Can Information Backfire? Facts, Beliefs, and the Health-Wealth Trade-off*

Shaun P. Hargreaves Heap†, Christel Koop†, Konstantinos Matakos†, Aslı Unan†, Nina Weber†

October 17, 2022

Abstract

Managing public health (and other) emergencies successfully requires managing citizens' beliefs and expectations, usually through information provision. But how effective is such communication? Is the provision of simple crisis-related facts sufficient to move peoples' beliefs and behavior, especially in the presence of multifaceted trade-offs and *differential threat perceptions*? We answer these questions with a survey experiment ($N = 4,618$) on a representative sample of US and UK residents. We randomize information provision on economic and health costs of the COVID-19 pandemic to assess public preferences over this trade-off –if people perceive one. In the baseline, we find that people strongly prioritize health over wealth, but the treatment effects suggest these priorities can change depending on whether public communication focuses more on the lives or income lost due to the pandemic. Information also has heterogeneous and polarizing effects that feed into pre-existing patterns. These results matter for policy design and encourage caution (e.g. in the speed of lifting restrictions) because individual differences in the relative health-wealth prioritization map onto compliance with government measures.

Keywords: COVID-19, health-wealth trade-off, information updating, communication, lockdown, policy compliance, backlash

JEL Codes: C90, D46, D78, D84, D91, H12, I12, J17

*We would like to thank the participants in seminars at the following institutions for useful comments and suggestions: King's College London, Newcastle University, Princeton University, German Armed Forces University. We also thank Riikka Savolainen, Jake Shapiro, Jonathan Portes, and Pierre-Louis Vezina for useful feedback. The data collection benefited from a grant from the King's Together Seed Fund and the research was approved by the King's Research Ethics Committee under REC ref. MRSP-19/20-18237. All authors contributed equally to the study.

†Department of Political Economy, King's College London, 30 Aldwych, London WC2B 4BG, UK. Correspondence addressed to: konstantinos.matakos@kcl.ac.uk.

"I know masks have become a partisan issue —but it's a patriotic act. Experts say wearing a mask from now until April will save more than 50,000 lives."

US President-elect Biden (`tweet @JoeBiden`, 01/17/2021)

I INTRODUCTION

When responding to major crises and emergencies policy makers and governments alike resort to behavioral interventions and extensive communication –at least at the very early stages– in order to modify citizens' behavior and mitigate the adverse consequences of such events. From financial and currency crisis (e.g. recall Draghi's famous 'whatever it takes' statement that basically stopped the run on the bonds of the EMU periphery) to natural disasters and aggregate health shocks transparent *communication of facts* has been portrayed as an effective means of modifying behavior and increasing compliance with policies that facilitate a coordinated aggregate response. Or in other words, communication in the form of factual information provision is seen as key to successfully managing such cases. But is such communication effective, especially when peoples' prior beliefs (on the severity of the situation) and threat perceptions vary substantially? Can provision of facts increase policy increase policy compliance *despite* the presence of significantly divergent views about reality?

The COVID-19 pandemic is an illustrative case in point, but need not be the only one. In the absence of pharmaceutical solutions, policy makers can only respond initially to new viruses like MERS, SARS, and COVID-19 with behavioural interventions to slow the spread of the virus.[1] COVID-19 is the most recent such virus, but it is unlikely to be the last. It also presents an interesting case of 'polarized reality' (see e.g. Alesina et al. 2020) regarding the threat it poses and the what the optimal course of action should be. As a result, such behavioural interventions pose a key challenge: they rely on citizens' compliance to be effective. Take for instance the mask mandates or lockdowns: their effectiveness rests solely on enough people following orders and complying with such restrictive measures. Thus, in the

[1] Behavioral interventions are also used extensively during the vaccine roll-out period in order to increase the take-up rates.

absence of strict enforcement, the effectiveness of similar (behavioral) non-pharmaceutical interventions relies extensively on information provision by policy makers (see e.g. Fetzer et al. 2020). Simply put, people comply only to the extend that they agree with such policies and, hence, information provision can help move such beliefs and forge agreement.

