Appendix A
[Online Working Paper]

The analysis of our data is undertaken using a Bayesian statistical model that allows us to evaluate the posterior probability that observed distributions of reports are indeed uniform.

The data are in the form of 100 tokens allocated by each subject in response to each question in a given wave. These tokens are allocated across 10 bins, which refer to interval outcomes for COVID-19 infections or deaths. Hence these data may be viewed as ordered, since tokens allocated to lower bins refer to fewer infections or deaths than tokens allocated to higher bins. The number of tokens allocated to each bin may be viewed as a frequency count, so that the implied likelihood of the observed data correctly reflects the intensity of reports about beliefs. We use the familiar ordered logit (or logistic) specification of this data generating process.

The prior we have, by design, is that tokens and beliefs will be allocated uniformly over the 10 bins for each question. This prior arises from our method of determining the intervals for each bin. One exception is for the elicitation of deaths in frame #3 of wave 5, due to an error in the calculation of intervals. Although this error did not radically change the intervals for this frame in comparison to other frames for this question in this wave, those bins did not reflect our uniform prior. Hence the data for this frame and question, in just this wave, are removed from our analysis.

In order to test our null hypothesis, we must define a ROPE that characterizes an interval around the posterior estimates of our model. To do this we undertake pre-estimation simulation of the estimates that we would obtain for the appropriate sample size we observed. These simulations considered random integer-valued allocations of tokens across the 10 bins by each (simulated) subject. For example, random allocations between 9 and 11 would have selected integers from the set {9, 10, 11} for each subject, with an expected average over enough simulations of 10, matching the uniform prior. For each such random allocation around 10, we estimate the model using the simulated data, and evaluate the 95% credible intervals of the model parameters around the estimates that would have been implied if every subject had strictly followed the prior and allocated 10 tokens to every bin. These credible intervals allow us to easily see how “regions of practical equivalence” in token allocations translate into regions of practical equivalence in estimates from the ordered logit model. We

can then repeat these random draws a large number of times, in our case 100 times, and find the ROPE from the largest difference in credible intervals over all bins and simulations. It is then a simple matter to return to the actual, observed data, estimate the ordered logit model, and compute the probability of the posterior estimates being in that ROPE, defined now over the estimates of the model using the observed data.

This approach to generating a ROPE is quite general. It is often the case that one can propose a ROPE directly in terms of the parameters defining the data generating process, or that they are proscribed by rules or regulations. For example, for bioequivalence the Food and Drug Administration recommends ROPE limits of 0.8 and 1.25 for the ratio of two means of different comparison distributions [19]. And ROPE limits are commonly used in actuarial and epidemiological calculations for risk management purposes, often varying with the expected size of the risk: for example, ±20% for moderate risk, ±5% for high risks, and ±50% for low risks [1]. But it is important to be able to undertake pre-posterior simulation when the mapping between "natural limits" for a ROPE and the corresponding limits for the underlying parameter estimates for some data generating process can be highly nonlinear.

Additional References

1. Little TA. Equivalence Testing for Comparability. BioPharm International. 2015;28:45-8.

Appendix B
[Online Working Paper]

This appendix shows the distributions of subject reports about COVID-19 prevalence from waves 2-6 of our study. The main text provides a general discussion of the results from these waves, and how they follow a similar pattern to that which we observed in wave 1.

