# Sample

We collected our data through an online survey using Google Forms. We e-mailed undergraduate, graduate and professional students of a large, private, not-for-profit university in the city of São Paulo, Brazil in May of 2020. The survey was open for 1 month for respondents and has 55 items and we provided no incentives and disclosed that it would help the university’s research group to understand media consumption and behavior during the challenging times of COVID-19. We also supplied the estimated time that it would take to respond the full survey (around 10 minutes). According to both the university’s and Brazilian ethical guidelines, since we are no intrusive questions, nor incentives to fill out the survey and respondents were invited to participate with the option to decline or drop out of the questionnaire at any time, it was not necessary IRB approval.

From the 55 survey’s items we used only 34 items to measure all of our variables:

* demographic variables: age (categorical 5-points) and gender;
* self-efficacy: 5 items using 5-point Likert scale to measure self-efficacy (gustavo we need say something about what items we used, I do not know the scale);
* fear: 2 items using 4-point intensity scale to measure how much respondents are afraid of COVID-19 infection (gustavo we need say something about what items we used, I do not know the scale);
* total media exposure: 1 item using a 4-point intensity scale regarding daily media consumption about COVID-19;
* media type: 4 types of media (newspaper, television, social media and health professionals) using a 5-item intensity scale regarding the frequency of media usage regarding media consumption about COVID-19; and
* protective behaviors: 20 items using 5-point Likert scale to measure the adoption of protective behaviors during COVID-19 (gustavo we need say something about what items we used, I do not know the scale).

Our sample is comprised of 7,554 respondents and the summary statistics are detailed in table X\_summarystats. The first column depicts the variable name with columns for central tendencies (mean and median) and dispersion measures (standard deviation, 1st quartile Q1 and 3rd quartile Q3) along with the minimum and maximal values. The age variable was discretized in 5 categories with respect to the age in years: (1) below 17; (2) between 18 and 30 (3) between 31 and 50; (4) between 51 and 70; and (5) over 70. In our sample, 28 percent were men and 72 percent were women. In terms of age distribution, 65 percent were between 18 and 30 years old, and 30 percent were between 31 and 50 years old. Thus, at least 95 percent of our sample was not in the over-60 category, which is the group with the greatest risk of contracting COVID-19. Regarding media type consumption, our sample prefers to consume television, followed by social media and medical professionals. Newspaper is the least preferable media type that the sample consumes.

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Variable | Mean | Standard Deviation | Minimum | Q1 | Median | Q3 | Maximum |
| Age | 2.36 | 0.58 | 1 | 2 | 2 | 3 | 5 |
| Sex Male | 0.25 | 0.45 | 0 | 0 | 0 | 1 | 1 |
| Self-Efficacy | 0.41 | 0.65 | -2 | 0 | 0.4 | 0.8 | 2 |
| Fear | 2 | 0.88 | 0 | 1.5 | 2 | 2 | 3 |
| Total Media Exposure | 1.58 | 0.78 | 1 | 1 | 1 | 2 | 4 |
| Media Type – Television | 2.7 | 1.3 | 0 | 2 | 3 | 4 | 4 |
| Media Type – Newspaper | 1.85 | 1.56 | 0 | 0 | 2 | 3 | 4 |
| Media Type – Social Media | 2.2 | 1.46 | 0 | 1 | 2 | 3 | 4 |
| Media Type – Health Professionals | 2.1 | 1.45 | 0 | 1 | 2 | 3 | 4 |
| Protective Behaviors | 2.8 | 0.55 | 0 | 2.57 | 2.91 | 3.19 | 3.81 |

Table X\_summarystats – Sample Summary Statistics.

# Variables

We present our variables and whenever possible the α between parenthesis is the Cronbach’s alpha for reliability measures for all items of the referred variable. Figure X\_framework illustrates our framework, hypotheses and variables.

