A behavioral intervention for encouraging the resumption of economic activity after COVID-19 lockdowns end: Evidence from a large scale-field experiment in China

Supplementary Information

Juan Palacios,1 Yichun Fan,1 Erez Yoeli 2, Jianghao Wang 1,3, Yuchen Chai 1, Weizeng Sun,4 David Rand,2 and Siqi Zheng 1

1Department of Urban Studies and Planning, Massachusetts Institute of Technology, MA, USA
2Sloan School of Management, Massachusetts Institute of Technology, MA, USA
3Institute of Geographic Sciences and Natural Resources Research, Chinese Academy of Science, China
4School of Economics, Central University of Finance and Economics, China
## Contents

A Robustness checks and supplementary analyses

| Section | Page |
|---------|------|
| A.1 Non-parametric comparisons of treatment and control group | 1 |
| A.1.1 Distribution of prior (Wednesday) and posterior beliefs (Friday) | 1 |
| A.1.2 Changes in beliefs over prior belief distribution | 2 |
| A.1.3 Correlation between beliefs regarding planned and actual behavior | 3 |
| A.1.4 Changes in behavior | 4 |
| A.2 Regression tables with different model specifications | 4 |
| A.3 Treatment effects as a function of bias in prior beliefs | 7 |
| A.4 Moderator analysis | 8 |
| A.4.1 Subgroup robustness check: Belief updates | 8 |
| A.4.2 Multiple hypothesis testing | 8 |
| A.4.3 Interaction effects | 10 |
| A.5 Normative channels and experimenter demand effects | 12 |

B Extra Manipulation I: Restaurant Certification

| Section | Page |
|---------|------|
| B.1 Construction information treatment | 15 |
| B.2 Estimated effects of restaurant certifications on individual behavior | 17 |

C Extra manipulation II: Social Preferences, Perceived Norms and Precautionary Behaviors

| Section | Page |
|---------|------|
| C.1 Construction treatment | 18 |
| C.2 Estimated effects of intervention precautionary behavior on individual behavior | 21 |

D Changes in Risk Perception

| Section | Page |
|---------|------|
| D.1 Risk aversion and risk perception | 21 |

E Initial conditions in park and restaurant settings

| Section | Page |
|---------|------|
| E.1 Comparison restaurant and park visits first week relative to Pre-COVID | 24 |
| E.2 Initial risk perception to visit public spaces | 24 |

F Prior beliefs over weeks

| Section | Page |
|---------|------|
|       | 26 |

G Supplementary methods: Correspondence survey responses with GPS data

| Section | Page |
|---------|------|
| G.1 GPS data collection: Sampling method and limitations | 27 |
| G.2 Correspondence coordinates reported restaurant visits with individual’s smartphone GPS coordinates | 28 |
| G.3 Time at home and reported visit to restaurant | 29 |
H Supplementary Notes

H.1 Description moderator scales .................................................. 31
H.2 Display treatment information .................................................... 31
H.3 Summary statistics and balance tests .......................................... 36
H.4 Recruitment strategy and sample attrition ................................... 39
H.5 Percentages used to construct descriptive norm .............................. 40
H.6 Compensation participants ....................................................... 41
A Robustness checks and supplementary analyses

In the main text, we describe the treatment effects based on an ordinary least square regression. In this section, we present the non-parametric comparisons between control and treatment group, specification checks (testing the treatment effects in non-linear models (i.e. Logit and Probit). In addition, we show the high correspondence in participants’ prior beliefs regarding planned and actual behavior of neighbors on our Wednesday survey. Finally, we also include the regression analysis which examines the differences in treatment effect over the distribution of participants’ bias in prior beliefs.

The results of the tests provides supportive evidence showing that the intervention worked by encouraging individuals with prior beliefs below the truth to visit restaurants, and was specially strong for those in the lowest part of the distribution of beliefs.

A.1 Non-parametric comparisons of treatment and control group

A.1.1 Distribution of prior (Wednesday) and posterior beliefs (Friday)

Figure A.1 shows the distribution of the difference between the posterior beliefs regarding the proportion of neighbors going to restaurants (left) and parks (rights) over the weekend and the true percentage included in the information treatment. The sample is restricted to the first week the participant was part of the experiment. The treatment group is substantially more compressed around the truth (zero bias) than the control group in their responses to our Friday questionnaire, after receiving the percentage of neighbors planning on going to restaurants and parks. This indicates that subjects in the treatment group believed that the information provided was truthful.

Figure A.1: Distribution of beliefs regarding the actual percentage of neighbors going to parks or restaurants

Panel a. Beliefs Restaurants  
Panel b. Beliefs Park

Notes: Sample restricted to the first week individuals took part in the experiment.
A.1.2 Changes in beliefs over prior belief distribution

Figure A.2 describes the average changes in beliefs from Wednesday to Friday separately for individuals receiving the descriptive norm information and those who did not.

Figure A.2: Changes in beliefs associated with descriptive norm over bias in prior beliefs

a Beliefs Restaurants

Note: i) The vertical axes in Panel display the change in subjects’ beliefs from Wednesday (prior) to Friday (posterior) regarding the percentage of neighbors who are actually going to go to restaurants (panel a) and parks (panel b) over the weekend. The horizontal axes display the bias in beliefs prior to our intervention, computed by subtracting subjects’ prior beliefs regarding the number of neighbors that are planning to go to restaurants from the true percentage of neighbors planning to go to restaurants. The true percentages, used to build the descriptive norm information, are computed from the proportion of individuals that reported in our Wednesday questionnaire to have plans to go to restaurants over the weekend. The sample includes only responses of individuals that were treated for the first time, and control subjects that were in the sample for one week. Dots represent the change in subjects’ beliefs about the % of others who are going out to restaurants, after our intervention, as a function of their bias prior to the intervention (prior belief - true percentage). The steeper the fitted line, the more subjects with beliefs different from the truth adjusted those beliefs; a -45 degree line would mean that our intervention had perfectly corrected beliefs. ii) The bar graph included in the right hand side of Panel A and Panel B indicate the average changes effects in beliefs for individuals with prior beliefs below and above the truth, separately; the error bars are the 95% confidence intervals.

b Beliefs Parks
A.1.3 Correlation between beliefs regarding planned and actual behavior

The goal of our experimental design is to isolate the causal effect of variation in beliefs regarding neighbors’ restaurant visits on one’s own restaurant visitation. A challenge we face is that such information must be provided prior to the weekend, before we know the actual visit decisions others.

We solve this problem following the experimental design by (39). We collect information on individuals’ beliefs about others’ planned visits to restaurants, as well as beliefs about others’ future actual turnout at restaurants.

With this design, we are able to elicit planned restaurant visitation (as opposed to actual visitation) prior to the weekend. This allows us to provide truthful information regarding others’ planned visitation to restaurant, plausibly affecting beliefs regarding others’ actual protest participation. Figure A.3 shows the high correspondence between beliefs regarding neighbors’ planned and actual behavior.

Figure A.3: Correspondence of prior beliefs (Wednesday survey) regarding planned and actual behavior of neighbors

Panel a. Beliefs Restaurants

Panel b. Beliefs Park

Notes: The binned scatter plots displays the relationship between prior beliefs regarding planned (x-axis) and actual (y-axis) behavior of neighbors, for the restaurant (left) and park (right) setting. Pearson correlation coefficients in beliefs regarding planned and actual behavior is 0.83 for restaurants and 0.86 for parks. As in our main results, the sample is restricted to the first week individuals took part in the experiment.
A.1.4 Changes in behavior

Figure A.4: Changes in behavior associated with descriptive norm
Panel a. Restaurant visits  Panel b. Park visits

Note: The figure shows the average differences in propensity to visit a restaurant between treatment and control group disaggregated by the distribution of prior beliefs regarding the number of neighbors planning to go to restaurants (relative to the true percentage used to construct the treatment information). Error bars describe mean ± standard error. Sample restricted to the first week individuals took part in the experiment.

Figure A.5: Changes in behavior associated with descriptive norm over distribution of prior beliefs
Panel a. Restaurant visits  Panel b. Park visits

Note: The figure shows the comparisons of treatment and control group in visits to restaurants (left) and parks (right) using local polynomial smooth plots (non-parametric). Shadowed area describe the 95 confidence interval. Sample restricted to the first week individuals took part in the experiment. Control group describes individuals receiving no information, treatment group describes individuals receiving descriptive norms.

A.2 Regression tables with different model specifications
Table A.1: Estimation Average Treatment Effects. Specifications: Linear Probability Model, Logit and Probit. Outcome: Individual visits a restaurant.

|               | (1)     | (2)     | (3)     | (4)     | (5)     | (6)     | (7)     | (8)     | (9)     |
|---------------|---------|---------|---------|---------|---------|---------|---------|---------|---------|
|               | LPM     | Logit   | Probit  | LPM     | Logit   | Probit  | LPM     | Logit   | Probit  |
| Descriptive Norm | 0.125** | 0.652** | 0.398** | 0.120** | 0.655** | 0.402** | -0.015 | -0.086 | -0.058 |
|               | (0.056) | (0.282) | (0.167) | (0.055) | (0.284) | (0.169) | (0.065) | (0.330) | (0.195) |
| [0.025]       | [0.021] | [0.017] | [0.031] | [0.021] | [0.017] | [0.815] | [0.795] | [0.765] |         |
| Prior Above Truth | 0.085   | 0.432   | 0.262   | 0.060   | 0.300   | 0.178   | 0.160   | 0.150   | 0.143   |
|               | (0.060) | (0.300) | (0.178) | (0.060) | (0.300) | (0.178) | (0.160) | (0.150) | (0.143) |
| Descriptive Norm | -0.137  | -0.710* | -0.441* | 0.084   | 0.417   | 0.249   | 0.104   | 0.089   | 0.077   |
| x Prior Above Truth | (0.084) | (0.417) | (0.249) | (0.104) | (0.089) | (0.077) |         |         |         |
| Observations  | 489     | 489     | 489     | 272     | 269     | 269     | 217     | 217     | 217     |
| R-squared / Pseudo R2 | 0.188   | 0.146   | 0.147   | 0.172   | 0.135   | 0.137   | 0.231   | 0.179   | 0.180   |
| Sample        | Full    | Full    | Full    | Below   | Below   | Below   | Below   | Above   | Above   |
|               | Sample  | Sample  | Sample  | Truth   | Truth   | Truth   | Truth   | Truth   | Truth   |
| Controls      | YES     | YES     | YES     | YES     | YES     | YES     | YES     | YES     | YES     |

Notes: Robust standard errors are in parentheses. P-values shown in square brackets. Asterisks indicate coefficient statistical significance level: ∗P < 0.10; ∗∗P < 0.05; ∗∗∗P < 0.01. Column (1), (4) and (7) display the coefficients from a linear probability model (LPM) estimation; Column (2), (5) and (8) display the coefficients from a Logit regression; and Column (3), (6) and (9) display the coefficients from a Probit regression. The dependent variable is a dummy variable indicating whether a respondent reported to going to a restaurant. All regressions include as controls the education level of individuals, district fixed effects, a dummy variable showing whether the individual went to restaurants at least once a week before the COVID-19 pandemic, experimental week fixed effects, whether the individual received the information treatment about the precautionary measures of restaurants, and a dummy variable indicating whether the individual reported in our Wednesday questionnaire to have plans to go to a restaurant over the weekend. Sample restricted to the first week each participant joined the experiment.