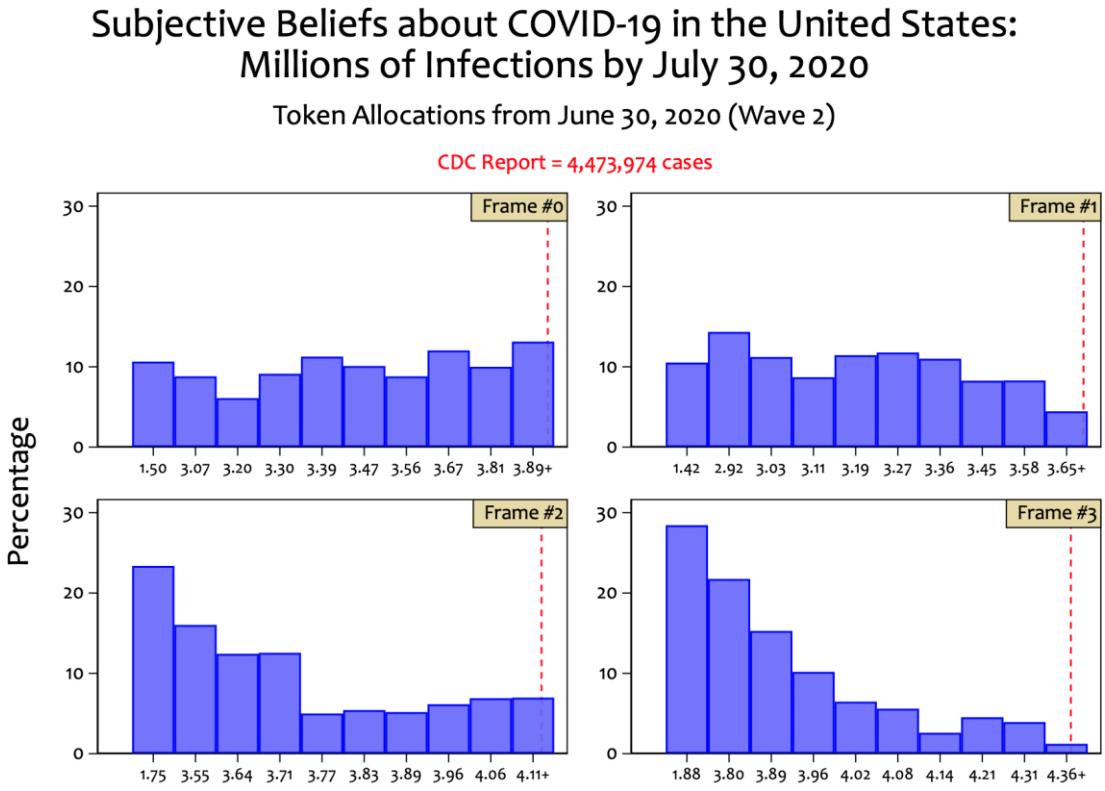

Figure B1: Beliefs about COVID-19 Infections in the U.S. by July 30, 2020

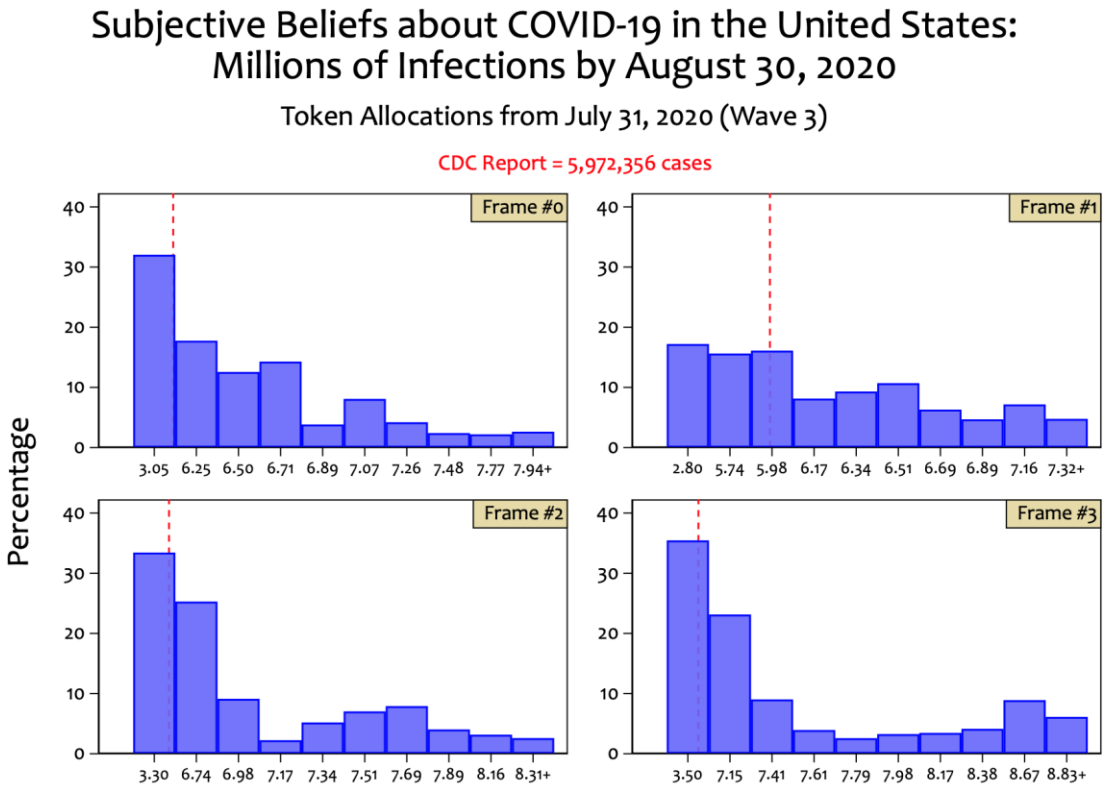

Figure B2: Beliefs about COVID-19 Infections in the U.S. by August 30, 2020