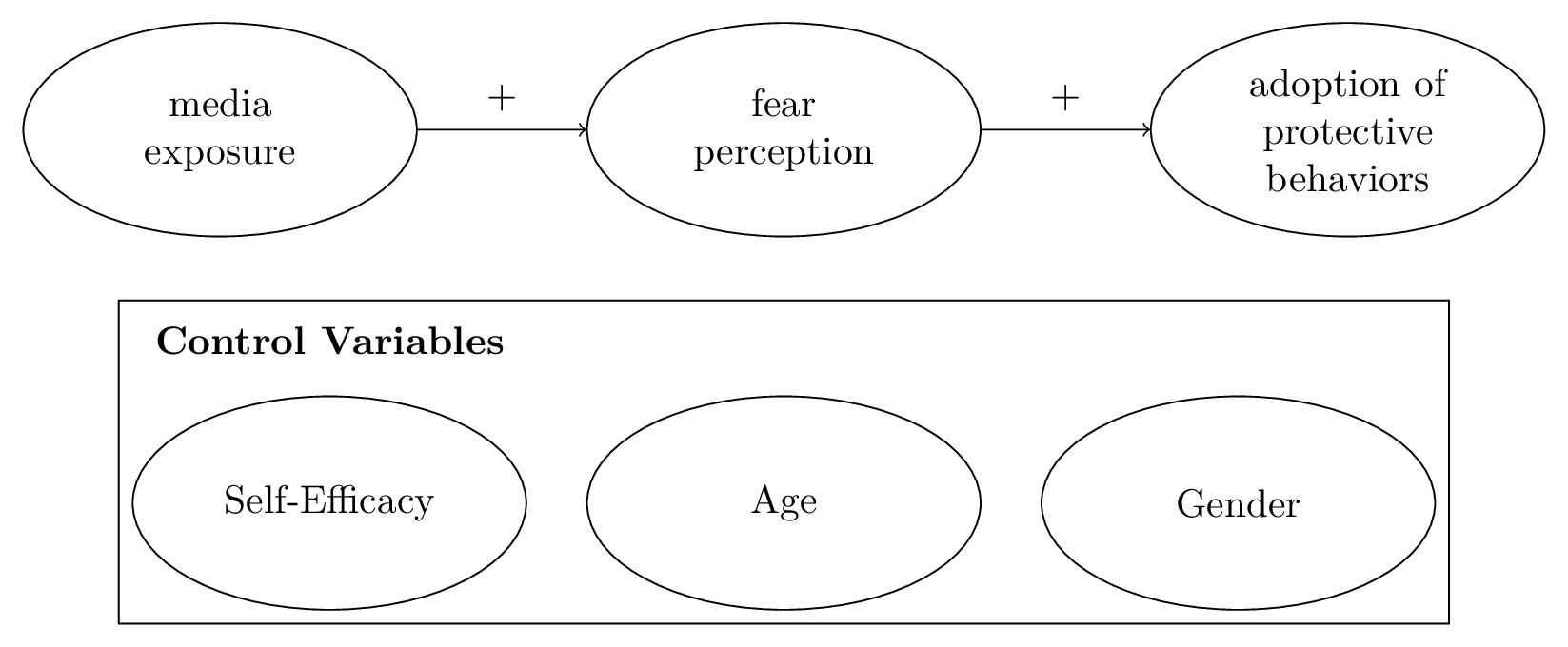
Our dependent variable is the adoption of protective measures. We asked respondents to indicate on a 5-point Likert scale from 0 to 4 how frequently they engage in protective behaviors. The answers range from “Never” to “Always”, where “Never” was coded as 0 and “Always” as 4. Examples of the items are “wash your hands with soap and water,” “avoid touching your mouth and nose with your hands,” “cough in your elbow,” “maintain at least a meter of distance from other people,” “avoid visiting friends and family members not living with you” and “put on a face mask when going outside” (α = .89).