In a pooled regression with an interaction term between the treatment dummy and a dummy denoting the park setting, we tested whether the estimated treatment effect differs significantly between the park and restaurant setting. We predict that this interaction should be negative, but whether it will be statistically significant from zero depends on the statistical power (But at least it should not be statistically greater than zero). We are underpowered for detecting such an interaction effect, which requires substantially more statistical power than the main effects. Indeed, the interaction term is negative, though not statistically significant in the sample of individuals with priors below the truth [interaction $\beta = -0.09$, p-value =0.23]. The interaction is marginally significant for those individuals with prior beliefs below one standard deviation of the truth, which is where our restaurant intervention really seems to have a large impact [interaction $\beta = -0.20$, P−value = 0.08]. Consistent with this, the 3-way interaction between treatment, domain, and prior beliefs [interaction $\beta = -0.1$, P−value = 0.06], suggesting that the interaction effect gets stronger for people further away from the truth (as one might expect from panel b in Figure 2 and 3 in the main text). Thus, we have only partial support for this secondary analysis suggesting that our intervention worked by increasing subjects’ perception that restaurants are safe.
Table A.2: Estimation Average Treatment Effects. Specifications: Linear Probability Model, Logit and Probit. Outcome: Individual visits a park

|                  | (1)     | (2)     | (3)     | (4)     | (5)     | (6)     | (7)     | (8)     | (9)     |
|------------------|---------|---------|---------|---------|---------|---------|---------|---------|---------|
|                  | LPM     | Logit   | Probit  | LPM     | Logit   | Probit  | LPM     | Logit   | Probit  |
| Descriptive Norm | 0.039   | 0.275   | 0.175   | 0.060   | 0.267   | 0.167   | 0.080   | 0.363   | 0.214   |
|                  | (0.050) | (0.255) | (0.157) | (0.060) | (0.258) | (0.159) | (0.067) | (0.299) | (0.181) |
| [0.438] [0.282]  | [0.266] | [0.322] | [0.301] | [0.295] | [0.232] | [0.225] | [0.235] |         |         |
| Prior Above Truth| 0.153***| -0.177  | -0.100  |         |         |         |         |         |         |
|                  | (0.057) | (0.277) | (0.168) |         |         |         |         |         |         |
| [0.007] [0.524]  | [0.553] |         |         |         |         |         |         |         |         |
| Descriptive Norm | -0.110  | 0.141   | 0.073   |         |         |         |         |         |         |
| × Prior Above Truth| (0.076) | (0.382) | (0.234) |         |         |         |         |         |         |
|                  | (0.007) | (0.524) | [0.553] |         |         |         |         |         |         |
| Observations     | 506     | 508     | 508     | 271     | 269     | 269     | 237     | 237     | 237     |
| R-squared        | 0.289   | 0.0531  | 0.0528  | 0.072   | 0.0500  | 0.0497  | 0.090   | 0.0699  | 0.0694  |
| Sample           | Full    | Full    | Full    | Below   | Below   | Below   | Below   | Above   | Above   |
| Controls         | YES     | YES     | YES     | YES     | YES     | YES     | YES     | YES     | YES     |

Notes: Robust standard errors are in parentheses. P-values shown in square brackets. Asterisks indicate coefficient statistical significance level: ∗∗∗P < 0.01; ∗∗P < 0.05; ∗P < 0.10. Column (1), (4) and (7) display the coefficients from a linear probability model (LPM) estimation; Column (2), (5) and (8) display the coefficients from a Logit regression; and Column (3), (6) and (9) display the coefficients from a Probit regression. The dependent variable is a dummy variable indicating whether a respondent reported to going to a park. All regressions include as controls the education level of individuals, district fixed effects, a dummy variable showing whether the individual went to restaurants at least once a week before the COVID-19 pandemic, experimental week fixed effects, whether the individual received the information treatment about the precautionary measures of restaurants, and a dummy variable indicating whether the individual reported in our Wednesday questionnaire to have plans to go to a restaurant over the weekend. Sample restricted to the first week each participant joined the experiment.

In a pooled regression with an interaction term between the treatment dummy and a dummy denoting the park setting, we tested whether the estimated treatment effect differs significantly between the park and restaurant setting. We predict that this interaction should be negative, but whether it will be statistically significant from zero depends on the statistical power (But at least it should not be statistically greater than zero). We are underpowered for detecting such an interaction effect, which requires substantially more statistical power than the main effects. Indeed, the interaction term is negative, though not statistically significant in the sample of individuals with priors below the truth [interaction β = -0.09, p-value =0.23]. The interaction is marginally significant for those individuals with prior beliefs below one standard deviation of the truth, which is where our restaurant intervention really seems to have a large impact [interaction β = -0.20, P-value = 0.08]. Consistent with this, the 3−way interaction between treatment, domain, and prior beliefs [interaction β = -0.1, P-value = 0.06], suggesting that the interaction effect gets stronger for people further away from the truth (as one might expect from panel b in Figure 2 and 3 in the main text). Thus, we have only partial support for this secondary analysis suggesting that our intervention worked by increasing subjects’ perception that restaurants are safe.
A.3 Treatment effects as a function of bias in prior beliefs

In the main text, Figure 2 displays the differences in posterior beliefs and behavior between treatment and control group. The figure is based on the following regression model:

\[ Y_{it} = \alpha T_i + \beta_1 T_i \times bias_i + \beta_2 T_i \times bias_i \times bias_i^2 + \delta_1 bias_i + \delta_2 bias_i^2 + X_i \Gamma + \epsilon_{it} \]  

(3)

where \( Y_{it} \) includes the two main outcomes of interest \{PosteriorBelief_{it}, Behavior_{it}\}. The variable \( PosteriorBelief_{it} \) is individual \( i \)'s belief about the proportion of neighbors who will actually go to the restaurants (or parks) elicited on Friday post-treatment survey in week \( t \). \( Behavior_{it} \) is a dummy variable indicating if individual \( i \) visits a restaurant (or park) on the weekend of week \( t \). \( T_i \) is a dummy variable taking a value of one for people who receive the descriptive norm treatment information and zero otherwise; \( bias_{it} \) is a continuous variable indicating the difference between \( i \)'s pre-treatment belief on Wednesday regarding proportion of neighbors planning to go to restaurants (or parks) and the actual percentage; \( X_i \) describes the set of control variables. To calculate the main treatment effects on behavior (\( Behavior_{it} \)), the list control variables include the planned behavior variable (measured on Wednesdays, taking the value of 1 if individual plans to visit a restaurant), strata variables (neighborhood fixed effects and a dummy variable indicating whether the individual went to restaurants at least once a week before the COVID pandemic), and education level (below college dummy indicator).

Figure A.6: Coefficient interaction terms associated with bias in prior beliefs

Panel a. Posterior Beliefs

Panel b. Behavior

Note: Dots describe the point estimates and bars describe the 95 confidence interval. Figure A.6 shows the coefficients associated with the interaction terms between the treatment indicator, the bias in individual’s prior beliefs (standardized), and the quadratic term of individuals’ previous beliefs.

i) Panel a describes the coefficients in the regression predicting individual posterior beliefs, and Panel b describes the regression of predicting individual behavior. We estimate separately the regressions for parks and restaurants. The coefficients in Panel a show the similarities between parks and restaurants. The interaction between individual’s bias in prior belief is negative and significantly different from zero. The interaction between the treatment indicator and the quadratic transformation of the bias is not statistically different from zero in restaurants and parks.

ii) Panel b shows the coefficients of predicting behavior. There are clear differences between parks and restaurants. In the restaurant regression, both coefficients are statistically different from zero. The two coefficients in parks are not statistically different from zero.
A.4 Moderator analysis

In our main text, we show the differential effect of the descriptive norm intervention between risk averse and risk tolerant participants. This section presents a series of checks to ensure that those differences are robust. We present the decomposition of treatment effects based on interaction terms as well as sub sample split. The interaction terms are tested in both discrete (dummy indicating individual is above median in risk tolerance) and continuous (standardized scale) measures. Finally, we correct for multiple hypothesis testing. Across all specifications and checks, risk tolerance (general, and health specific risk aversion) stand out as a key moderator of the treatment effects of the descriptive norm intervention. For completeness, we also included the full moderation analysis for all the pre-specified moderators (RCT ID: AEARCTR-0005644).

A.4.1 Subgroup robustness check: Belief updates

In this subsection, we test if people with different risk tolerance updated posterior belief differently. If risk perception is the key factor moderating the effects of our descriptive norm intervention, we should see people of both group corrected their belief effectively after receiving our intervention. Figure A.7 show the estimated coefficients in Equation 1 describing the changes associated with our treatment dummy $T$ on posterior belief (panel a) and behavior (panel b) for the different subsamples of risk aversion (above and below sample median). The results indicate that all subgroups updated their beliefs to a similar extent, yet the beliefs did not translate into actual behavior for all. Specifically, only more risk tolerant people (in both general and health-specific domains) changed behavior after adjusting their beliefs that more people are dining out (Panel b, Figure A.7).

A.4.2 Multiple hypothesis testing

Table A.3 presents the p-values adjusted for multiple hypothesis testing in the subgroup analysis. We follow the methods presented in (31). Table A.3 shows that the effect of the descriptive norm intervention among subjects who had low beliefs and were risk averse was robust to multiple comparisons (adjusted P-value < 0.05).
The figures investigate the differences in the average treatment effect of our intervention on belief and actual behavior for people with different risk tolerance, in the general and health domain. Sample restricted to individuals with prior belief below the truth. (a) shows the treatment effect on belief about other’s behaviors, and (b) shows the treatment effect on actual dining out behavior. The error bars describe the 95% confidence intervals.

Table A.3: Multiple hypothesis test adjusted p-value. Outcome: Probability of visiting a restaurant (1=Yes)

| Risk preference (general)                      | TE (unadjusted) | p values (multiplicity adjusted) |
|-----------------------------------------------|-----------------|---------------------------------|
| Prior below truth and risk averse             | 0.006           | 0.933                           |
| Prior below truth and risk seeking            | 0.242           | 0.005***                        |
| Prior above truth and risk averse             | 0.002           | 0.986                           |
| Prior above truth and risk seeking            | 0.007           | 0.948                           |

| Risk preference (health)                      | TE (unadjusted) | p values (multiplicity adjusted) |
|-----------------------------------------------|-----------------|---------------------------------|
| Prior below truth and risk averse             | 0.016           | 0.848                           |
| Prior below truth and risk seeking            | 0.214           | 0.013**                         |
| Prior above truth and risk averse             | 0.032           | 0.734                           |
| Prior above truth and risk seeking            | 0.041           | 0.686                           |

Note: The multiple hypothesis testing is based on List et al. (2019).
A.4.3 Interaction effects

The results of how risk preference moderates treatment effect presented in the main text are based on sub-group analysis. Here, we also test interaction terms on the original analysis, instead of running separate regressions by subgroups. The regression formula is described in Equation 2, Methods Section. Since the treatment effect mainly took place within the subgroup who have initial beliefs below the truth (i.e., underestimated the number of neighbors going to restaurants), we focus on this group for the analysis of risk moderator.