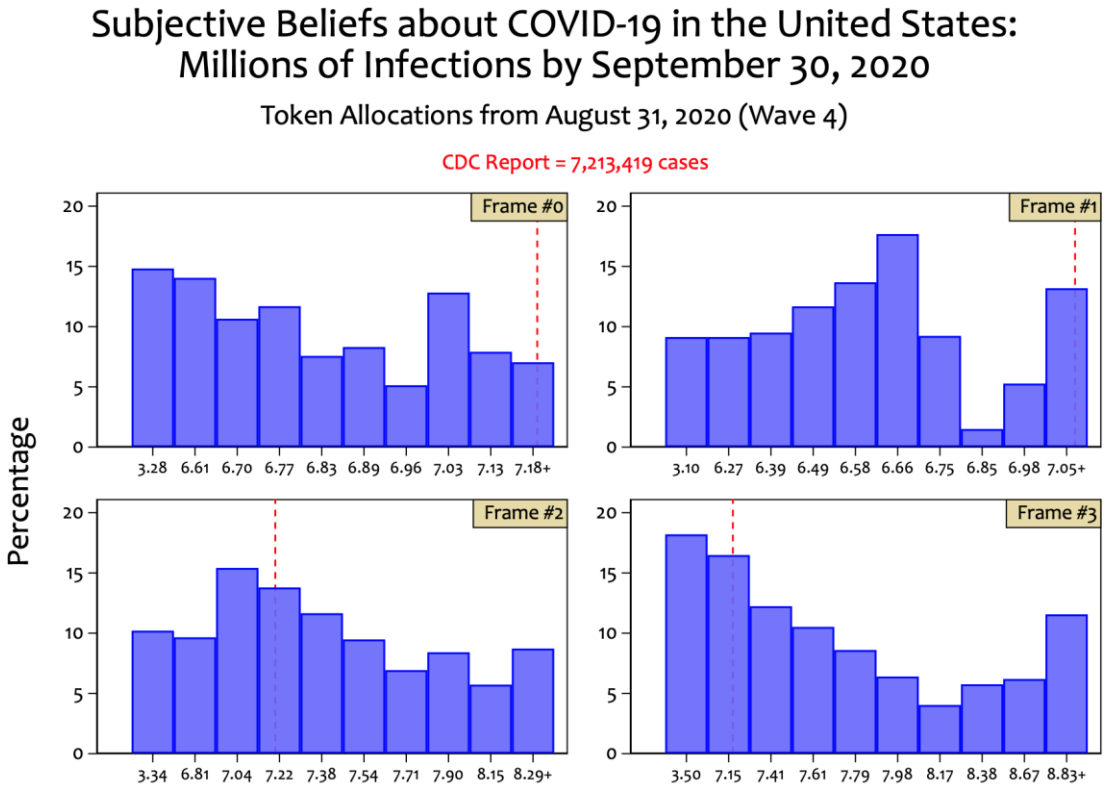

Figure B3: Beliefs about COVID-19 Infections in the U.S. by September 30, 2020

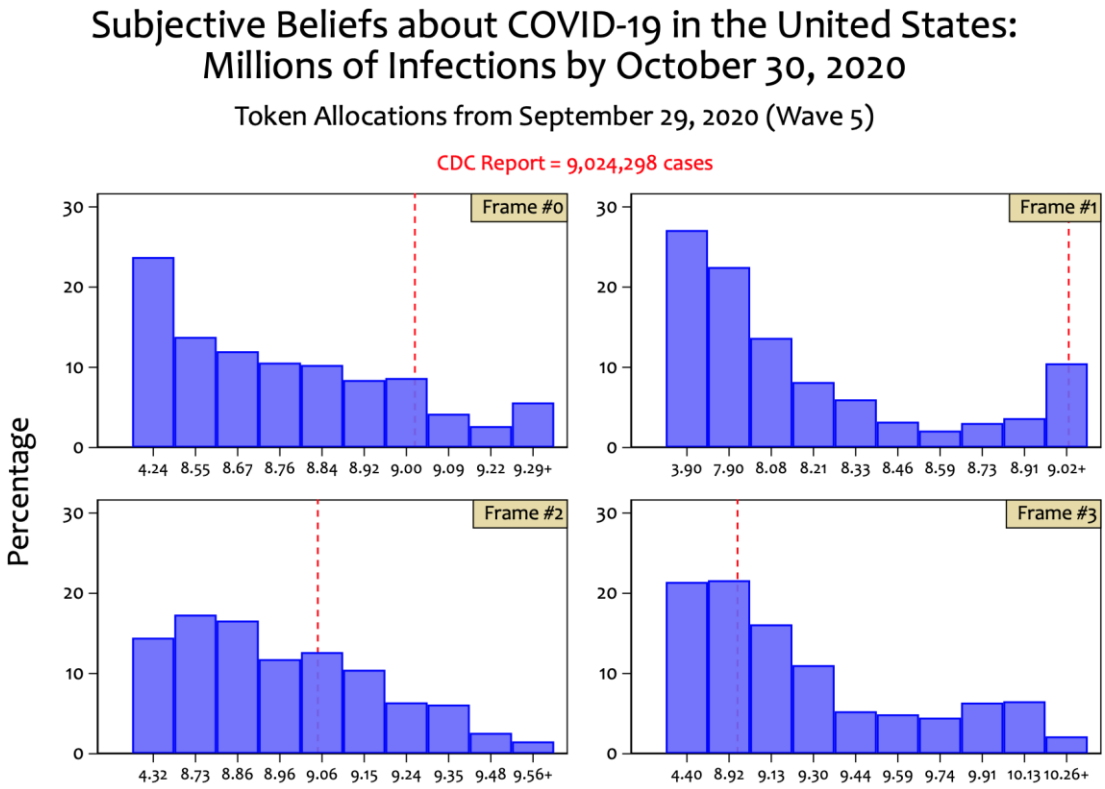

Figure B4: Beliefs about COVID-19 Infections in the U.S. by October 30, 2020

