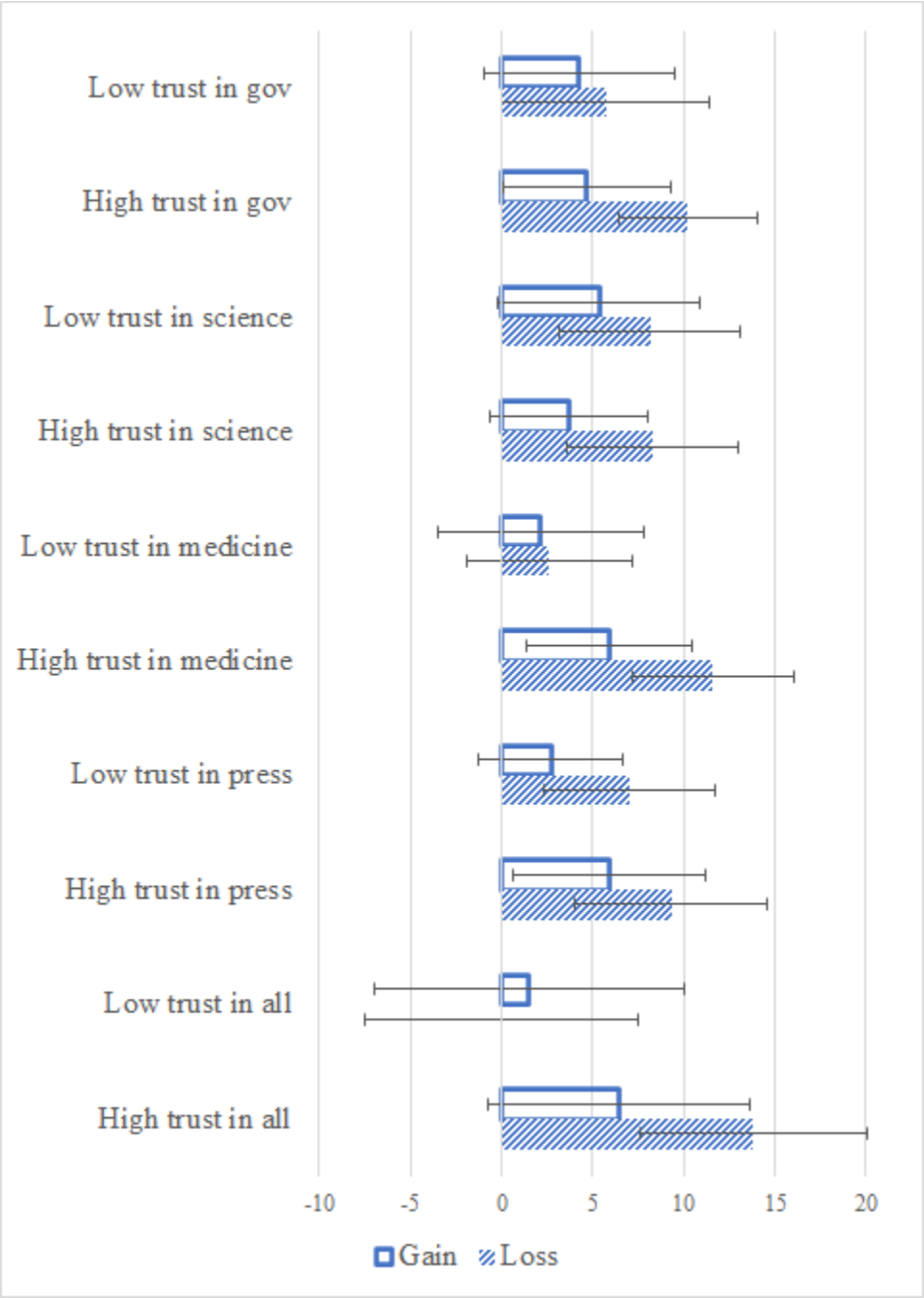
Note: Sample is restricted to those who did not receive the second dose of COVID-19 vaccine. Control variables include age, education level, marital status, work status, self-reported health, pre-treatment vaccination intention, first-dose status, belief in the chance of experiencing severe symptoms because of COVID-19, belief in the chance of experiencing moderate symptoms because of COVID-19, and questionnaire ordering indicators. Bars represent treatment effect estimates based on interaction models. Capped lines represent 95% confidence intervals.

Figure A7: Effects on VI (30% neighbors) by Trust Levels



Note: Sample is restricted to those who did not receive the second dose of COVID-19 vaccine. Control variables include age, education level, marital status, work status, self-reported health, pre-treatment vaccination intention, first-dose status, belief in the chance of experiencing severe symptoms because of COVID-19, belief in the chance of experiencing moderate symptoms because of COVID-19, and questionnaire ordering indicators. Bars represent treatment effect estimates based on interaction models. Capped lines represent 95% confidence intervals.