First, we use a dummy variable to represent whether an individual is risk tolerant (i.e., risk tolerance score above median) and estimate the coefficient associated with its interaction with the treatment dummy. Supplementary Figure A.8 displays the estimation results. The results show that, no matter we measure risk in general realm or health related realm, treatment effect of descriptive norm is significantly larger for people who are more risk tolerance in the general domain [interaction effect general risk: $\beta = 24.7 \text{ pp}; 95\% \text{ CI: } 2.36 \text{ to } 47.04 \text{ pp}; P = 0.031$] and health-specific domain [interaction effect health risk: $\beta = 22.2 \text{ pp}; 95\% \text{ CI: } 0.25 \text{ to } 44.15 \text{ pp}; P = 0.048$].

Second, we estimate Equation 2 with a continuous and standardize measure of risk tolerance. Results displayed in the last two rows of the Supplementary Figure A.8. The estimation results suggest that one standard deviation increase in general risk tolerance increases the treatment effect by 14.3 [95% CI: 3.5 to 25.0 pp; $P < 0.01$]. Similarly, one standard deviation increase in risk tolerance in health-related realm increases the treatment effect by 10.4 pp [95% CI: 0 to 21.2 pp; $P = 0.058$].

**Other moderators.** Finally, for completeness, we present here the estimated interaction terms of the treatment with the other moderators collected in our surveys. In our pre-analysis plan (RCT ID: AEARCTR-0005644), we specified several key moderators affecting people’s perceptions and preferences towards going out: with socio-demographics (age, gender, having kids or not, education, income), economic preferences (risk preference, time preference, altruism), knowledge about coronavirus, and community trust. In total, we consider 13 potential moderators and denote each moderator with a dummy variable indicating high or low value in that specific dimension based on their median values, except for categories with specific meanings like education (i.e., college educated), gender, having kids at home.

Our pre-analysis plan includes three additional variables that we did not include in our moderator analysis for the following reasons. First, it includes a measure of objective risk based on the proximity of cases to the home location of participants as additional moderator. However, the cases were already zero for several weeks when the experimental phase of the study began, preventing us from exploring the heterogeneity in treatment effects along this dimension. Finally,
Figure A.8: Moderating effect of risk preferences

Note: The figure presents the results of differences in treatment effects over risk preferences. Blue and red dots show the results for general and health related risk preferences respectively. Error bars reflect 95% confidence intervals. The sample is restricted to those subjects whose prior beliefs were below the true percentage (i.e., underestimated number of neighbors planning to go to restaurants).

i) The first two rows display the results of the subgroup analysis in which the main regression (Equation 1) was run separately for risk averse (below median) and risk tolerant (above median) subjects.

ii) The last two rows presents the results of the analysis based on interaction terms, as described in Equation 2. In the interaction term analysis, we first interact the treatment dummy with a dummy variable taking the value of one if the subject stated a level of risk tolerance above the median of the sample. Secondly, we interacted the treatment dummy with the standardized scale of risk tolerance.

we listed car ownership and anxiety about coronavirus as moderators. However, the answers of participants were mostly concentrated in one specific answer (i.e. most of our participants had a car and had around middle level of anxiety towards coronavirus), hindering the decomposition of the effects with enough participants in each subgroup.

To estimate the heterogeneous treatment impacts moderated by all these variables, we first decompose individuals into two sub-samples according to whether they have higher or lower prior belief (i.e., the value of $AboveTruth_{it}$ in Eq[1]). Second, we run 13 separate regressions for each moderator by estimating the following modified version of Equation 4:

$$Y_i = \alpha T_i + \beta^k T_i \times M_{ik} + \gamma^k M_{ik} + X_i \Gamma + \epsilon_i$$

(4)

Here $M_{ik}$ denotes the value of the $k^{th}$ moderator for individual $i$. Our parameter of interest
\( \beta^k \), denotes whether people who have higher value in the \( k^{th} \) moderator respond significantly differently to our treatment. Again, we restrict the analysis to the week when people enter our experiment for the first time to avoid the confounding impacts of the previous weeks’ information treatment.

Among participants with prior beliefs below the truth, only individuals with high (above median) level of risk tolerance in the general and health-specific domains increased their restaurant visits. Besides risk aversion, education also acts as an important moderator. Low educated (no college education) did not adjust their behavior when learning that more neighbors than they believed planned to dine out, while college graduates were 30.59 percentage points [95% CI: 2.98 to 55.22; \( P = 0.015 \)] more likely to adjust behaviors in response to the treatment.

For people who overestimated the percentage of their neighbors going out, prosociality stands out as the only significant moderator (Figure A.9b). More prosocial participants (above median) reduced their visits to both restaurants and parks when learning that fewer neighbors than they believe are planning to visit such places. In contrast, people with weak (below median) social preferences did not adjust their behaviors. However, when including the moderator as a (continuous) standardized variable the interaction term with the treatment dummy becomes non-significant for participants with prior beliefs above the truth (Figure A.10a).

**Figure A.9: Heterogeneous treatment effects on restaurant visits (dummy moderators)**

**a.** Subsample with prior beliefs below the truth **b.** Subsample with prior beliefs above the truth

Note: The figure displays the estimated coefficients of the interaction term between the moderator dummy and treatment indicator in Eq. 4. The effects of all moderators are obtained through separate regressions. Dots describe point estimates, and the darker (lighter) bars describe the 90% (95%) confidence intervals.

A.5 Normative channels and experimenter demand effects

The impact of descriptive norms on the likelihood to visit a restaurant could be in part due to changing normative expectations. Participants could interpret their own belief errors (once corrected on Friday) as somehow meaning that they should, normatively, act in the direction
Figure A.10: Heterogeneous treatment effects on restaurant visits (continuous moderators)

a. Subsample with prior beliefs below the truth

b. Subsample with prior beliefs above the truth

Note: The coefficients display the interaction term between the continuous indicators of moderators (standardized) and treatment indicator in Eq. 4. The effects of all moderators are obtained through separate regressions. Dots describe point estimates, and the darker (lighter) bars describe the 90% (95%) confidence intervals.

of the updated beliefs. Participants’ beliefs could reflect the degree to which it is seen as right, wrong, moral, fair, etc. to go to restaurants, and our intervention might motivate subjects to act by changing these beliefs (33). In addition, the main outcome variable is self-reported, adding scope for experimenter demand effects - i.e., when survey responses are influenced by what individuals in the sample think the research team wants them to say.

We test whether the treatment group is more likely than the control group to believe that the research team wanted them to go (or not go) to restaurants (or parks) over the weekend. This allows us to test whether the descriptive norm changes individual beliefs about what is a desirable behavior. Table A.4 shows the comparison between treatment and control group in the answers to the question “Do you think the researchers have a particular inclination in this research?”. The results show that participants in the treatment and control group do not differ in their belief regarding the intentions of the research team to encourage or discourage people to go to restaurants or parks.

Finally, we also looked for evidence that prosociality moderates belief and behavior change, which would suggest that our descriptive norm intervention was operating by changing normative expectations (33). However, we found no such evidence. The treatment effects do not differ significantly between individuals with high and low levels of prosociality in our subsample of individuals with prior beliefs below the truth (Supplementary Figure A.9 and Figure A.10).
Table A.4: Impact descriptive norm on beliefs regarding experimenter intentions between treatment and control group

|                                           | (1) | (2) | (3)  |
|------------------------------------------|-----|-----|------|
|                                           | Mean| Mean| p-value |
| Treatment                                | 0.49| 0.45| 0.47 |
| Control                                  | 0.50| 0.50|      |
| I do not think they have a particular inclination | 0.21| 0.20| 0.85 |
| Yes, wanted to encourage people to go to restaurants or parks in the post pandemic period | 0.21| 0.18| 0.62 |
| I am not sure                            | 0.07| 0.15| 0.06*|
|                                          | 0.26| 0.36|      |

Notes: The table shows the answers to the question “Do you think the researchers have a particular inclination in this research?”, included in a questionnaire at the end of the experimental phase. Standard errors are displayed in square brackets. Column (1) displays the average for the treatment group, Column (2) displays the average for the control group, and Column (3) displays the p-value associated with a t-test of differences in mean between treatment and control group. *** p<0.01, ** p<0.05 and * p<0.1.

The sample is restricted to participants whose prior beliefs are below the true percentage of neighbors planning to visit a restaurant in their neighborhood (i.e. those who responded to the descriptive norm intervention by increasing their restaurant visits). The results for the full sample of participants are consistent with those in using only the individuals with beliefs below the truth. In the full sample, there are no differences between control and treatment group in their beliefs regarding the intentions of the experimenters in the study (p-value > 0.4).
B Extra Manipulation I: Restaurant Certification

As described in the main text, we tested an additional intervention which highlights the measures of restaurants to minimize the transmission of the virus. This section describes the intervention in detail, and displays the main results, together with the interaction with our main treatment (descriptive norm). The results show that the intervention generated a positive, but non-significant effect on restaurant visits.

B.1 Construction information treatment

On February 20th 2020, the largest restaurant portal in China, Dianging.com (Chinese version of Yelp), created a voluntary certification program (The Meituan Safe Dining Certificate) for restaurants to signal clients that they take extensive precautionary measures to minimize the risk of infection to COVID-19 among their customers.

Restaurants need to fulfill the "Standards for Online Disclosure of Health Service Information of Catering Merchants" published by Dianping, take pictures, and upload to the Dianping to go through the reviewing process. After getting approved, the certification will appear on the main page of Dianping.com in the format of a yellow tag. Restaurants can lose their certification status if they stop sending the daily reports to Dianping, or a customer is filing a complaint about the safety of the restaurant. The measures that restaurants must take to get the certificate include the disinfection dining space, kitchen, and restaurant equipment daily; temperature testing and recording for all employees, allowing them to work only if their temperature is normal; ensuring that all employees must wear face-mask and gloves at all times; and have appropriate table location setting the appropriate distance between dining tables and chairs (See below the full list of requirements to be "Meituan Safe Dining Certificate").

We constructed a database of the 5,000 most popular restaurants in the city, including their location, opening and safe-dining certification status. In our sample, 18% of restaurants in the city are certified, remaining stable over the five weeks of the experiment. Every Friday, during the experiment, individuals in the treatment group were informed about the presence of certified restaurants in the neighborhood (i.e. urban district) of their place of residence through the app. In addition, we constructed a website where participants in the treatment group could see the name of restaurants in the city, their location, and whether the restaurant was certified as a "Meituan Safe Dining Certificate" at the time of the visit (Figure B.1).

Requirements or obtain "Meituan Safe Dining Certificate":

(i). Publish disinfection record: Disinfect dining space on a daily basis and publish disinfection records.
(ii). Temperature testing for all members: Temperature testing and recording for all employees and only allow working if temperature is normal.

(iii). All members must wear face-mask: All employees must wear face-mask; When serving the customers, both sides need to wear face-mask

(iv). Appropriate table location setting: Appropriate distance between dining tables and chairs. Tidy dining environment.

(v). Clean kitchen environment: Kitchen utensils such as cooking equipment, dining equipment, and cloth must be clean and properly disinfected.

(vi). Isolated food serving: Food servers must wear face-masks and gloves, and dishes need to be covered by a lid or plastic sheet.

(vii). Disinfect equipment.