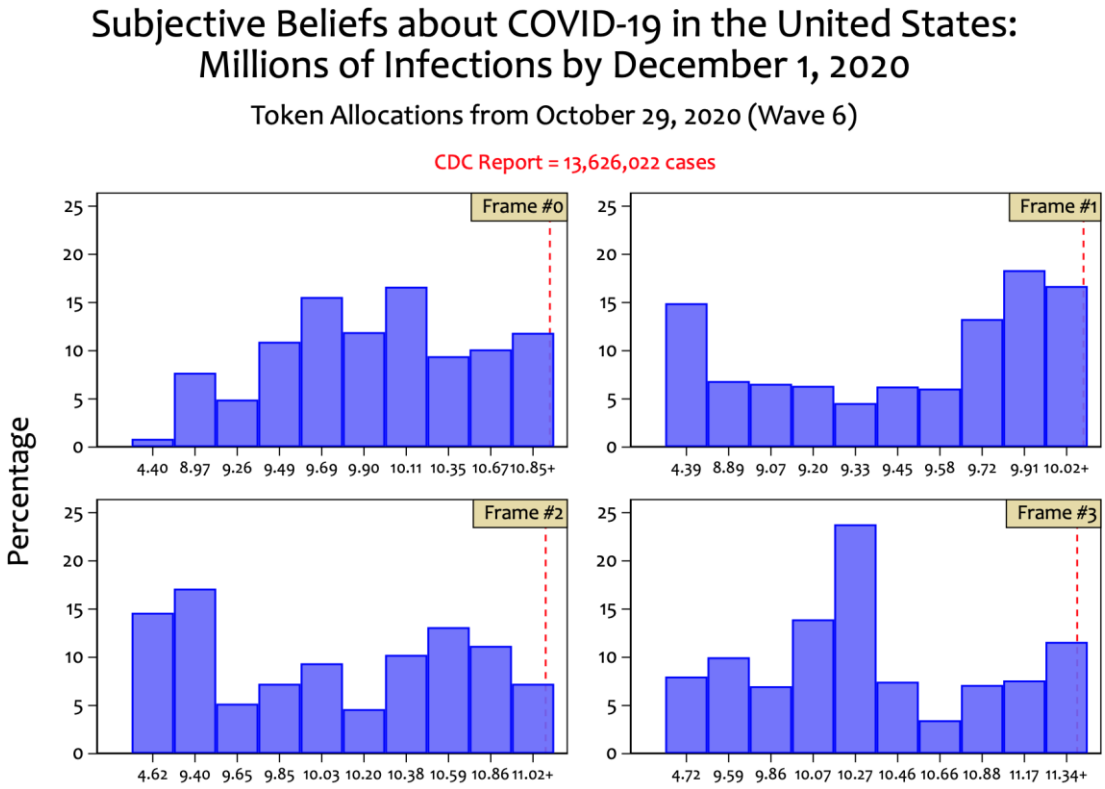

Figure B5: Beliefs about COVID-19 Infections in the U.S. by December 1, 2020

Appendix C
[Online Working Paper]

This appendix shows the distributions of subject reports about COVID-19 mortality from waves 2-6 of our study. The main text provides a general discussion of the results from these waves, and how they follow a similar pattern to that which we observed in wave 1.