Figure A8: Effects on VI (70% neighbors) by Trust Levels



Note: Sample is restricted to those who did not receive the second dose of COVID-19 vaccine. Control variables include age, education level, marital status, work status, self-reported health, pre-treatment vaccination intention, first-dose status, belief in the chance of experiencing severe symptoms because of COVID-19, belief in the chance of experiencing moderate symptoms because of COVID-19, and questionnaire ordering indicators. Bars represent treatment effect estimates based on interaction models. Capped lines represent 95% confidence intervals.

B Prosocial Motives and Social Interaction Concerns

Table A4 presents the treatment effects estimated by equation (3). In particular, for the vaccination intention under the "Free" condition that mirrors the policy of the Pakistani government, each of the individual messages is effective in increasing vaccination intention. Furthermore, the differences between the "loss" and the "gain" messages are significant when the framing messages are combined with messages that emphasize selfish gains to prosocial behaviors (Treatments $g+$ and $l+$).

Adding statements that emphasize strategic, selfish aspect of prosocial behaviors increase treatment effects when applied to "loss" treatment (Treatments l and $l+$) but not to the "gain" treatment (Treatments g and $g+$). In fact, emphasizing strategic concerns to the "gain" treatments seems to backfire, lowering treatment effects (Treatments g and $g+$).[16] None of the statistical tests comparing the marginal contribution of strategic messages are significant, however, except in the case of the outcome with hypothetical cost of 3,000 rupees ($H_0 : \beta_g = \beta_{g+}$ and $H_0 : \beta_l = \beta_{l+}$).

B.1 The Model

There are at least two main reasons explaining why agents' vaccination intention is affected by the likelihood of infecting others. The first explanation is that subjects are purely prosocial. They care about the health of others. The second relevant explanation is that not getting vaccinated reduces an agent's social interactions. Hence, to preserve her social interactions, an agent might care about transmitting COVID-19 to others. In this extension, we extend our model to account for these two mechanisms, and explain how we can disentangle them through our experimental design.

[16] This would be consistent with Bénabou and Tirole (2003), who demonstrated theoretically that extrinsic motivation can crowd out intrinsic motivation.

We assume that the "prosocial" cost that subjects feel when they infect others can be decomposed into two components:

$$c_{ps} = c_A + c_{SI}, \tag{27}$$

where $c_A \geq 0$ is the "Altruistic (A)" cost of infecting others, whereas $c_{SI} \geq 0$ is the "Social Interactions (SI)" cost of infecting others. Infecting others decreases a subject's utility for two reasons. First, when the subject is altruistic, she feels a cost $c_A \geq 0$ when she infects someone. Second, when the subject cares about meeting others, she feels a cost $c_{SI} \geq 0$ when she infects someone because this would lead to fewer social interactions.

To disentangle whether subjects react to the treatment by pure altruistic concerns, or because they wish to preserve their social interactions, we conducted an experiment with four treatment arms, T_g, T_{g+}, T_l, and T_{l+}.

Subjects in both T_g and T_{g+} are provided with information about the likelihood of transmitting the disease to others. Both T_g and T_{g+} emphasize the gain associated with getting vaccinated compared to not getting vaccinated. However, subjects in treatment T_{g+} gets the following additional sentence:

> Also, others may be more willing to meet you in person without fear of infection. Your social life could be protected if you are vaccinated.

Similarly, subjects in both T_l and T_{l+} are provided with information about the likelihood of transmitting the disease to others. Both treatments emphasize the loss associated with not getting vaccinated compared to a situation where one is vaccinated. However, subjects in treatment T_{l+} gets the following additional sentence:

> Also, others may be less willing to meet you in person given that they fear that you will transmit them the virus. Your social life could be interrupted if you are not vaccinated.

We assume that while subjects in treatment T_g only focus their attention on the purely altruistic dimension of their vaccination decisions, subjects in treatment T_{g+} focus their attention on both the purely altruistic and the social interaction dimensions of their vaccination decisions. Hence, only treatments T_{g+} and T_{l+} induce a gain-loss utility over both the purely altruistic and the social interaction dimensions of vaccination decisions. In contrast, treatments T_g and T_l induce a gain-loss utility over the purely altruistic dimension of vaccination decision. In these settings, we obtain a slightly modified version of Proposition 1:

Proposition 2 *The likelihood $q_i(t_i) \in [0,1]$ that subject i gets vaccinated is*

$$\begin{aligned} & q_i(T_{k+}) \qquad (28) \\ =& F\left(-\pi + c\,\mathbb{E}[p(v_j,0) - p(v_j,1)] + (c_A + c_{SI})(1 + \eta(1 + (\lambda - 1)r(T_{k+})))\,\mathbb{E}[p(0,v_j) - p(1,v_j)]\right), \end{aligned}$$

and $q_i(T_k)$ is given in Proposition 1, with $k \in \{g,l\}$, $r(T_g) = r(T_{g+}) = 0$, and $r(T_l) = r(T_{l+}) = 1$.