Before our experiment, Dianping had already made the certificate publicly available on its website. The information is thus publicly accessible to all our participants. However, our treatment reduces the information processing costs and increases the attention to the safety features of restaurants in the post-pandemic period. The results from the end of the survey questionnaire indicate that the intervention was indeed successful in providing new information to participants. In the final week, we include a set of questions to test how many of our participants knew about the ”Meituan Safe Dining Certificate”. The results indicate that only 24% percent of the control group knew about the certificate (Table B.1). In the treatment group, 75% of individuals were aware of the certificate used to construct our treatment.
Table B.1: Knowledge about safe-dining certificate used in the experiment for the construction of information treatment

|                         | (1) Mean/SD | (2) Mean/SD | (3) t-test |
|-------------------------|-------------|-------------|-----------|
| Treatment               |             |             |           |
| Use Meituan to get safe dining information (1=Yes) | 0.36 [0.48] | 0.30 [0.46] | 0.23      |
| Know Meituan safe dining certificate (1=Yes)       | 0.74 [0.44] | 0.24 [0.43] | 0.00***   |
| Observations            | 204         | 205         |           |

Notes: Column 1 and 2 reports mean values, with standard deviations in brackets. Column (3) the t-statistic of differences in means between the control and treatment group. Asterisks indicate coefficient statistical significance level: 

* $P < 0.10$; ** $P < 0.05$; *** $P < 0.01$.

B.2 Estimated effects of restaurant certifications on individual behavior

The information treatment providing the information about restaurant certification did not increase significantly the propensity of individuals in our sample to visit a restaurant (Table B.2):

Table B.2: Treatment Effects Restaurant Certification

|                         | (1) | (2) | (3) | (4) | (5) | (6) |
|-------------------------|-----|-----|-----|-----|-----|-----|
| Restaurant Certificate  | 0.055 | 0.062 | 0.045 | 0.060 | 0.060 | 0.064 |
|                          | (0.041) | (0.043) | (0.057) | (0.057) | (0.066) | (0.068) |
| Descriptive Norm         | 0.125** | 0.120** | -0.015 |
|                          | (0.056) | (0.055) | |
| Observations             | 489 | 489 | 272 | 272 | 217 | 217 |
| R-squared                | 0.184 | 0.188 | 0.157 | 0.172 | 0.231 | 0.231 |

Notes: Robust standard errors in parentheses. ** $p < 0.01$, * $p < 0.05$, $p < 0.1$. Columns (1)-(3) presents the estimation results for the full sample sample of participants; Column (3)-(4) individuals whose prior beliefs where below the true percentage of neighbors planning to go to restaurants ,used to construct the descriptive norm. Column (5)-(8) present the results for the sample of individuals whose prior beliefs where above the true percentage of neighbors planning to go to restaurants ,used to construct the descriptive norm.
C Extra manipulation II: Social Preferences, Perceived Norms and Precautionary Behaviors

C.1 Construction treatment

We pre-registered (RCT ID: AEARCTR-0005644) an additional intervention that was not reported in the main text. For completeness, we included a description of the intervention and the main effects on behavior in this section. In week 5, we randomly relocated half of the individuals in experimental arm 2 and experimental arm 4 to a new treatment which expands the descriptive norm with information about the precautionary measures of neighbors to protect each other against the virus, and includes a message to encourage the participant to do the same.

In the last week of the study, we expanded the original descriptive norm on the treatment group with a message describing the stated efforts that their neighbours were taking to protect each other (e.g., wearing face-masks and keeping safe distance in public spaces, and doing temperature checks before going out when feeling uncomfortable, below).

Questionnaire:

We asked both participants in our sample and 400 individuals recruited from the survey company WJX. All responses are based on a 7 point Likert scale from never to always.

(i). I wear a face-mask when I go to the restaurants and will only take it off when I am eating food.

(ii). I use hand sanitiser before I touch anything in the supermarket, restaurant or street.

(iii). When I am in the street or parks and I encounter another person, I try to avoid being too close and keep a safe distance of 6 feet (1.5 meters).

(iv). I check my temperature whenever I feel uncomfortable, and I do not leave the house if I have fever.

(v). I try to go to the park or restaurant at a time when I expect few people to be there.

Table C.1 displays the summary statistics for each question separately. The table shows that most individuals in our sample were undertaking precautionary measures on a frequent basis.

Wording Information Treatment. We added to our main treatment, the following text highlighting the prosociality of neighbors:

"The neighbors in your urban district are working together to protect each other from the Coronavirus so that everyone in the society can enjoy a normal leisure life. The statistics say virtually everyone confirms that they do not go out if they feel sick, and they wear face-masks
Table C.1: Summary statistics questions

|                              | Mean | SD  | % Responses ≥ 5 | N  |
|------------------------------|------|-----|------------------|----|
| Wearing face mask            | 6.17 | 1.33| 88.21%           | 992|
| Use hand sanitiser           | 4.56 | 1.83| 54.72%           | 992|
| Maintain safe distance       | 5.41 | 1.48| 72.58%           | 992|
| Test temperature and not go out | 6.19 | 1.18| 89.72%           | 992|
| Change time going out        | 5.67 | 1.34| 79.94%           | 992|

whenever they are outside the house. The mass majority also report maintaining safe distance with others in public space, wearing face-masks in restaurants as long as not eating, and going out at a time when less people are expected to be there.”

The message made salient that the vast majority of their neighbors were executing important precautionary measures, emphasizing that they were doing so to protect others, and encouraging participants to do the same.

Figure C.1: Display treatment information descriptive norms
Panel a. Original version treatment
Panel b. New version treatment
We estimate the effects of the information treatment using two specifications, separately for parks and restaurants:

$$Y_{it} = \alpha_C T_{Prosocial_i} + \delta AboveTruth_{it} + \mu Plan_{it} + X_i \Gamma + \epsilon_{it}$$ (5)

where $Y_{it}$ takes the value of one if individual $i$ reported in our end-week questionnaire to have visited a park or restaurant over the weekend of our last experimental week (i.e. week 5), $T_{Prosocial_i}$ denotes that individual $i$ received the message with the precautionary measures of neighbors, $AboveTruth_{it}$ takes the value of one if individual $i$ reported in our mid-week questionnaire beliefs regarding the proportion of neighbors going to parks or restaurants over the last experimental weekend, correspondingly. $Plan_{it}$ takes the value of one if an individual reports in our mid-week questionnaire to have plans to go to restaurants or parks over the weekend. $X_i$ includes individual controls. The parameter of interest in eq. [5], $\alpha_C$, describes the differences in restaurant visitation in the last experimental week between individuals receiving the prosocial information treatment together with the descriptive norm and those receiving only the original descriptive norm.

Besides cross subject comparisons, we estimate the following model using the within-subject variation separately for restaurants and parks over the weeks of the experiment:

$$Y_{it} = \alpha_W T_{Prosocial_i} + \delta AboveTruth_{it} + \mu Plan_{it} + \lambda_i + \epsilon_{it}$$ (6)

where $Y_{it}$ describes whether individual $i$ visited a restaurant or park in experimental week $t$, and $T_{Prosocial_i}$ denotes that individual $i$ received in the last week of the experiment the message with the precautionary measures of neighbors. $Plan_{it}$ takes the value of one if individual $i$ reports in our mid-week questionnaire to have plans to go to restaurants or parks over the weekend in week $t$. $AboveTruth_{it}$ takes the value of one if individual $i$ reported in our mid-week questionnaire beliefs regarding the proportion of neighbors going to parks or restaurants, correspondingly, over the weekend in week $t$. The regression includes individual fixed effects, $\lambda_i$, controlling for any time invariant characteristic of individuals. Equation [6] exploits the differences in visits over time associated with the change in treatment status for individuals in the prosocial treatment, who switched in the last experimental week (i.e. week 5) from the original descriptive norm treatment to the expanded version included the precautionary measures of their neighbors to protect others. Robust standard errors are clustered at the individual level.
C.2 Estimated effects of intervention precautionary behavior on individual behavior

The results indicate that the intervention did not have any significant impact on the behavior of individuals (See Table C.2).

Table C.2: Treatment effects on restaurant visitation. Manipulation perceived prosociality

| Prosocial Treatment | (1) | (2) | (3) | (4) | (5) | (6) |
|---------------------|-----|-----|-----|-----|-----|-----|
|                     | CROSS | WITHIN | CROSS | WITHIN | CROSS | WITHIN |
|                     | -0.0632 | -0.0709 | -0.224* | -0.147 | 0.0810 | 0.000291 |
|                     | (0.0780) | (0.0906) | (0.114) | (0.138) | (0.118) | (0.118) |
| Observations        | 138 | 969 | 65 | 461 | 73 | 508 |
| Low Prosoc. Sample  | YES | YES | YES | YES | NO | NO |
| High Prosoc. Sample | YES | YES | NO | NO | YES | YES |

Robust standard errors reported in parenthesis. *P < 0.10; **P < 0.05; ***P < 0.01. The sample includes only individuals that received the descriptive norm first four weeks of the experiment. In this group of participants, half of them were re-assigned to the expanded treatment highlighting the precautionary measures of their neighbors. The CROSS columns display the estimation results of the cross sectional regression using only the observations from fifth week of the experiment. The WITHIN columns display the estimation results of regressions with individual fixed-effects including individual fixed effects. The controls in the CROSS models include district fixed effects, week fixed effects and a dummy variable indicating that the participant went at least once a week to a restaurant in the pre-COVID period, age, age squared, gender, a dummy indicating the subject does not have a college degree, a dummy indicating the participant’s income is below the median level, a dummy variable indicating that the individual stated that he was planning to go to a park in our Wednesday questionnaire. The controls in the WITHIN models include a dummy variable indicating that the individual stated that he was planning to go to a park in our Wednesday questionnaire, a dummy variable indicating whether individuals’ beliefs regarding the number of neighbors going to parks over the weekend was above the truth. The last rows indicate the number of observations, whether the sample included the sample of participants with weak (below median), and strong (above median) social preferences. While non-significant for those subjects with stronger social preferences (above median) [95% CI: -8.33, 62.36, P = 0.129], participants with weaker social preferences (below median) significantly reduced dining out [95% CI: -67.01 to -7.57, P = 0.015] after receiving the manipulation. These individuals are those who were initially non-responsive when learning that fewer neighbors than expected were going to restaurants in our main treatment, and were doing less social distancing measures in public space.

D Changes in Risk Perception

In this section we examine how the descriptive norm impacted the level of risk perception of dining out. The question was included in our Friday survey, just after the elicitation of posterior beliefs. The results show the descriptive norm had a marginal impact on the risk perception of individuals. Those subjects with priors below the truth reported higher risk perception after receiving the descriptive norm intervention (p-value = 0.093). We consider this as suggestive evidence that the intervention made individuals perceive restaurants as safer.
Table D.1 presents the estimations from the following regression models.