The latter has important implications for policy success during COVID-19: if the tightening or relaxation of measures does not accord with citizens' evaluations over the severity of the threat or the value of lives saved by such interventions, and whether restrictions are a cost worth paying, then success of such policies is bound to be limited. In fact, the whole debate around the presence (or not) of a health-wealth trade-off rests on this point; if a large enough segment of the population does not agree with lockdown relaxations because it sees no such trade-off or because it values lives saved at an extra-ordinary rate, easing them to restart the economy will not prove very effective (cite evidence new LSE paper). Simply put, citizens will not comply and will not modify their behavior. But such communication does not take place in political vacuum. From the early days of the pandemic response and public's attitudes to it have been heavily polarized ranging from denial of the severity of it or refusing to follow NPIs such as social distancing (or wear masks) –cite here the PNAS paper by Grossman and the the Milosh et al study).

So a key question arises: can communication of factual information reduce opinion polarization and disagreement (see Geanakoplos and Polemarchakis 1982) and, thus, move societies to the welfare-improving "cooperative" equilibrium. In other words, can the *unilateral* provision of information by governments and policy makers move societies to the Pareto superior outcome? Past studies have found that communication facilitates coordination in common interest games with positive spillovers and strategic complementarities –exactly as in the case of the COVID-19 pandemic (see e.g. Cooper et al. 1992; Ellingsen and Östling 2010). But what if information provision is de facto restricted to one party? Do such positive results hold? The answer to these matters not only for policy purposes and welfare but also bares important conceptual and theoretical implications.

This paper aspires to addresses this gap. To answer this question, we conduct a survey experiment where we employ a methodological innovation –we embed an information treatment –we provide facts about the economic and health consequences of COVID-19– into a preference elicitation mechanism, namely the relative prioritization of health over wealth.

The usual method for eliciting a relative preference over the health-wealth trade-off –if people perceive such– involves, in one way or another, seeing how people price lives saved. This is sometimes done implicitly through, for example, the wage premium enjoyed by riskier types of employment or explicitly through contingent or stated preference studies (see e.g. Bateman et al. 2002). In this paper, we adopt the latter approach by conducting a survey experiment on US and UK citizens ($N = 4{,}618$) that includes a sequence of binary options (A and B each time) where each option (A, B) combines a number of lives lost with an average income loss. The sequence is constructed in the manner of a Holt and Laury (2002) risk elicitation exercise, only in this case the trade-off between options is not between risk and return, it is between lives lost and income lost. Thus, any person who has a preference ordering over these health and wealth outcomes will switch at some point between option A to option B and where they switch provides a lower bound estimate for their valuation of a COVID-19 life. The experimental element of our survey comes from dividing our participants into three groups. One is our control group that we use to test for subjects revealing the same switch point across the two occasions they make these binary option decisions. The other two groups receive an information prompt either on expected COVID-19 deaths or on expected COVID-19 income losses before they make the binary option decision for a second time. This enables us to test whether the provision of simple facts

The information element of our survey thereby assess: i) how citizens currently value COVID-19 lives saved; ii) whether such valuations —and any individual differences behind them– influence policy effectiveness; and iii) how these valuations might evolve over time as the pandemic unfolds and both COVID-19 deaths and economic costs.[2] ... blah blah about

[2]Our information treatment serves, in fact, a dual purpose. In addition to probing the dynamic evolution of people's priors over the economic and health costs of COVID-19 as they become more informed over time,

what we do (we can take parts from existing paper). [KM: Shaun, Nina I need help here]

Our answer is not necessarily so! We find that while information provision does move some people in the expected direction it also generates a backlash effect. Further examining individual differences on who is more likely to prioritize the economy over the health and who is more likely to over/under-estimate the lives lost to COVID-19 reveal interesting partisan patterns. Not only (Trumpists etc etc) are more likely to prioritize the economy and under-estimate deaths but their reaction to information provision is not the expected one. — and here we can add something about those who reject authority and the informational backlash and the polarizing effects of information.