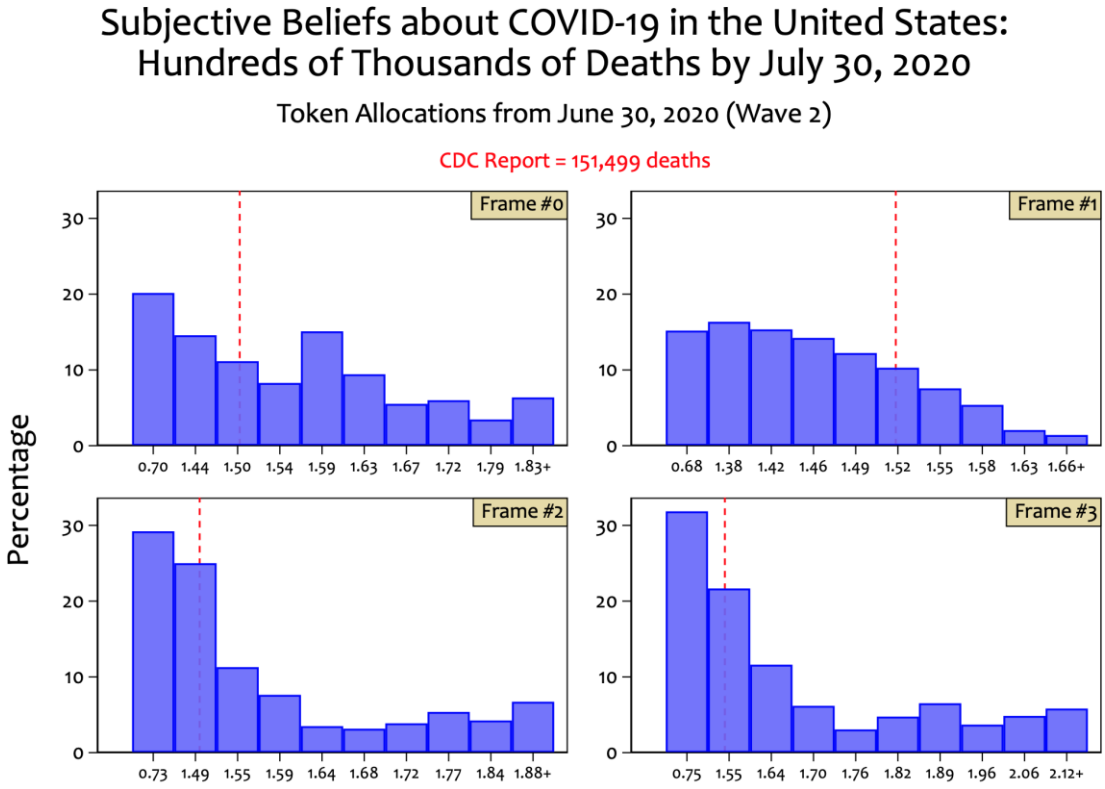

Figure C1: Beliefs about COVID-19 Deaths in the U.S. by July 30, 2020

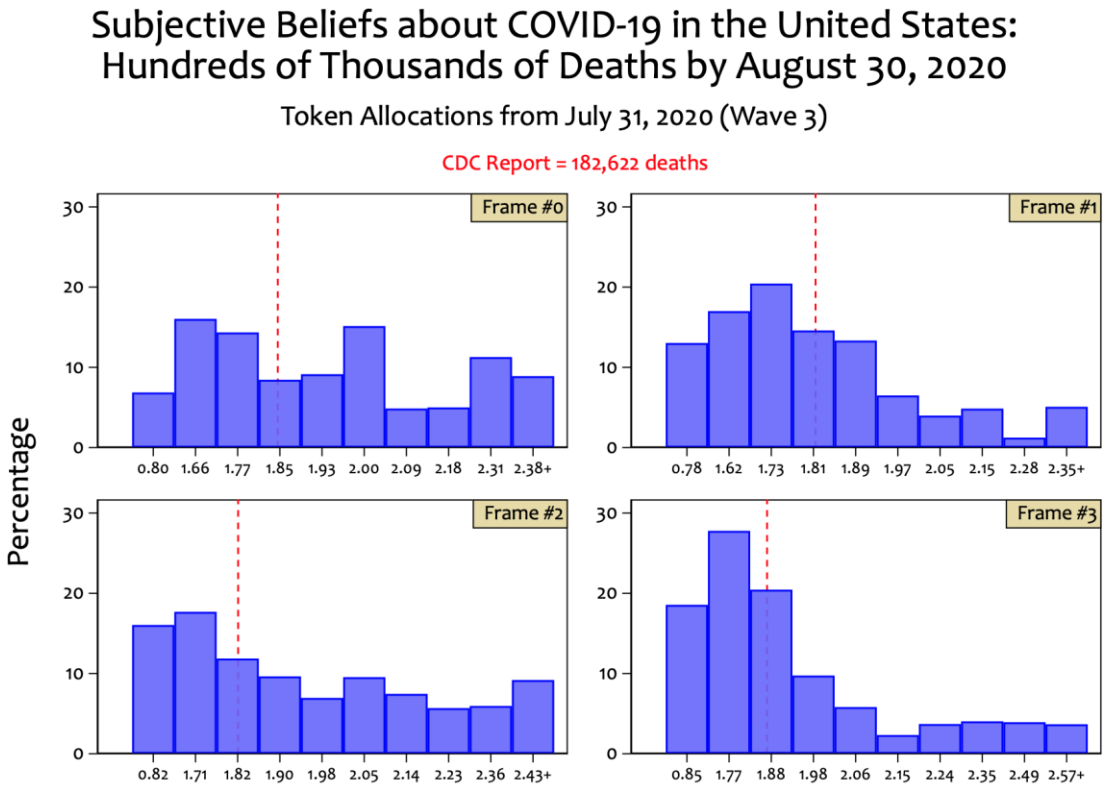

Figure C2: Beliefs about COVID-19 