Column 1 presents the results restricting the sample to the first week of a participant in the experiment:

\[
PostRisk_{i1} = \alpha T_i + \beta T_i \text{AboveTruth}_{i1} + \delta \text{AboveTruth}_{i1} + PreRisk_{i1} \Lambda + X_i \Gamma + \epsilon_i \tag{7}
\]

Where \(PostRisk_{i1}\) describes the reported risk level on the Friday of the first week that participant \(i\) joined the experiment, after receiving the treatment information; \(PreRisk_{i1}\) describes the level of risk that participant \(i\) reported on Wednesday. \(T_i\) is a dummy variable taking a value of one for people who receive the treatment information (i.e., the information of other neighbors’ planned behaviors), and zero otherwise; \(AboveTruth_{i1}\) is a dummy variable indicating an individual reported on Wednesday beliefs regarding proportion of neighbors planning to go to restaurants (or parks) higher than the true percentage, based on the calculations by the research team; \(X_i\) describes the set of control variables (district fixed effects, pre-COVID restaurant frequency, calendar week fixed effects, dummy indicating whether the individual received the information about the restaurant certification and a dummy variable indicating that the individual has a college degree). Sample is restricted to the week individuals joined the experiment.

| Table D.1: Treatment Effects on Risk Perception of dining out | (1) | (2) | (3) |
|-------------------------------------------------------------|-----|-----|-----|
| | Full Sample | Below Truth | Above Truth |
| Descriptive Norm | -0.160* | -0.156* | 0.145 |
| | (0.095) | (0.094) | (0.110) |
| | [0.093] | [0.098] | [0.188] |
| Prior Above Truth | -0.187* | | |
| | (0.100) | | |
| | [0.063] | | |
| Descriptive Norm | 0.316** | | |
| \(\times\) Prior Above Truth | (0.142) | | |
| | [0.027] | | |
| Restaurant Certification | -0.108 | -0.078 | -0.129 |
| | (0.070) | (0.092) | (0.111) |
| | [0.122] | [0.398] | [0.248] |
| Observations | 506 | 279 | 227 |
| R-squared | 0.426 | 0.439 | 0.435 |

Note: Robust standard errors reported in parenthesis. \(\ast P < 0.10; \ast\ast P < 0.05; \ast\ast\ast P < 0.01\).
D.1 Risk aversion and risk perception

Table D.2: Moderating effect of risk aversion on risk perception

| Moderator                          | (1)  | (2)  | (3)  | (4)  |
|------------------------------------|------|------|------|------|
| High vs low general risk tolerance| 0.254| 0.329| 0.44 |      |
| High vs low health risk tolerance  | 0.110| 3.666| 0.976|      |

Note: The table describes how risk aversion moderates the impact of the descriptive norm on risk perception of individuals. We included in the analysis general risk aversion, and risk aversion in the health domain. Column (1) describes the moderator, Column (2) displays the estimated coefficient associated with the interaction between the descriptive norm treatment and the corresponding risk aversion scale, Column (3) displays robust standard errors, and Column (4) includes the associated p-value. The results indicate both coefficients are statistically not different from zero. ∗P < 0.10; ∗∗P < 0.05; ∗∗∗P < 0.01.
E Initial conditions in park and restaurant settings

In the main text, we describe the different impacts of the descriptive norm intervention in restaurants and parks. In this section, we describe in detail the differences in average visits and risk perception at the beginning of the experiment.

E.1 Comparison restaurant and park visits first week relative to Pre-COVID

Table E.1 presents the percentage of individuals in our control group, receiving no descriptive norms nor the restaurant certification, that reported going to restaurants and and parks over the weekend, together with the percentage of individuals reporting planning to go to these two places, and their reported visit frequency in pre-COVID. In our control sample, the percentage of individuals going to restaurants went down by 36.5 percent \((1 - \frac{31.18}{49.10})\) in our first experimental week, relative to the the reported pre-COVID normal visitation patterns. In parks, the percentage went up by 13.9 percent \((\frac{37.63}{33.04} - 1)\) in our first experimental week, relative to the the reported pre-COVID normal visitation patterns.

Table E.1: Comparison visits to parks and restaurants in first week of experiment with pre-COVID rates

|                  | Observations | Percentage |
|------------------|--------------|------------|
| **Restaurant**   |              |            |
| Went over the first weekend | 93           | 31.18      |
| Planned to go    | 112          | 31.25      |
| Went at least once a week before COVID-19 | 112          | 49.10      |
| **Park**         |              |            |
| Went over the first weekend | 93           | 37.63      |
| Planned to go    | 112          | 59.82      |
| Went at least once a week before COVID-19 | 112          | 33.04      |

Notes: The figures in the table are based on data for the first experimental week 1 for the group receiving no information treatment (no descriptive norm, and no information about the restaurant certification).

E.2 Initial risk perception to visit public spaces

In the initial questionnaire, participants were asked to state their perceived risk of visiting 7 different places on a 5-point Likert scale from low to high risk: parks, restaurants, public transport, office, taxi, and cinema. Panel a in Figure E.1 shows the average risk perception in our sample for each place included in our initial questionnaire.

Panel b in Figure E.1 shows the proportion of individuals that state a strictly higher value in
the risk scale for restaurants than for the corresponding place, indicating that going to restaurants is riskier than going to the place described in the bar in the initial questionnaire.

Figure E.1: Initial perceived risk to visit public spaces

Panel a. Average risk perception

Panel b. Percentage participants perceiving Restaurants Riskier than Venue

Notes: i) Panel a: Bars indicate the average risk perception in our sample for each place included in our initial questionnaire. The scales rank from 1 (lowest risk) to 5 (highest risk).
ii) Panel b: Bars indicate the proportion of individuals that state a strictly higher value in the risk scale for restaurants than for the corresponding place, indicating that going to restaurants is riskier than going to the place described in the bar.
F Prior beliefs over weeks

In the main text, we describe that individuals stayed in the sample for several weeks reporting their beliefs and behaviors. Here, we show how the influence of the descriptive norm intervention in the prior beliefs (reported in our Wednesday survey) of the treatment group in the weeks following the first intervention.

Figure F.1: Changes in prior beliefs over experimental weeks

![Figure F.1](image)

Notes: Figure displays the distribution of beliefs regarding the percentage of neighbors planning to go to restaurants over the weekend reported on our Wednesdays surveys separately for control (left) and treatment (right) group, in every week of the experiment. The figure shows how the beliefs of the treatment group are more concentrated around the true percentage of neighbors for week 2 on. In contrast, the distribution of beliefs of the control group is wider and not centered around the truth. This suggests that prior beliefs are influenced by the descriptive norm information in previous weeks.
Supplementary methods: Correspondence survey responses with GPS data

During the experiment, we retrieved participants’ GPS coordinates using the smartphone app designed for the purpose of the study. In this section we present the different tests that we undertook to confirm the validity of our survey measures. First, we explore the correspondence of the GPS coordinates of participant’s smartphones with the restaurant that they reported in our Sunday questionnaire during the time of the weekend that they reported being there (e.g. Sunday afternoon). Second, we test whether the proportion of time the GPS coordinates of participants are at their home location is lower for those individuals who report being at a restaurant at that day/time.

G.1 GPS data collection: Sampling method and limitations

In order to protect the privacy rights of individuals, and comply with the conditions of Ethical Study protocol from the Committee on the Use of Humans as Experimental Subjects at Massachusetts Institute of Technology (Protocol ID 2002000100), participants provide their coordinates on a voluntary basis by giving the app access to their smartphone’s GPS.

In addition, the major phone operating systems in China (Android, Tencent and iOs) implemented a number of restrictions over the past years to protect the privacy and battery power of their devices, making the passive data collection of GPS coordinates increasingly challenging for app developers. To collect data uninterruptedly the producer of the smartphone app needs to reach agreements to each phone manufacturer separately. These agreements tend to be granted to a small number of popular apps (e.g. WeChat), and are not for new unknown apps, like those designed for the purpose of academic research. Therefore, the app designed for the purpose of this study could share the GPS coordinates with the server only when it is running in the background. This requires individuals to open the app regularly (i.e. approximately once a day, but it depends on the exact specifications of the phone) to keep constant transmission of the GPS data. We incentivized the opening of the app with a series of lotteries among those that open the app, and let it run in the background during the weekend. However, due to these technical difficulties, the sample of individuals reporting GPS data is significantly smaller (N=175, 28% of the full sample) than our full sample of participants and therefore we are limited to use the GPS as a validation tool, instead of a behavioral outcome as a main dependent variable.

In addition, in order to save the battery of our participants while ensuring high quality of data, we implemented two GSP sampling frequencies (details displayed in Table G.1).

Finally, to ensure data quality we applied the following filters to the raw GPS data:

- The percentage of GPS data points in a single day for a participant should be larger
Table G.1: GPS Sampling frequency

| Time of the day | Proportion of minutes sampled | Sample Frequency |
|-----------------|------------------------------|-----------------|
| 7 am to 9 am    | 10 out of 10 mins            | 2 seconds       |
| 9 am to 5 pm    | 2 out of 10 mins             | 2 seconds       |
| 5 pm to 7 pm    | 10 out of 10 mins            | 2 seconds       |
| 7 pm to 8 pm    | 2 out of 10 mins             | 2 seconds       |

Notes: The GPS data records are stored in a Mongodb database together with a user’s ID, to connect the records to the survey responses and treatment status of the participant.

than 80%. There are several ways for the cellphone to generate a GPS point: receiving signals from satellites, receiving signal from WIFI, and using inertia measurement unit to infer the location based on the last precise position. These generated discrepancies in precision. According to the description from AMAP (The provider of the cellphone positioning API), satellites usually provide the highest precision since WIFI tend not to be of sufficient quality. To ensure the precision of the data, we included only data that were generated by satellites. We set a quality threshold of 80% to extract the data from days that users provided which contains enough high quality data points to ensure accuracy of latter analysis.

- The number of data points provided for each period should be more than 50% of the maximum data points that can be provided. We defined 3 periods based on the time in a day: 7AM-9AM (Morning), 9AM-17PM (Noon), 17PM-19PM (Evening).

The low number of observations limits the use of the data as dependent variable in statistical analysis. In total, 175 participants provided good quality GPS data, providing sufficient GPS data during a part of a day morning, afternoon or evening in at least one of the days of the experiment.

G.2 Correspondence coordinates reported restaurant visits with individual’s smartphone GPS coordinates

We use the GPS data to validate participants self reported visits. In particular, we investigate whether we can observe participant’s phone GPS coordinates near (i.e. 200 meters) the restaurant they reported in their weekend report. Given that participants provide the GPS on a voluntary basics, we can only validate the users’ self reported visits using the GPS data from participants who provided good GPS data uninterruptedly during the time of the weekend that individual reported their visit to the restaurant - i.e. Friday dinner time, Saturday lunch time, Saturday dinner time, Sunday lunch time or Sunday dinner time.
From the surveys, we constructed a dataset describing whether a participant went to a restaurant over the weekend, the name of the restaurant, the dates, and the time of day (evening or noon) the restaurant was visited. We implemented different filters to the GPS data to ensure the data quality and the correspondence with the time the participant reported he visited the restaurant.

In total, we compiled a dataset with 58 visits to restaurants reported by the 175 participants who managed to provide GPS data of sufficient data quality. We classified each entry into two categories: entries in which (1) the coordinates of the reported restaurant were within 200 meters of the participants’ phone; and entries in which (2) the coordinates of the reported restaurant were farther than 200 meters of the participants’ phone. The results show that 95% (55/58) of entries had a GPS location within 200 meters of the reported location. This indicates that the vast majority of our respondent’s GPS coordinates were within 200 meters of the restaurant they claimed to visit over the weekend, at the time that they reported doing so, supporting the validity of the survey outcome variables.