We also find very high –if not extreme– prioritization of health over welath at the onset of the pandemic in both the UK and the US.[3] We also find that such high prioritization of health is likely to change (fall) as information on income losses mounts. Our findings therefore suggest optimal policies and communication strategies that wish to balance marginal benefits and costs should take note of this dynamic evolution of the health-wealth trade-off as income losses accumulate during a pandemic (or any other crisis). This is the general conclusion of our paper. The more specific conclusion regarding COVID-19 is that the early spring lockdowns in the UK and USA yielded a positive net-benefit. The latter is an important conclusion in the ongoing policy debate over these lockdowns and can explain initial compliance despite weak enforcement.

But will this high prioritization of health persist as the COVID-19 pandemic unfolds and how will it be affected by communication and information provision? To answer this question, we first probe the robustness of subjects' revealed health-wealth preferences in a

it also helps us assess the effectiveness of (mostly behavioral) non-pharmaceutical policy interventions that rely extensively on information provision for the successful containment of COVID-19 and other pandemics in the future (see e.g. Fetzer et al. 2020).

[3]Taken at face value, our elicited valuations are a factor of 10 times greater than the base value suggested by the OECD (2012) for usual policy evaluation and it is a factor of 3 times greater than the Environmental Protection Agency figure that was used in a recent utilitarian social welfare function approach to the evaluation of lockdowns (Hall et al. 2020). In a follow-up study, where we use the conventional approach to deriving such valuations we, again, find very similar figures. While our main analysis does not hinge upon those elicited valuations, their size and their consistency across methods and over time is telling.

variety of ways. For instance, we ask our subjects to make the set of decisions on the binary options a second time and we check whether they plausibly reveal the same switch point on both occasions. We find that most of our subjects do. We then consider whether their strong preferences for health over wealth have behavioral consequences that policy makers need to take into account. The point is that it is often argued that people only comply with policies in so far as they agree with them (Tyler 2006). If this was the case for COVID-19 policies, then policy makers need to take account of *actual* citizens' preferences because these, in turn, will affect their likelihood of compliance with policies. In other words, responding to citizens' preferences matters not only for electoral popularity of any policy; it could also matter for the likely effectiveness of that policy in achieving its objectives. We find evidence of such a behavioral correlate with the valuations revealed in our survey: individual (self-reported) likelihood of compliance with government guidelines in both countries is predicted by individual valuations of COVID-19 lives saved. Our findings of such high prioritization of health together with its connection with compliance plausibly helps explain why the initial lockdowns were so successful in reducing virus transmission (Flaxman et al. 2020; Hsiang et al. 2020).

In practice, it is *as if* people's high valuations of health outcomes internalized the significant externalities that pandemics like COVID-19 entail (e.g. staying home and healthy protects others as well). Such estimates could be an indication of the public's strongly pro-social behavior during a pandemic. Alternatively, it can be that the shape of peoples' utility function –if the have one– varies with own-health (Finkelstein et al. 2013), thus generating significant spillovers from health to wealth. This finding is important in supporting our general conclusion. It suggests that it is only when policies are compatible with citizens' preferences during a pandemic we can expect to enjoy *voluntary* compliance. This is another reason why the informational element of our survey is important: policy makers need to know whether peoples' preference for health over wealth is likely to persist (or will fall) as the experience of income losses accumulates.

In sum, our information treatment helps us assess the effectiveness of (mostly behavioral) non-pharmaceutical policy interventions that rely extensively on information provision for the successful containment of COVID-19 and other pandemics in the future. Our answer is that the glass is half-full: communication is not the silver bullet of such interventions. In fact, in some instances excessive communication can have the opposite effects. Our study thus points to the limitations of mitigation strategies that rely exclusively on soft communication and information provision as they are unlikely to modify behavior significantly.