Our independent variable is media exposure. It was measured on 4-point intensity scale that asked respondents to indicate how much time they spend daily consuming media regarding COVID-19. It had the following values: 1 for “Less than 1 hour”, 2 for “1-2 hours”, 3 for “3-5 hours” and 4 for “More than 5 hours”. We also asked respondents for media type consumption regarding COVID-19 information for 4 different media types: television, newspaper, social media and health professionals. We measured media type preferences with a 5-point Likert scale in the same manner as the protective behaviors, i.e. from 0 (“Never”) to 4 (“Always”).

Our mediator variable is fear. We asked respondents two questions regarding how much they are afraid of either being infected or having a close relative being infected by COVID-19. It was measured on a 4-point intensity scale which ranges from 0 ( “I am not afraid”) to 3 (“Very afraid”) (α = 0.77).

We controlled for self-efficacy asking 5 questions about … (gustavo need help). We measured the responses using 5-point Likert scale that ranges from -2 (“Strongly Disagree”) to 2 (“Strongly Agree”) while 0 was the neutral point (“Neither Disagree or Agree”) (α = 0.51).

We introduced two socio-demographic variables as control variables. Age, despite being ordinal, was measured as a continuous variable. Gender was measured as a dummy variable with women coded 1 and men coded 0.

Figure X\_framework – Model Framework.

# Data and Software

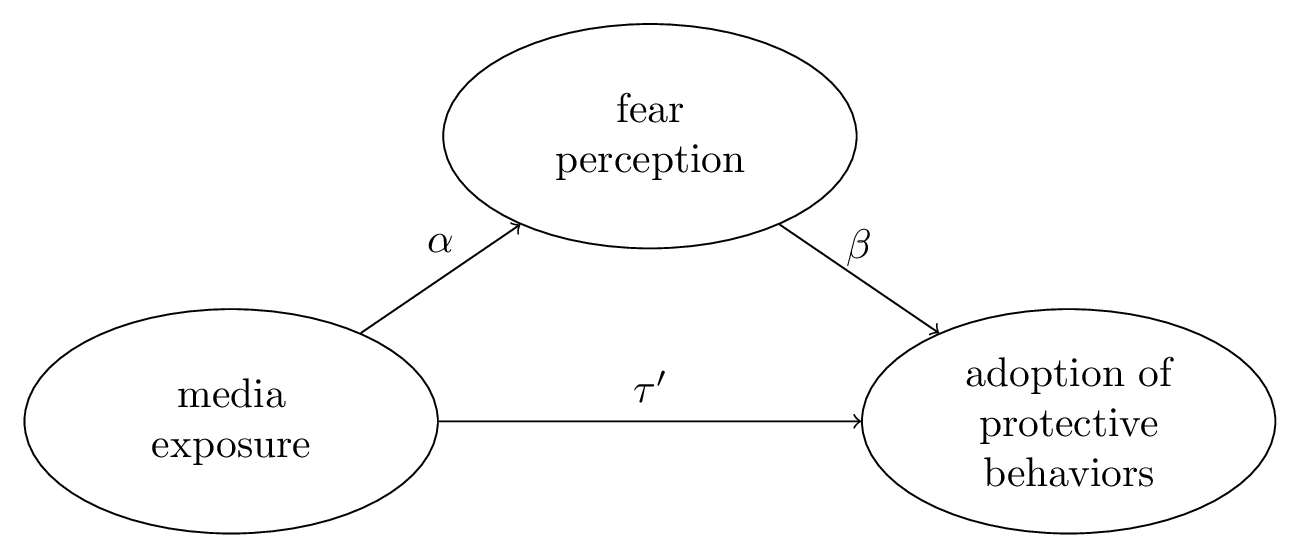
All of our data and code is available in a public repository at GitHub (<https://github.com/LabCidades/COVID-Media>). We used solely Julia language (Bezanson et al., 2017) version 1.6. To read, manipulate and prepare the data for analyses we used the package DataFrames.jl version 1.2. For the images and visualizations we used packages Makie.jl with Cairo backend and AlgebraOfGraphics.jl both versions 0.6. Cronbach’s alphas were calculated using StatsBase.jl version 0.33. Finally, the Bayesian modeling and Markov chain Monte Carlo sampling were conducted using Turing.jl (Ge, Xu & Ghahramani, 2018) version 0.19.