G.3 Time at home and reported visit to restaurant

Individuals in our sample were asked to provide their home address in the initial questionnaire. During dinner time (5 pm to 7 pm) over the weekend, we calculated the share of time at home of a participant by calculating the proportion of GPS data points within 200 meters of their home coordinates. Table G.2 describes the results of a regression of proportion of time participants’ GPS coordinates are at the home location on a dummy variable taking the value of one if the individual reported to be at a restaurant at that time:

\[ \text{GPSinHomeLocation}_{d,i} = \beta \text{ReportedRestaurantVisit}_{d,i} + \Omega' X_{d,i} + \lambda_d + \mu_i + \epsilon_{d,i} \]  \hspace{1cm} (8)

where \( \text{GPSinHomeLocation}_{d,i} \) describes the time the coordinates of individual \( i \) in day \( d \), during dinner time. \( \text{ReportedVisittoRestaurant}_{d,i} \) takes the value of 1 if individual \( i \) reported going to a restaurant in day \( d \) for dinner; \( \lambda_d \) and \( \mu_i \) are day and individual fixed effects, correspondingly. \( \beta \) is our coefficient of interest describing the differences in time at home for those reporting going to a restaurant in day \( d \) and those who did not report a restaurant visit on day \( d \) and those who did not report a restaurant visit on day \( d \). Standard errors \( \epsilon_{d,i} \) are clustered at the individual level. The list of controls \( (\omega' X_{d,i}) \) includes the education level of individuals, district fixed effects, a dummy variable showing whether the individual went to restaurants at least once a week before the COVID-19 pandemic, a dummy variable indicating that participants’ income is below median, experimental week fixed effects, and a dummy variable indicating whether

---

1We focused on the dinner time, rather than lunch time, since participants are less likely to be at the office
whether the individual received the additional information treatment about the precautionary measures of restaurants.

Table 4 \textsuperscript{G.2} presents different model specifications of Eq. 4.3 In all specifications, with different sets of controls and fixed effects, the results consistently show the time at home is 11 pp \[P<0.05\] lower when participants reported going to restaurants. This provides additional evidence on the correspondence of the GPS and self-reported measures.

Table G.2: Estimated changes in time at home measured by GPS when reporting a visit to a restaurant

|                      | (1) | (2) | (3) | (4) |
|----------------------|-----|-----|-----|-----|
| Time at home         |     |     |     |     |
| Reported Rest. Visit (1=YES) | -0.109** | -0.103** | -0.112** | -0.114** |
|                      | (0.0467) | (0.0457) | (0.0530) | (0.0538) |
| Observations         | 549 | 549 | 549 | 549 |
| R-squared            | 0.094 | 0.099 | 0.118 | 0.118 |
| Day FE               | YES | YES | YES | YES |
| Individual FE        | NO | NO | YES | YES |
| Strata Controls      | NO | YES | NO | YES |
| Extra Controls       | NO | YES | NO | YES |

Notes: Robust standard errors are in parentheses. Asterisks indicate coefficient statistical significance level: \( \ast P < 0.10; \ast \ast P < 0.05; \ast \ast \ast P < 0.01 \). All estimations include day fixed effects. The Strata controls include district fixed effects, a dummy variable showing whether the individual went to restaurants at least once a week before the COVID-19 pandemic. The regression include as "Extra" controls the education level of individuals, age, age square, district fixed effects, a dummy variable showing whether the individual went to restaurants at least once a week before the COVID-19 pandemic, a dummy variable indicating that participants' income is below median, whether the individual received the information treatment about the precautionary measures of restaurants. Sampling time restricted to the evening time on Friday, Saturday and Sunday. The \( R^2 \) of each regressions, displayed at the bottom of the table, is about 10%. The remaining 90% of variation in the dependent variable (% time at home by the individual inferred by GPS) is driven by a number of activities not collected in our weekend diaries (which only asked about the visits to parks and restaurants over the weekend). Visits to any other venue out of the home of participants (e.g. shopping malls, supermarkets, walking in the street, visits to the doctor, working at the office, commuting etc.) will explain this remaining variation in the time at in the sample.
H Supplementary Notes

H.1 Description moderator scales

This subsection describes the wording of the scales to measure individual risk preferences and the rest of moderating scales included in the analysis.

| Variable               | Question                                                                                      | Source |
|------------------------|-----------------------------------------------------------------------------------------------|--------|
| Risk Aversion (General)| Please tell me, in general, how willing or unwilling you are to take risks, using a scale from 0 to 5, where 0 means you are "completely unwilling to take risks" and 5 means you are "very willing to take risks." 0 completely unwilling to take risks - 10 very willing to take risks | 38     |
| Risk Aversion (Health) | One can evaluate different areas of your daily life differently. How do you evaluate your attitude towards risk regarding your health? 0 completely unwilling to take risks - 10 very willing to take risks | 35     |
| Social Preferences     | How willing are you to give to good causes without expecting anything in return?              | 35     |
|                        | 0 completely unwilling - 10 very willing                                                      |        |
| Time preferences       | How willing are you to give up something that is beneficial for you today in order to benefit more from that in the future? | 35     |
| Community Trust        | Generally speaking, would you say that most people can be trusted or that you cannot be too careful in dealing with people? What proportion of people living in your neighborhood do you are trustworthy? | 37     |
| Official Media Trust   | What do you think is the proportion of true information in news from the official media?     |        |
| COVID Knowledge        | Example question: In which of the following channels do you think the coronavirus will spread (multiple choices)? Skin contact, droplets,... |        |

H.2 Display treatment information

This section describes the communication with subjects during the experimental week. The individuals in the control group received the same information pieces as those in the treatment group, except for the descriptive norm information (i.e. percentage of neighbors planning to go to restaurants over the weekend). We included the original display in Chinese as well as the translated version in English (see next page).
Welcome screen: thank you for participating in this follow-up survey. In this follow-up survey, we wish to update you on the plans of our survey participants in your urban district. Recall that you guessed that for participants living in your urban district:

- Out of 100 survey participants intend to go to the restaurant this weekend;
- Out of 100 survey participants intend to go to the park this weekend.

We have analyzed all responses to the survey to give you the true number of people that intend to go to restaurants and parks this weekend. For each 100 people filling the questionnaire living in your urban district,

- Plan to the restaurant this weekend;
- Plan to the park this weekend.

The neighbors in your urban district are working together to protect each other from the Coronavirus so that everyone in the society can enjoy a normal leisure life.

The statistics say virtually everyone wears face masks whenever they are outside the house, and confirms that they do not go out if they feel sick. The mass majority also report maintaining safe distance with others in public spaces, and wearing face masks in restaurants as long as not eating.

When you go out, please continue to protect your neighbours as they are doing for you.

| Text                                                                 | Subsample     |
|----------------------------------------------------------------------|---------------|
| [Wechat]                                                            | All Sample    |
| Welcome screen: thank you for participating in this follow-up survey. In this follow-up survey, we wish to update you on the plans of our survey participants in your urban district. Recall that you guessed that for participants living in your urban district: |               |
| ![emoji Wednesday Response]] out of 100 survey intend to go to the restaurant this weekend; |               |
| ![emoji Wednesday Response]]out of 100 survey participants intend to go to the park this weekend. |               |
| Find out the real: [Link]                                           |               |
| We have analyzed all responses to the survey to give you the true number of people that intend to go to restaurants and parks this weekend. For each 100 people filling the questionnaire living in your urban district, | Descriptive Norm Treatment Group |
| [[Information Descriptive Norm]] plan to the restaurant this weekend. | Experimental arm 2 and 4, See Methods Section |
| For each 100 people filling the questionnaire, [[Information Descriptive Norm]] plan to the park this weekend. |               |
| The neighbors in your urban district are working together to protect each other from the Coronavirus so that everyone in the society can enjoy a normal leisure life. The statistics say virtually everyone wears face masks whenever they are outside the house, and confirms that they do not go out if they feel sick. The mass majority also report maintaining safe distance with others in public spaces, and wearing face masks in restaurants as long as not eating. When you go out, please continue to protect your neighbours as they are doing for you. | Manipulation Prosociality Treatment Group (50% Descriptive Norm group in Week 5, Experimental arm 2 and 4, See Methods Section) |
Remember that we offered you:

We offered you 2 RMB bonus payment for accurately guessing the survey participants living in your urban district who would actually go to a restaurant and go to a park this weekend.

Remember that in the survey of Wednesday, you guessed that:
[embedded individual value] out of 100 survey participants living in your urban district would actually go to the restaurant this weekend;
[embedded individual value] out of 100 survey participants living in your urban district would actually go to the park this weekend;
Perhaps your views about other’s behaviors change.

Please spend one minute to answer this questionnaire so that we can know your thoughts now: [Link to Survey]
[WJX Platform]
Perhaps your views about other’s behaviors change.

Please guess for every 100 people living in your urban district that took our survey, how many will actually go to the restaurants and parks this weekend.
(If your guess is within 2 people of the truth, you will earn a bonus payment of 2 RMB for each)

- Restaurant: [Answer] out of 100 respondents living in my urban district will go to restaurants
- Park: [Answer] out of 100 respondents living in my urban district will go to parks

Thank you for your sharing with us your beliefs.
Figure H.1: Display descriptive norm treatment information and collection of posterior beliefs

**Control Group**

**Treatment Group**

**Note:** The figure displays the Friday-survey for the control (left) and treatment group (right). The only difference between the treatment and control group is that we provided the treatment group the percentage of neighbors planning to go to restaurants and parks over the weekend (in red numbers). The rest of the information was the same: Both groups were reminded of their prior beliefs, and the monetary rewards for the right guests. Finally both, control and treatment group, were requested to fill in their beliefs regarding the proportion of neighbors they thought were going to visit a restaurant and a park in the text boxes at the bottom of the image.
Our research group is examining very regularly Dianping website to give you the most valuable information of restaurant businesses around you. Some restaurants in the urban district are certified by Dianping as “Meituan Safe Dining Restaurants” by today, indicating that they adopted all satisfactory precautionary strategies to prevent COVID-19 virus infection, including frequent disinfection, temperature testing, and maintaining appropriate distance between tables, etc.

Restaurants need to continuously work hard to maintain the certificate. They must send photos to Meituan everyday to prove that they satisfy all the seven safe dining criteria (listed in the picture below). Customers can report to the Meituan platform whenever they detect a violation, and the restaurant will get removed from the list.

How risky do you feel about dining out at a nearby restaurant this weekend? (7-level Likert scale)

For the next questions, we would like you to think about the current situation in your urban district:

How risky do you feel about going to a nearby park this weekend? (7-level Likert scale)

We have created a list of restaurants with certification status of thousands of restaurants in the entire city.

Do you want us to share the most updated information about the restaurants over the weekend?

Yes

No

Here is the link to the website. [Link]

Thank you very much for your participation. We will inform you about the total payment you have earned from Wednesday and today’s surveys – the payment will be deposited to your virtual account via the SXCX APP.

We will also inform you about future study opportunities, and we look forward to seeing you again soon! Feel free to contact us via WeChat «Add WeChat ID» if you have questions regarding this study.

Have a good weekend and don’t forget to set up your phone to participate in the weekend lottery!
H.3 Summary statistics and balance tests

This section presents the summary statistics and shows the balance tests across treatment arms by regressing baseline individual characteristics on indicators for the three treatment arms and the strata variables, and conduct F-tests that all treatment coefficients equal zero.

For the 38 variables analyzed, F-statistics are significant at 5 percent for only two variables, namely, a binary variable indicating whether the respondent had no college degree, and the trust in social media. Controlling for those variables do not change the significance nor the magnitude of our main estimates. As it is a key socio-demographic variable, we include education as a control in all regressions in our sample.