Thus, our main contribution is to alert policy makers to the perverse (and sometimes polarizing) effects that –even factual– information provision and expert communication might have in the context of highly polarized and politicized health crisis (Adida et al. 2020). In that respect President Biden's quote is very relevant: a careful communication strategy is needed if we are to break the polarization wall. Our final contribution comes from the evidence regarding the information treatment because it also connects with the emerging literature on polarization over facts and the politicization of health and other crises (see e.g. Alesina et al. 2020; Barrera et al. 2020; Aksoy et al. 2020; Grossman et al. 2020). In this further respect, our evidence points to another aspect in which the COVID-19 crisis might be unique but its consequences could be lasting.

Thus, our main contribution is to alert policy makers to the current high VSLs during COVID-19, that these valuations affect policy compliance and how they will likely evolve over time. In so far as these results can be taken to apply to all pandemics, they suggest a general conclusion for pandemic policy making: start with a very high VSL, compared with the normal, off-the-shelf valuations, and allow for the VSL to fall as income losses mount. In the particular case of COVID-19, and unlike the off-the-shelf VSLs, the high pandemic VSL implies that the early spring lockdowns in the UK and the US likely produced a large net-benefit. Our final contribution comes from the evidence regarding the information treatment because it also connects with the emerging literature on polarization over facts and the politicization of health and other crises (see e.g. Alesina et al. 2020; Barrera et al. 2020; **?**;

Aksoy et al. 2020; Grossman et al. 2020). In this further respect, our evidence points to another aspect in which the COVID-19 crisis is unique whereas its consequences may be lasting.

The rest of the paper is organized as follows: In the next section, we present our research design and the empirical estimation strategy in more detail. Section 3 presents our key findings and section 4 concludes.

II RESEARCH DESIGN AND METHODS

DATA AND SAMPLING

To conduct the online experiment, we teamed up with Prolific Academic, a web-based panel with about 35,500 participants in the United States (US) and 44,600 participants in the United Kingdom (UK) as of May 2020. Our quota-based sample was recruited between Friday 17 and Tuesday 21 April 2020: at the end of the week when both the UK and the US were predicted to hit peak deaths (IHME 2020b,c) To generate samples for the US and the UK, we used the US Current Population Survey (US Census Bureau 2018), the 2011 UK Census (ONS 2011), and the Scotland's Census 2011 (NRS 2011) in order to achieve a representative sample. We excluded Northern Ireland from the survey. We created a total of 170 subgroups weighted based on age, gender, region and work status. Table 1 and 2 in the SM appendix B, section 1 are the stratification tables for the UK and US, respectively. We targeted a total sample size of 2,500 respondents for each country, of which we were able to reach 95.36% in the UK and 89.08% in the US. In sum, 2,385 and 2,233 respondents participated in the UK and the US, respectively. Table 3 in the SM appendix B, section 1 reports the subgroups that we could not fill our quotas completely on Prolific and thus weighted accordingly in our analysis to ensure representativeness. All of these subgroups were within the oldest two age groups of which Prolific has disproportionately fewer active participants.

The average completion time was 33.6 minutes and respondents earned on average £3.08 for their participation. The survey was pre-registered on https://osf.io/qtes9/files/ and approved by the King's Research Ethics Committee under REC ref. MRSP-19/20-18237. Further details on sampling, as well as the full survey instrument that we used are available online. The data and code used for the analysis will be made available online at Harvard's Dataverse for replication purposes upon acceptance for publication.

EXPERIMENTAL DESIGN

Our survey experiment consisted of three parts. In the first part, in the spirit of Holt and Laury (2002), respondents were asked to make a sequence of eight binary decisions between pairs of health (expected number of COVID-19-related deaths per 1 million of the population) and wealth (expected average household income loss due to measures to prevent transmission of COVID-19) outcomes over the a period of 3 months. Figure 1 shows the actual sequence of eight decisions between these pairs given to UK and US respondents. Respondents read a short text on how restrictions on personal movements help contain the spread of coronavirus and save lives but with a cost of disrupting and lowering economic activity. They were then presented with those eight decisions sequentially.[4] In each of the eight decisions, they clicked on the option that they think has the best combination.