Deaths in the U.S. by August 30, 2020

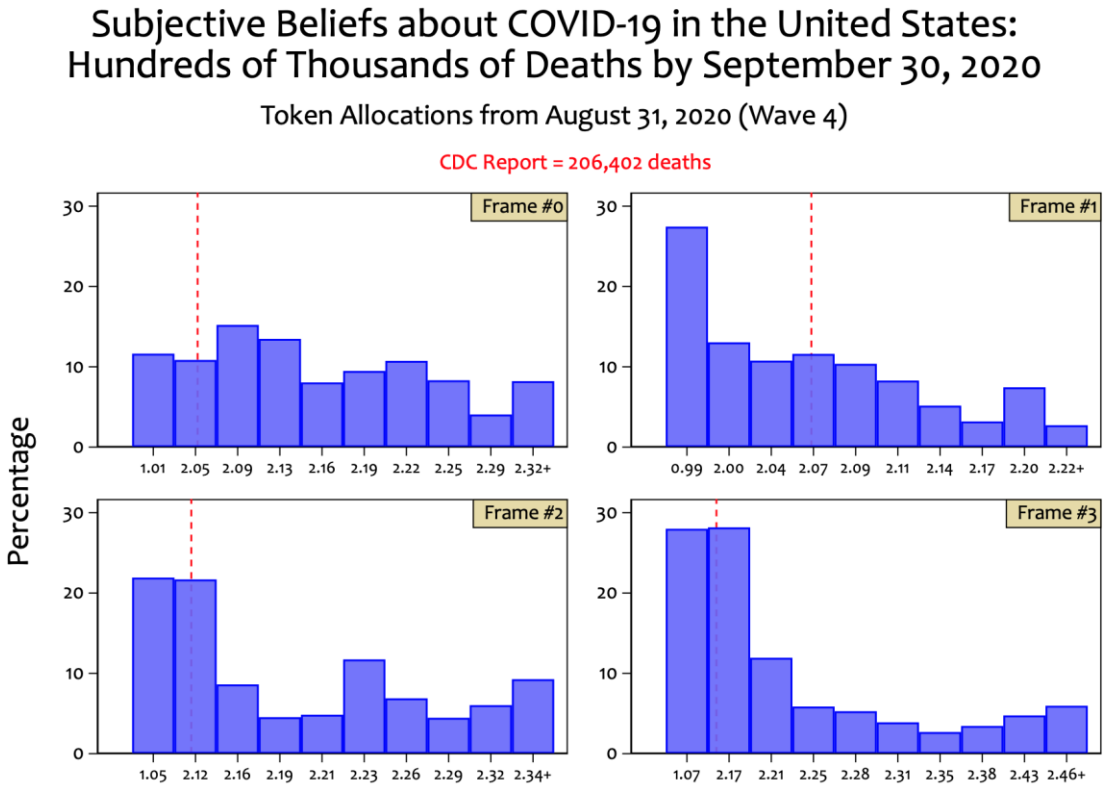

Figure C3: Beliefs about COVID-19 Deaths in the U.S. by September 30, 2020

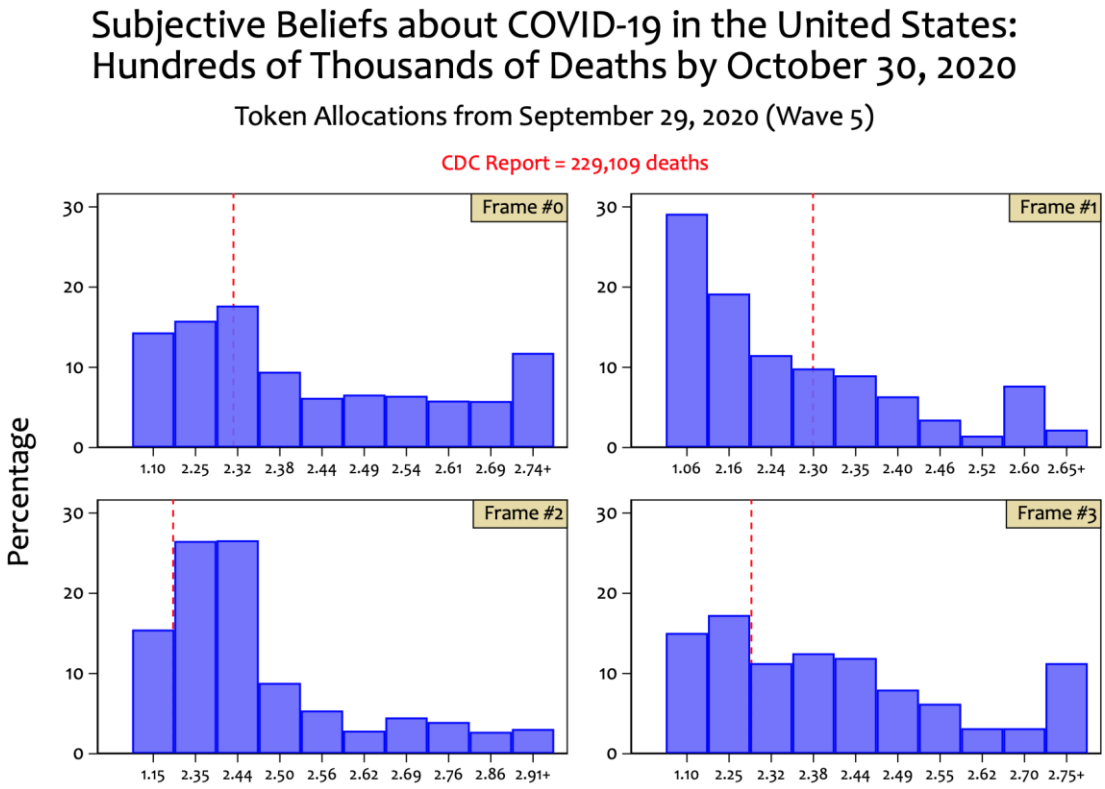