All parameter estimates and regression models were estimated using a Bayesian approach (Gelman et al., 2013; van de Schoot et al., 2021). Specifically we used weakly informative priors and used the state-of-the-art Markov chain Monte Carlo sampler with No-U-Turn (NUTS) algorithm available at Turing.jl with 4 chains and 2,000 iterations using the first 1,000 as warmup. To assess convergence and reliability in our parameter values we used the R̂ (rhat) statistics that measures how well the Markov chains have mixed. More specifically, it measures the within- and between-chain variance. Values close to 1 are desired, and a threshold of 1.01 is widely used in the literature to indicate convergence (Gelman et al., 2013; van de Schoot et al., 2021).

Finally, in all of our models we standardized the variables to a canonical/standard normal distribution, i.e. mean 0 and standard deviation 1. This transformation has two advantages. First, it makes for the sampler to estimate parameter values and sample the model, thus decreases the model run time considerably. Second, since our variables have different scales with some being measured as 4-point and others 5-point scales, it makes easier and more intuitive to compare effect sizes across variables since they are standardized to the same scale, which effectually makes the comparisons using standard deviations.

# Mediation Test

We tested our mediation using a single-level mediation model (Yuan and MacKinnon, 2009; VanderWeele, 2016). Informally, this is also known as the “Sobel test”. We isolate the three variables: dependent, independent and moderator. In our case the dependent is adoption of protective behaviors, independent is media exposure and mediator is fear perception. The mediation test proceeds as following and depicted in figure X\_mediationtest. We estimate concurrently in a regression model the direct effects (τ) and the indirect effects (α and β). Mathematically the indirect effects in the product between the effect of the independent variable onto the mediator (α) and the effect of the mediator onto the dependent variable (β). The total effects is the sum of the direct and indirect effects. Finally, we measure the strength of the mediation by comparing how much of the total effects are indirect versus direct.

Figure X\_mediationtest – Mediation Test.

For the mediation test model we used for all priors a student t distribution with 3 degrees of freedom and 0 as location parameter and 1 as scale parameter (since the variables all have mean 0 and standard deviation 1). This applies to both intercepts for the mediator (fear perception) and the dependent (adoption of protective behaviors) variables, and also the effects coefficients’ of independent variable onto the dependent variable (τ), independent variable onto mediator (α), and mediator variable onto dependent variable (β) we also used as prior a student t distribution with 3 degrees of freedom and 0 as location parameter and 1 as scale parameter. For the measurement errors related to the Gaussian/normal likelihood in the linear regression for the mediator (fear perception) and the dependent (adoption of protective behaviors) we used a exponential distribution with rate 1 as a prior.

Mathematically, the mediation model can be stated as following:

12§latex§$$
\begin{aligned}
\alpha_{\text{med}} &\sim \text{Student}(3, 0, 1) \\
\alpha_{\text{dep}} &\sim \text{Student}(3, 0, 1) \\
\beta_{\text{indep\_dep}} &\sim \text{Student}(3, 0, 1) \\
\beta_{\text{indep\_med}} &\sim \text{Student}(3, 0, 1) \\
\beta_{\text{med\_dep}} &\sim \text{Student}(3, 0, 1) \\
\sigma_{\text{med}} &\sim \text{Exponential}(1) \\
\sigma_{\text{dep}} &\sim \text{Exponential}(1) \\
\text{med} &\sim \text{Normal}(\alpha_{\text{med}} + \text{indep} \cdot \beta_{\text{indep\_med}}, \sigma_{\text{med}}) \\
\text{dep} &\sim \text{Normal}(\alpha_{\text{dep}} + \text{indep} \cdot \beta_{\text{indep\_dep}} + \text{med} \cdot \beta_{\text{med\_dep}}, \sigma_{\text{dep}}) \\
\text{direct} &= \beta_{\text{indep\_dep} \\
\text{indirect} &= \beta_{\text{indep\_dep}} \cdot  \beta_{\text{dep\_dep} \\
\text{total} &= \text{indirect} + \text{indirect}
\end{aligned}
$$
§png§600§TRUE§