As stated briefly above, in the second part, the experimental element of the survey comes from our test of the stability of these revealed preferences for health versus wealth when provided with expert communication in the form of information prompts. We asked respondents to make those eight decisions a second time divided into three groups. After the first round of these decisions and prior to treatment, they engaged in an unrelated task. They were further asked to answer questions regarding their likely estimates of COVID-19 deaths and income loss given the current lockdown. Respondents were then divided randomly into three sub-groups and, before the second round of decisions, one sub-group

[4]We slightly deviated from the pre-registered figures in our decisions menu due to further elasticity calculations following our pilot.

Figure 1: Decisions for preference elicitation

	United Kingdom		United States	
	Option A	Option B	Option A	Option B
Decision 1	445 lives lost per million, £2,700 average disposable income loss	460 lives lost per million, £2,750 average disposable income loss	320 lives lost per million, $4,000 average disposable income loss	335 lives lost per million, $4,150 average disposable income loss
Decision 2	412 lives lost per million, £2,500 average disposable income loss	431 lives lost per million, £2,420 average disposable income loss	310 lives lost per million, $3,850 average disposable income loss	325 lives lost per million, $3,740 average disposable income loss
Decision 3	383 lives lost per million, £2,300 average disposable income loss	393 lives lost per million, £2,200 average disposable income loss	247 lives lost per million, $3,670 average disposable income loss	256 lives lost per million, $3,500 average disposable income loss
Decision 4	360 lives lost per million, £2,150 average disposable income loss	367 lives lost per million, £2,020 average disposable income loss	213 lives lost per million, $3,500 average disposable income loss	219 lives lost per million, $3,300 average disposable income loss
Decision 5	300 lives lost per million, £2,000 average disposable income loss	305 lives lost per million, £1,850 average disposable income loss	200 lives lost per million, $3,300 average disposable income loss	204 lives lost per million, $3,100 average disposable income loss
Decision 6	240 lives lost per million, £1,900 average disposable income loss	243 lives lost per million, £1,750 average disposable income loss	188 lives lost per million, $3,120 average disposable income loss	192 lives lost per million, $2,820 average disposable income loss
Decision 7	230 lives lost per million, £1,800 average disposable income loss	232 lives lost per million, £1,640 average disposable income loss	177 lives lost per million, $2,350 average disposable income loss	180 lives lost per million, $2,000 average disposable income loss
Decision 8	210 lives lost per million, £1,550 average disposable income loss	210 lives lost per million, £1,450 average disposable income loss	165 lives lost per million, $1,950 average disposable income loss	165 lives lost per million, $1,800 average disposable income loss

received information about predicted COVID-19 deaths and another received information on predicted income losses due to COVID-19 mitigation (lockdown) measures. Both predictions came from 'experts'. The control group heard a short piece of instrumental music instead of receiving information. If individuals change how they prioritize health over wealth in one or both information treatments (as compared with the control), this suggests priorities will change in predictable ways as the experience of death and economic loss unfolds in the coming weeks and months. In so far as there are no treatment effects and individuals do not change their revealed priorities significantly between the first and second round, the results point to coherence and stability in priorities in the face of changing information. The third part contained some standard demographic questions.

Treatment information

Our treatment consists of two types of information prompts that are shown to the survey respondents. The first prompt provides information about estimated lives that will be lost (in the US and the UK) by August 2020 according to the IMHE (IHME 2020a). The second prompt provides respondents with information on expected income (GDP) loses based on estimates presented by the IMF (IMF 2020). We present the exact wording of the two information treatments in the survey instrument available online.

Finally, we asked a series of demographic and attitudinal questions. This enabled us to assess whether individual compliance with current lockdown measures is indeed predicted by individual differences in the valuation of health over wealth. We also tested for which individual objective characteristics (e.g., age and income) and subjective ones (e.g., risk tolerance and fear, which we measured as an index capturing respondents' perception of the threat of COVID-19) help predict these individual differences in the valuation of health versus wealth.