By comparing the effects of treatments T_l and T_{l+}, or the effects of treatments T_g and T_{g+}, we can see whether subjects are primarily reacting to purely altruistic concerns, or to which extent their decision is also affected by social interaction concerns.

C Quantile Treatment Effects

Taking advantage of continuous outcome measures, we estimate conditional quantile functions using quantile regression. The model is

$$\hat{\beta}_\tau = \arg\max_\beta \sum \left(\rho_\tau \left(Y_i - X_i\beta\right)\right), \qquad (29)$$

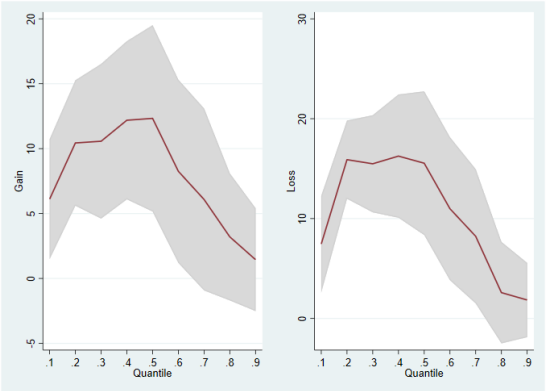
where

$$\hat{q}_\tau = \arg\min_{\rho} \sum_{i=1}^{n} \rho_\tau \left(y_i - q \right) \tag{30}$$

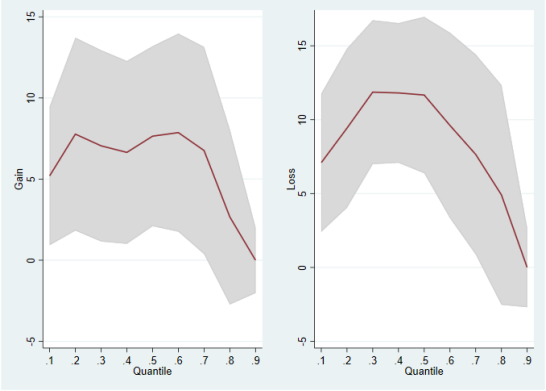
$$= \arg\min_{\rho} \left[(\tau - 1) \sum_{y_i < q} \left(y_i - q \right) + \tau \sum_{y_i \geq q} \left(y_i - q \right) \right]. \tag{31}$$

Figure A9 presents $\hat{\beta}_\tau$ for each decile. The estimates tend to be larger between the 20th and 50th quantiles, suggesting that the treatments act on the margin of decision. The quantile treatment effects are larger and more precise for vaccines provided at lower costs. Figure A10 also shows similar findings for the four-arm treatments.