Finally, an overall F-test in a seemingly unrelated regression (SUR) specification across the 38 regressions yields a p-value on the F-statistic of 0.68, indicating that we cannot reject the hypothesis of equality across all of the treatment arms and control group, indicating that the randomization created largely comparable groups.

Column 1 in \[H.1\] reports mean values for the control group, with standard deviations in brackets. Columns 2 to 4 report the coefficients from separate regressions in which a dependent variable is regressed on the full set of treatment indicators and stratification variables (i.e., district, and a dummy variable that indicates that the individual went to the restaurant at least once a week in the pre-pandemic period on average). Column 5 reports the p-values of F-tests of whether the treatment coefficients are jointly equal to zero. An overall F-test in an SUR specification across the 38 regressions yields a p-value on the F-statistic of 0.68, indicating that we cannot reject the hypothesis of baseline equality across all of the treatment arms and control group.
### Table H.1: Balance tests

|                          | (1) |   | (2) |   | (3) |   | (4) |   | (5) |   | p-value | F-test |
|--------------------------|-----|---|-----|---|-----|---|-----|---|-----|---|---------|--------|
| Control Descriptive     |     |   | 31.12 | 0.91 | 1.38 | 0.09 | 0.26 |
| Norm                     |     |   | [7.12] | (0.81) | (0.83) | (0.75) |
| Restaurant. Desc. Norm   |     |   | 0.92 | 0.04 | 0.02 | 0.06 | 0.25 |
| Certificate AND Rest. Cert |     |   | [0.27] | (0.03) | (0.03) | (0.03) |
| Low Education (1=Yes)    |     |   | 0.27 | 0.11** | -0.03 | 0.02 | 0.07* |
| [0.45]                  |     |   | (0.05) | (0.05) | (0.05) |
| Low Income (1=Yes)       |     |   | 0.56 | -0.01 | 0.02 | 0.05 | 0.70 |
| [0.50]                  |     |   | (0.06) | (0.06) | (0.06) |
| Owns Car (1=Yes)         |     |   | 0.77 | -0.01 | 0.03 | -0.05 | 0.52 |
| [0.42]                  |     |   | (0.05) | (0.06) | (0.06) |
| Prior Proportion Rest. Visitors |     |   | 37.58 | -2.52 | -10.96*** | -0.73 | 0.00*** |
| [23.79]                 |     |   | (2.67) | (2.55) | (2.76) |
| Prior Proportion Park Visitors |     |   | 51.62 | -1.39 | -6.05*** | 0.82 | 0.08* |
| [24.43]                 |     |   | (2.69) | (2.90) | (2.77) |
| Planning Visit Park (1=Yes) |     |   | 0.56 | 0.04 | 0.03 | -0.08 | 0.14 |
| [0.50]                  |     |   | (0.06) | (0.06) | (0.06) |
| Planning Visit Rest. (1=Yes) |     |   | 0.35 | 0.08 | -0.00 | 0.00 | 0.35 |
| [0.48]                  |     |   | (0.05) | (0.05) | (0.05) |
| Risk Shopping            |     |   | 3.66 | 0.05 | 0.08 | 0.02 | 0.90 |
| [0.96]                  |     |   | (0.11) | (0.11) | (0.11) |
| Office Risk              |     |   | 2.98 | 0.11 | 0.04 | 0.01 | 0.68 |
| [0.94]                  |     |   | (0.10) | (0.11) | (0.11) |
| Movie Risk               |     |   | 4.34 | -0.02 | 0.09 | 0.00 | 0.59 |
| [0.74]                  |     |   | (0.09) | (0.08) | (0.08) |
| Taxi Risk                |     |   | 3.18 | 0.17 | 0.16 | 0.04 | 0.23 |
| [0.91]                  |     |   | (0.10) | (0.11) | (0.10) |
| Park Risk                |     |   | 2.72 | 0.07 | 0.13 | -0.01 | 0.55 |
| [0.93]                  |     |   | (0.10) | (0.11) | (0.10) |
| Restaurant Risk          |     |   | 4.10 | 0.12 | 0.24*** | 0.11 | 0.07* |
| [0.81]                  |     |   | (0.09) | (0.09) | (0.09) |
| Trust General            |     |   | 1.67 | 0.02 | 0.04 | 0.01 | 0.95 |
| [0.62]                  |     |   | (0.07) | (0.07) | (0.07) |
| Trust Family             |     |   | 4.74 | -0.10 | -0.07 | -0.15** | 0.09* |
| [0.54]                  |     |   | (0.06) | (0.05) | (0.06) |
| Trust Colleagues         |     |   | 3.64 | -0.04 | -0.02 | -0.07 | 0.79 |
| [0.63]                  |     |   | (0.07) | (0.07) | (0.07) |
| Trust Neighborhood       |     |   | 2.94 | -0.04 | 0.03 | -0.07 | 0.60 |
| [0.67]                  |     |   | (0.07) | (0.08) | (0.08) |
| Trust Strangers          |     |   | 2.19 | -0.14 | -0.18 | -0.03 | 0.16 |
| [0.82]                  |     |   | (0.09) | (0.09) | (0.10) |

Robust standard errors are in parentheses.

Asterisks indicate coefficient statistical significance level: * P < 0.10; ** P < 0.05; *** P < 0.01.
Table H.1 (cont.). Balance Tests

|                                | (1)     | (2)     | (3)     | (4)     | (5)     |
|--------------------------------|---------|---------|---------|---------|---------|
|                                | Control | Descriptive | Restaurant. | Desc. Norm | p-value | AND Rest. Cert | F-test |
| Prop. Neighbors Trustworthy    | 3.56    | 0.01    | 0.06    | 0.02    | 0.96    |
|                                | [1.08]  | (0.12)  | (0.12)  | (0.13)  |         |
| Trust Money Neighbors          | 1.71    | -0.15   | -0.08   | -0.01   | 0.17    |
|                                | [0.63]  | (0.07)  | (0.07)  | (0.08)  |         |
| Trust Money Stranger           | 1.25    | -0.08   | -0.06   | 0.05    | 0.14    |
|                                | [0.50]  | (0.06)  | (0.06)  | (0.06)  |         |
| Trust TV                       | 4.41    | 0.04    | -0.05   | 0.04    | 0.84    |
|                                | [0.98]  | (0.10)  | (0.12)  | (0.11)  |         |
| Trust Official Media           | 4.41    | 0.03    | -0.02   | 0.03    | 0.97    |
|                                | [0.93]  | (0.10)  | (0.11)  | (0.10)  |         |
| Trust Websites                 | 2.90    | 0.08    | 0.06    | 0.07    | 0.88    |
|                                | [0.97]  | (0.11)  | (0.11)  | (0.11)  |         |
| Trust Weibo                    | 3.16    | -0.05   | -0.13   | -0.01   | 0.76    |
|                                | [1.08]  | (0.12)  | (0.13)  | (0.12)  |         |
| Trust Wechat                   | 2.72    | 0.33*** | 0.20    | 0.23    | 0.02**  |
|                                | [0.97]  | (0.11)  | (0.11)  | (0.11)  |         |
| Trust Wechat Official Account  | 3.26    | 0.23*   | 0.05    | 0.05    | 0.26    |
|                                | [1.11]  | (0.12)  | (0.13)  | (0.12)  |         |
| Trust Zhihu                    | 3.01    | -0.03   | -0.15   | -0.13   | 0.54    |
|                                | [1.09]  | (0.12)  | (0.12)  | (0.12)  |         |
| COVID Quiz Score               | 9.14    | -0.01   | 0.36    | 0.18    | 0.29    |
|                                | [2.03]  | (0.23)  | (0.22)  | (0.24)  |         |
| Norm Distancing                | 3.36    | 0.07    | 0.23*   | 0.25**  | 0.11    |
|                                | [1.12]  | (0.12)  | (0.12)  | (0.12)  |         |
| Norm Protect Others            | 4.77    | -0.03   | -0.02   | 0.02    | 0.81    |
|                                | [0.50]  | (0.06)  | (0.06)  | (0.06)  |         |
| Risk Preference General        | 3.45    | -0.12   | 0.28    | -0.29   | 0.39    |
|                                | [2.81]  | (0.31)  | (0.32)  | (0.32)  |         |
| Risk Preference Health         | 3.24    | -0.25   | 0.10    | -0.02   | 0.71    |
|                                | [2.65]  | (0.30)  | (0.30)  | (0.30)  |         |
| Time Preference Long Reward    | 7.03    | 0.06    | -0.08   | 0.02    | 0.97    |
|                                | [2.59]  | (0.29)  | (0.30)  | (0.29)  |         |
| Charity Preference             | 6.43    | -0.15   | 0.00    | 0.00    | 0.95    |
|                                | [2.68]  | (0.30)  | (0.30)  | (0.30)  |         |
| Time Preference Delay          | 3.67    | -0.31   | -0.14   | -0.01   | 0.80    |
|                                | [3.16]  | (0.35)  | (0.35)  | (0.36)  |         |

Robust standard errors are in parentheses.

Asterisks indicate coefficient statistical significance level: * P < 0.10; ** P < 0.05; *** P < 0.01.
H.4 Recruitment strategy and sample attrition

This section describes the procedure to recruit subjects, and provides an overview of the refreshment samples collected over the experimental weeks. In the main text, we report the results based on the first week individuals are in the sample (i.e. column (1)), where individuals’ priors are not being shaped by the descriptive norms in previous weeks.

The majority of individuals in our initial contact lists were extracted from a list of participants in a survey conducted in Zhengzhou July 2019 by the research team on air pollution and commuting behavior. Individuals voluntarily left their cell phone number for future research projects.

In addition, we complemented the list with a series of ads in the Chinese social network Wechat, which broadly advertised the research project as studying the “impact of Covid-19”. We sent out recruitment text messages with an APP installation link and rewards for participating illustrated. If they agree to participate, the screening questionnaires and informed consent are included during the APP registration.

Out of 5,047 phone calls, 475 individuals never responded to the call. 1228 accepted to install the app and participate in the experiment. 609 individuals did not fill any questionnaire, leaving a total of 619 valid participants. Table H.1 for the number of individuals in the sample in each week of the experiment.

| Experimental Week | New Comers | Return Sample | Total |
|-------------------|------------|---------------|-------|
| Week 1            | 450        | 0             | 450   |
| Week 2            | 81         | 406           | 487   |
| Week 3            | 47         | 474           | 521   |
| Week 4            | 12         | 496           | 508   |
| Week 5            | 32         | 344           | 376   |
H.5 Percentages used to construct descriptive norm

Table H.2 displays the percentages used to construct the descriptive norm information treatments.