Table X\_mediationtest shows the direct, indirect and total effects estimates. For each effect we estimated the mean and median central tendencies, the standard deviation as a dispersion measure, and also the 95% density interval using the 2.5% and 97.5% percentiles. We can see that the indirect effects are much higher than the direct effects which indicates a strong mediation effect of fear between media exposure and adoption of protective behaviors.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Effects | Mean | Standard Deviation | Median | 2.5% percentile | 97.5 percentile |
| Direct | -0.034 | 1.549 | -0.016 | -3.319 | 3.027 |
| Indirect | 0.042 | 0.003 | 0.042 | 0.035 | 0.048 |
| Total | 0.008 | 1.549 | 0.026 | -3.278 | 3.07 |

Table X\_mediationtest – Mediation Test Effect Estimates.

# Model

As with our mediation test, we also used weakly informative priors for the parameters of our model. First, both the mediator (fear perception) and dependent (adoption of protective behaviors) variables have an intercept with a prior as a student t distribution with 3 degrees of freedom and 0 as location parameter and 1 as scale parameter. The coefficients for the independent, mediator and control variables have as prior also a student t distribution with 3 degrees of freedom and 0 as location parameter and 1 as scale parameter. For the measurement errors related to the Gaussian/normal likelihood in the linear regression for the mediator (fear perception) and the dependent (adoption of protective behaviors) we used a exponential distribution with rate 1 as a prior.

Mathematically, the regression model can be stated as following:

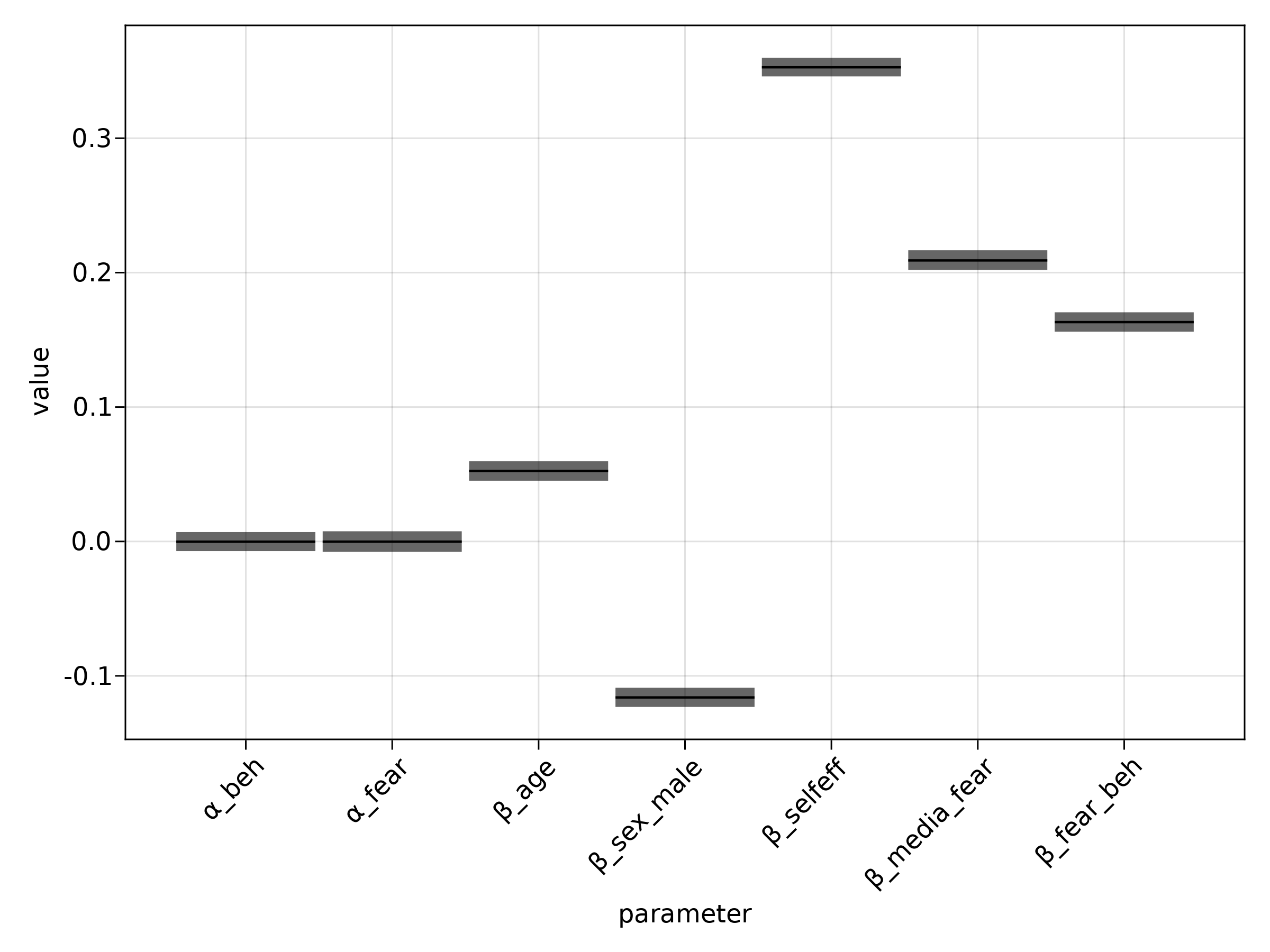
12§latex§$$
\begin{aligned}
\alpha_{\text{med}} &\sim \text{Student}(3, 0, 1) \\
\alpha_{\text{dep}} &\sim \text{Student}(3, 0, 1) \\
\beta_{\text{indep\_med}} &\sim \text{Student}(3, 0, 1) \\
\beta_{\text{med\_dep}} &\sim \text{Student}(3, 0, 1) \\
\beta_{\text{control}} &\sim \text{Student}(3, 0, 1) \\
\sigma_{\text{med}} &\sim \text{Exponential}(1) \\
\sigma_{\text{dep}} &\sim \text{Exponential}(1) \\
\text{med} &\sim \text{Normal}(\alpha_{\text{med}} + \text{indep} \cdot \beta_{\text{indep\_med}} + \text{control} \cdot \beta_{\text{control\_med}}, \sigma_{\text{med}}) \\
\text{dep} &\sim \text{Normal}(\alpha_{\text{dep}} + \text{med} \cdot \beta_{\text{med\_dep}} + \text{control} \cdot \beta_{\text{control\_dep}} , \sigma_{\text{dep}})
\end{aligned}
$$
§png§600§TRUE§