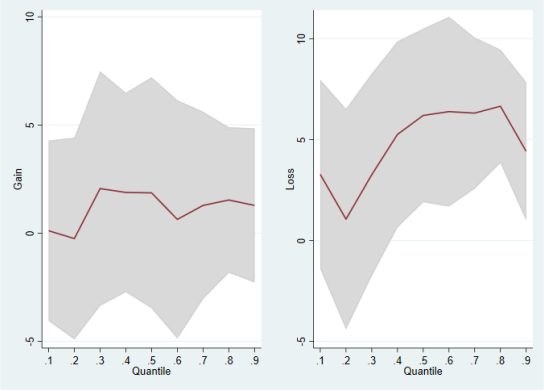
Figure A9: Quantile Treatment Effects



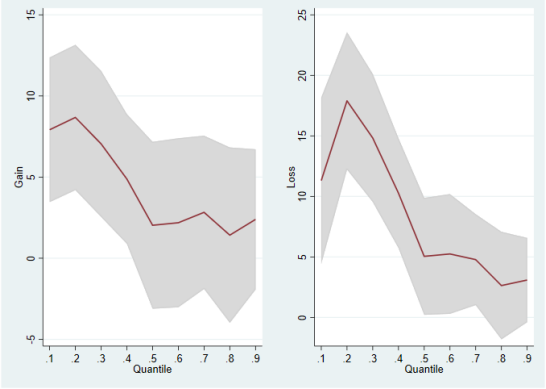
(a) if provided for free



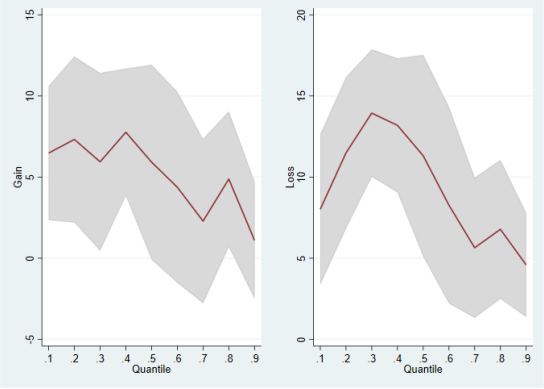
(b) if cost was 1500 rp



(c) if cost was 3000 rp



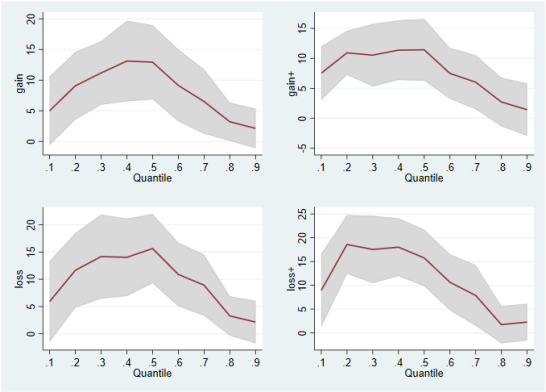
(d) if 30% got it



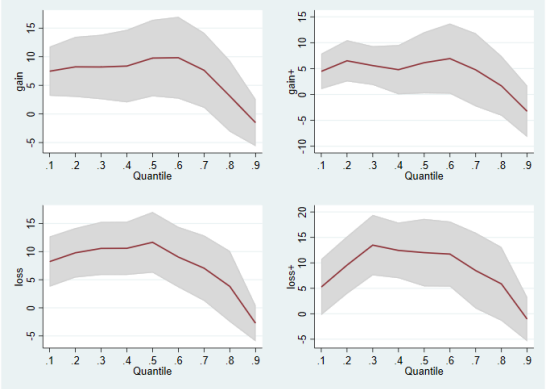
(e) if 70% got it

Note: Control variables include age, education level, marital status, work status, self-reported health, pre-treatment vaccination intention, first-dose status, and questionnaire ordering indicators.

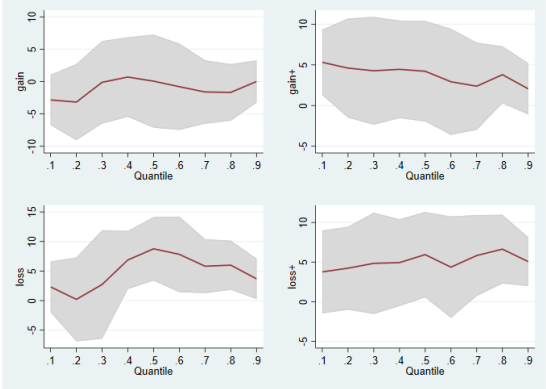
Figure A10: Quantile Treatment Effects (Four Arm)



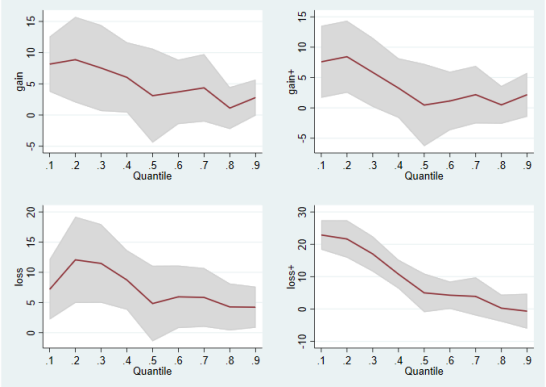
(a) if provided for free



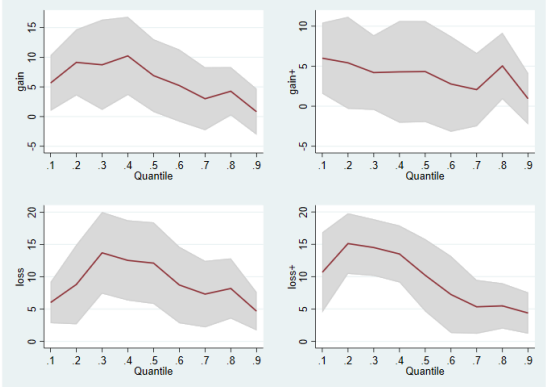
(b) if cost was 1500 rp



(c) if cost was 3000 rp



(d) if 30% got it



(e) if 70% got it

Note: Control variables include age, education level, marital status, work status, self-reported health, pre-treatment vaccination intention, first-dose status, and questionnaire ordering indicators. *gain*: prosocial-gain treatment; *gain+*: prosocial-gain-strategic treatment; *loss*: prosocial-loss treatment; *loss+*: prosocial-loss-strategic treatment.