Figure C4: Beliefs about COVID-19 Deaths in the U.S. by October 30, 2020

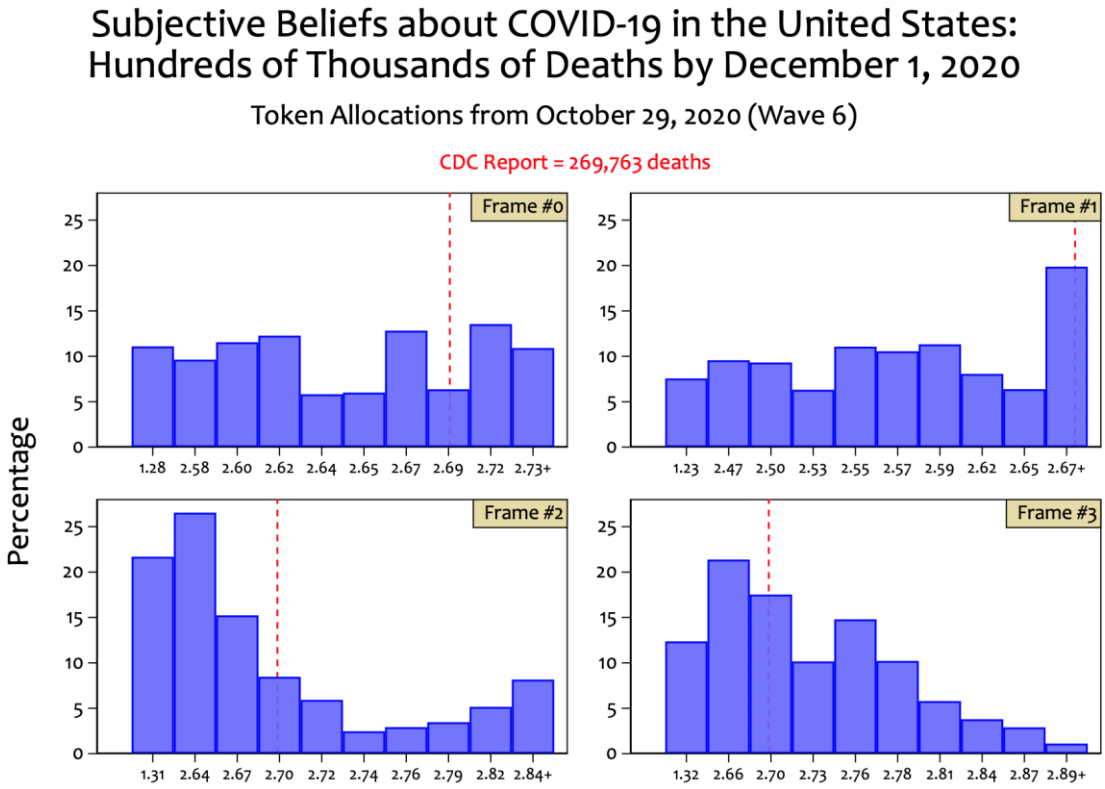

Figure C5: Beliefs about COVID-19 Deaths in the U.S. by December 1, 2020