Table H.2: Percentage individuals planning to visit restaurants in urban district in sample

|                 | District 1 | District 2 | District 3 | District 4 | District 5 |
|-----------------|------------|------------|------------|------------|------------|
| Panel a. Percentage neighbors planning to visit restaurants |            |            |            |            |            |
| Week 1          | 30         | 45         | 29         | 32         | 40         |
| Week 2          | 36         | 35         | 37         | 32         | 39         |
| Week 3          | 39         | 47         | 25         | 40         | 39         |
| Week 4          | 38         | 51         | 41         | 32         | 40         |
| Week 5          | 41         | 33         | 30         | 51         | 49         |

|                 | District 1 | District 2 | District 3 | District 4 | District 5 |
|-----------------|------------|------------|------------|------------|------------|
| Panel a. Percentage neighbors planning to visit parks     |            |            |            |            |            |
| Week 1          | 53         | 51         | 49         | 51         | 51         |
| Week 2          | 61         | 62         | 45         | 47         | 53         |
| Week 3          | 50         | 59         | 44         | 54         | 57         |
| Week 4          | 41         | 56         | 38         | 38         | 48         |
| Week 5          | 52         | 49         | 33         | 55         | 55         |
This section describes the monetary compensations that participants received for the completion of each task requested in the study.

| Item                                      | Monetary Compensation | Condition                        | Time                        |
|-------------------------------------------|-----------------------|----------------------------------|-----------------------------|
| Survey 1                                  | 20 RMB                | Completed                        | Start from Mar 22           |
| Survey 2                                  | 15 RMB                | Completed                        | Start from Mar 29           |
| Prior belief (weekly)                     | 2 RMB                 | Completed by Thursday night      | Weekly Wed-Thurs            |
| Treatment and posterior belief (weekly)    | 2 RMB                 | Completed by Friday night        | Weekly Friday               |
| Weekend actual behavior (weekly)          | 2 RMB                 | Completed                        | Weekly Sunday night         |
| Pop-up happiness questions                | 1 RMB                 | Completed                        | Weekly on Saturday          |
| End survey                                | 10 RMB                | Completed                        | Start from May 11           |
| Correct belief rewards                    | 1 RMB                 | Correctly guessed neighbor percentage (within 2 percentage points to the truth) | Start from March 30. All rewards for correct guesses were given at the end of the experiment. |
| GPS data provision                        | 2 RMB/day             | If a user provide any data, money will be top up his account | Since March 16th |
| GPS Weekend Lottery                       | 1000 RMB for two participants that provide valid GPS over the weekend | Participants who provide enough data for Friday night(30 mins), Saturday noon and night(30 mins each, 60 mins in total) and Sunday noon and night(30 mins each, 60 mins in total) can enter the lottery | Since March 30th |
References

[1] Solomon Hsiang, Daniel Allen, Sebastien Annan-Phan, Kendon Bell, Ian Bolliger, Trinetta Chong, Hannah Druckenmiller, Andrew Hultgren, Luna Yue Huang, Emma Krassovich, Pely Lau, Jaecheol Lee, Esther Rolf, Jeanette Tseng, and Tiffany Wu. The Effect of Large-Scale Anti-Contagion Policies on the Coronavirus (COVID-19) Pandemic. Nature, page 2020.03.22.20040642, jan 2020.

[2] Seth Flaxman, Swapnil Mishra, Axel Gandy, H Juliette T Unwin, Thomas A Mellan, Helen Coupland, Charles Whittaker, Harrison Zhu, Tresnia Berah, Jeffrey W Eaton, Christl A Donnelly, Steven M Riley, Michaela A C Vollmer, Neil M Ferguson, and Lucy C Okell. Estimating the effects of non-pharmaceutical interventions on COVID-19 in Europe. Nature, 2020.

[3] World Bank. Global Economic Prospects. Technical Report June, Washington, DC, 2020.

[4] BEA. Gross Domestic Product, 2nd Quarter 2020 (Advance Estimate) and Annual Update, 2020.

[5] Austan Goolsbee and Chad Syverson. Fear, lockdown, and diversion: Comparing drivers of pandemic economic decline 2020. Journal of Public Economics, 193:104311, 2021.

[6] Raj Chetty, John N Friedman, Nathaniel Hendren, and Michael Stepner. How did covid-19 and stabilization policies affect spending and employment? a new real-time economic tracker based on private sector data. Technical report, National Bureau of Economic Research, 2020.

[7] Kathleen Dooling. The advisory committee on immunization practices. updated interim recommendation for allocation of covid19 vaccine.united states, december 2020, 2020.

[8] The Times of Israel. Israel aims to vaccinate 25% of population in a month; 250,000 had shots so far, 2020.

[9] Melissa Eddy and Marc Santora. E.U. Starts Effort to Vaccinate 450 Million, 2020.

[10] Biden-Harris Transition. The Biden-Harris plan to beat COVID-19, 2020.

[11] Thiemo Fetzer, Lukas Hensel, Johannes Hermle, and Christopher Roth. Coronavirus perceptions and economic anxiety. The Review of Economics and Statistics, pages 1–36, 2021.

[12] Jeffrey V Lazarus, Scott C Ratzan, Adam Palayew, Lawrence O Gostin, Heidi J Larson, Kenneth Rabin, Spencer Kimball, and Ayman El-Mohandes. A global survey of potential acceptance of a covid-19 vaccine. Nature medicine, pages 1–4, 2020.

[13] Thiemo R Fetzer, Marc Witte, Lukas Hensel, Jon Jachimowicz, Johannes Haushofer, Andriy Ivchenko, Stefano Caria, Elena Reutskaja, Christopher P Roth, Stefano Fiorin, Margarita Gomez, Gordon Kraft-Todt, Friedrich M Gotz, and Erez Yoeli. Global behaviors and perceptions at the onset of the covid-19 pandemic. National Bureau of Economic Research Working Paper, 2020.

[14] Austan Goolsbee. Twitter. https://twitter.com/austangoolsbee/status/1347548098978062336, 08-January 2021.

[15] Nancy H.L. Leung, Daniel K.W. Chu, Eunice Y.C. Shiu, Kwok Hung Chan, James J. McDevitt, Benien J.P. Hau, Hui Ling Yen, Yunguo Li, Dennis K.M. Ip, J. S,Malik Peiris, Wing Hong Seto, Gabriel M. Leung, Donald K. Milton, and Benjamin J. Cowling. Respiratory virus shedding in exhaled breath and efficacy of face masks. Nature Medicine, 26(5):676–680, 2020.

[16] H. Wesley Perkins and Alan D Berkowitz. Perceiving the community norms of alcohol use among students: Some Research implications for campus alcohol education programming*. Substance Use and Misuse, 21(9-10):961–976, 1986.

[17] Michael Hallsworth, John A List, Robert D Metcalfe, and Ivo Vlaev. The behavioralist as tax collector: Using natural field experiments to enhance tax compliance. Journal of public economics, 148:14–31, 2017.

[18] Noah J Goldstein, Robert B Cialdini, and Vladas Griskevicius. A room with a viewpoint: Using social norms to motivate environmental conservation in hotels. Journal of consumer Research, 35(3):472–482, 2008.
[19] Kimberly S. Wolske, Kenneth T. Gillingham, and P. Wesley Schultz. Peer influence on household energy behaviours. *Nature Energy*, 5(3):202–212, 2020.

[20] Robert C. Dempsey, John McAlaney, and Bridgette M. Bewick. A critical appraisal of the social norms approach as an interventional strategy for health-related behavior and attitude change. *Frontiers in Psychology*, 9(NOV):1–16, 2018.

[21] J.J.; et al VanBavel. Using social and behavioural science to support COVID-19 pandemic response. *Nature Human Behaviour*, pages 1–50, 2020.

[22] Michael Muthukrishna, Thomas J H Morgan, and Joseph Henrich. The when and who of social learning and conformist transmission. *Evolution and Human Behavior*, 37(1):10–20, 2016.

[23] John Beshears, James J Choi, David Laibson, Brigitte C Madrian, and Katherine L Milkman. The effect of providing peer information on retirement savings decisions. *The Journal of Finance*, 70(3):1161–1201, 2015.

[24] Christina Annette Gravert and Linus Olsson Collentine. When nudges aren’t enough: Incentives and habit formation in public transport usage. *CEBI WORKING PAPER SERIES*, Working Paper 10/19, ISSN 2596-44TX, 2019.

[25] Ariella S Kristal and Ashley V Whillans. What we can learn from five naturalistic field experiments that failed to shift commuter behaviour. *Nature Human Behaviour*, 4(2):169–176, 2020.

[26] Mark A Andor, Andreas Gerster, Jörg Peters, and Christoph M Schmidt. Social norms and energy conservation beyond the us. *Journal of Environmental Economics and Management*, 103, 2020.

[27] P Wesley Schultz, Jessica M Nolan, Robert B Cialdini, Noah J Goldstein, and Vladas Griskevicius. The constructive, destructive, and reconstructive power of social norms. *Psychological Science*, 18(5):429–434, 2007.

[28] Eli P. Fenichel. Economic considerations for social distancing and behavioral based policies during an epidemic. *Journal of Health Economics*, 32(2):440–451, 2013.

[29] Eli P. Fenichel, Carlos Castillo-Chavezb, M. G. Ceddiac, Gerardo Chowellb, Paula A. Gonzalez Parraea, Graham J. Hickling, Garth Holloway, Richard Horan, Benjamin Morin, Charles Perrings, Michael Springborn, Leticia Velazquez, and Cristina Villalobos. Adaptive human behavior in epidemiological models. *Proceedings of the National Academy of Sciences of the United States of America*, 108(15):6306–6311, 2011.

[30] Seth Benzell, Avinash Collis, and Christos Nicolaides. Rationing Social Contact During the COVID-19 Pandemic: Transmission Risk and Social Benefits of US Locations. *Proceedings of the National Academy of Sciences*, pages 2–4, 2020.

[31] John A. List, Azeem M. Shaikh, and Yang Xu. Multiple hypothesis testing in experimental economics. *Experimental Economics*, 22(4):773–793, 2019.

[32] Gordon Kraft-Todd, Erez Yoeli, Syon Bhanot, and David Rand. Promoting cooperation in the field. *Current Opinion in Behavioral Sciences*, 3:96–101, 2015.

[33] David G Rand, Erez Yoeli, and Moshe Hoffman. Harnessing reciprocity to promote cooperation and the provisioning of public goods. *Policy Insights from the Behavioral and Brain Sciences*, 1(1):263–269, 2014.

[34] Heejung Kim and Hazel Rose Markus. Deviance or uniqueness, harmony or conformity? a cultural analysis. *Journal of personality and social psychology*, 77(4):785, 1999.

[35] Richard E Nisbett and Takahiko Masuda. Culture and point of view. *Proceedings of the National Academy of Sciences*, 100(19):11163–11170, 2003.

[36] Michele J Gelfand, Joshua Conrad Jackson, Xinyue Pan, Dana Nau, Dylan Pieper, Emmy Denison, Munegith Dogher, Paul AM Van Lange, Chi-Yue Chiu, and Mo Wang. The relationship between cultural tightness–looseness and covid-19 cases and deaths: a global analysis. *The Lancet Planetary Health*, 2021.
[37] Jon M. Jachimowicz, Salah Chafik, Sabeth Munrat, Jaideep C. Prabhu, and Elke U. Weber. Community trust reduces myopic decisions of low-income individuals. Proceedings of the National Academy of Sciences of the United States of America, 114(21):5401–5406, 2017.

[38] Armin Falk, Anke Becker, Thomas Dohmen, Benjamin Enke, David Huffman, and Uwe Sunde. Global evidence on economic preferences. The Quarterly Journal of Economics, 133(4):1645–1692, 2018.

[39] Davide Cantoni, David Yang, Noam Yuchtman, and Jane Zhang. Protests as strategic games: Experimental Evidence from Hong Kong’s Authoritarian Movement. The Quarterly Journal of Economics, 134(2):1021–1077, 2019.