Table X\_model shows for all estimated parameters the mean and median central tendencies, the standard deviation as a dispersion measure, and also the 95% density interval using the 2.5% and 97.5% percentiles. Our intercepts for both the mediator (α\_med) and dependent variables (α\_dep) are estimated to be 0 which is expected since we standardized all variables to mean 0 and standard deviation 1. For the effects of the independent variable (media exposure) onto the mediator variable (fear perception) the estimated parameter (β\_indep\_med) has a positive association with 95% density interval between 0.19 and 0.23, which means that keeping all other effects fixed we shall expect an increase/decrease of around 0.2 standard deviation in fear perception for every increase/decrease of 1 standard deviation in media exposure. The parameter (β\_med\_dep) that measures the effect of the mediator variable (fear perception) onto the dependent variable (adoption of protective behaviors) has also a positive 95% density interval between 0.15 and 0.18, which means that keeping all other effects fixed we shall expect an increase/decrease of around also 0.2 standard deviation in adoption of protective behaviors for every increase/decrease of 1 standard deviation in fear perception. Finally, the control variables we see that age has a positive association and sex male a negative association with adoption of protective behaviors; one thing to note is the profound positive association between self-efficacy and adoption of protective behaviors: the 95% density interval spans from 0.35 to 0.37.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Parameter | Mean | Standard Deviation | Median | 2.5% percentile | 97.5% percentile |
| α\_med | 0 | 0.01 | 0 | -0.021 | 0.021 |
| α\_dep | 0 | 0.01 | 0 | -0.02 | 0.02 |
| β\_indep\_med | 0.216 | 0.011 | 0.216 | 0.194 | 0.237 |
| β\_med\_dep | 0.163 | 0.011 | 0.163 | 0.142 | 0.184 |
| β\_control\_med\_age | -0.083 | 0.011 | -0.083 | -0.105 | -0.061 |
| β\_control\_med\_sex\_male | -0.184 | 0.011 | -0.1845 | -0.206 | -0.163 |
| β\_control\_med\_selfeff | 0.051 | 0.011 | 0.051 | 0.029 | 0.072 |
| β\_control\_dep\_age | 0.053 | 0.011 | 0.053 | 0.032 | 0.073 |
| β\_control\_dep\_sex\_male | -0.116 | 0.011 | -0.11 | -0.123 | -0.096 |
| β\_control\_dep\_selfeff | 0.353 | 0.01 | 0.353 | 0.333 | 0.373 |
| σ\_med | 0.955 | 0 | 0.949 | 0.94 | 0.97 |
| σ\_dep | 0.907 | 0 | 0.903 | 0.894 | 0.922 |

Table X\_model – Model Parameter Estimates.

Also, in Figure X\_model we can see the boxplot of the parameters values with the middle line representing the parameter median value and the boxes’ lower and upper bounds representing the 95% density interval of the parameters value, i.e. the 2.5% and 97.5% percentiles respectively.

Figure X\_model – Boxplot of Model Estimate Parameters.

Also, we tested the effects of different media types consumption regarding COVID-19 information for 4 different media types: television, newspaper, social media and health professionals. We tested using a hierarchical varying-intercept variable where each of the media types is a group-level intercept while also removing the media exposure independent variable. We found that media type has no effect in fear, thus we kept our original model with media exposure as a independent variable.

# **References**

Bezanson, J., Edelman, A., Karpinski, S., & Shah, V. B. (2017). Julia: A fresh approach to numerical computing. *SIAM Review*, *59*(1), 65–98. <https://doi.org/10.1137/141000671>

Ge, H., Xu, K., & Ghahramani, Z. (2018). Turing: A Language for Flexible

Probabilistic Inference. *International Conference on Artificial Intelligence and Statistics*, 1682–1690. <http://proceedings.mlr.press/v84/ge18b.html>

Gelman, A., Carlin, J. B., Stern, H. S., Dunson, D. B., Vehtari, A., & Rubin, D. B. (2013). *Bayesian Data Analysis*. Chapman and Hall/CRC.

van de Schoot, R., Depaoli, S., King, R., Kramer, B., Märtens, K., Tadesse, M. G., Vannucci, M., Gelman, A., Veen, D., Willemsen, J., & Yau, C. (2021). Bayesian statistics and modelling. *Nature Reviews Methods Primers*, *1*(1), 1–26. https://doi.org/10.1038/s43586-020-00001-2

VanderWeele, T. J. (2016). Mediation Analysis: A Practitioner’s Guide. *Annual Review of Public Health*, *37*, 17–32. https://doi.org/10.1146/annurev-publhealth-032315-021402

Yuan, Y., & MacKinnon, D. P. (2009). Bayesian mediation analysis. *Psychological Methods*, *14*(4), 301–322. <https://doi.org/10.1037/a0016972>