EMPIRICAL STRATEGY

To estimate our main treatment effects, we analysed the data using two statistical forms – an ordinary least squares (OLS) regression and a logistic regression – in order to identify the causal effects of our treatment and how they interacted with respondent i's estimate of deaths and income lost. Treatment assignment to one of the two groups (plus the control group) was fully randomized. Such analysis allows us to understand which of the variables has a significant impact on health-wealth prioritization.

We estimated the following two basic empirical models, whereby HL_i i is the *change* (between the two rounds) in respondent i's preference over health and wealth, δ_1 is the treatment effect, δ_2 the effect of respondent i's estimate of deaths interacted with the health treatment (T1), δ_3 the effect of respondent i's estimate of the income loss interacted with the wealth treatment (T2) and ϵ_i the error term. In all our main specifications we used

population weights (as specified above) in order to able to make inference for the general US and UK populations. We also clustered our standard errors at the regional level (US States and UK NUTS-2 areas). Formally, we estimate the following equations:

$$HL_i = \beta_0 + \delta_1 treatment + \epsilon_i \tag{1}$$

$$HL_i = \beta_0 + \delta_1 treatment + \delta_2(t1 \times deathest_i) + \delta_3(t2 \times incomeest_i) + \epsilon_i \tag{2}$$

Given that we are interested in within subject changes between pre- and post-treatment preferences we do not control for demographics in our main estimation. Section 3a (in the SM appendix B[5]) includes the main treatment effects with demographic controls. Parameters δ_1, δ_2, and δ_3 capture the causal estimates of our treatment effects. Random assignment to treatment ensures the causal interpretation of OLS estimates. The results of our main analysis are reported in Figures 4 and 3 below (and in more detail in Tables 1-4 of SM A, section 1b). We also report a series of robustness checks in SM A, section 2 and in SM B, section 3, including alternative coding of the main variables and ATE estimates (between subject design).

Outcomes

Our main outcome variable is derived from the Holt and Laury (2002) elicitation process and captures the relative subjects' prioritization of health over wealth (for more details see SM Appendix A, section X.x). Briefly, if a respondent values both life and income and has a preference ordering over their various combinations, they should choose option A in Decision 1 and option B in Decision 8. This is because, in Decision 1, A dominates B in both the health and wealth outcomes, whereas in Decision 8, B weakly dominates A as both have the same death outcome, but B is better on income loss. In the intermediate Decisions

[5]Appendix B is available online.

2-7, option A has the better health outcome and option B has the better wealth outcome. As subjects move through Decisions 2-7, the health advantage of A over option B becomes progressively smaller in terms of deaths avoided per unit of income lost. In this way, a person with a preference ordering will switch from option A to B as they progress through Decisions 1-8. Where they switch indicates how strongly they prioritize health over wealth: the later the switch, the stronger the preference for health over wealth (Holt and Laury 2002; Bateman et al. 2002). Subjects that exhibit canonical economic preferences over health and wealth should switch *at most once* in this sequenced of decisions. Subjects who never switch –we call them *All A's'*– exhibit an extreme preference for health[6] We present more details

We measure our DV in five different ways. We have two categories of outcomes: a) binary ones (Switching Up; Switching Down; Down from maximum value of life (VoL); Up to maximum VoL) and b) a continuous one (VoL). We detail each one of them (and how we computed them) in SM A, section 1a. In order to collect the outcome information, we simply analyzed the responses that subjects gave in the two parts of the survey that contained the eight binary decisions. The exact phrasing of those binary decisions and the questions used to collect the outcome data can be found in the survey instrument available online.

But how does our elicitation method compare with the one used to estimate the value of statistical life (VSL)? In the standard stated preference elicitation of the value of statistical life (VSL), an individual is usually asked to place a valuation on an intervention like a safety procedure, an air quality regulation or medical procedure in so far as it affects that individual's own life chances: E.g. how much would you pay for an intervention that reduced your chances of dying from air pollution from 3 in 100,000 to 2 in 100,000? The economy-wide benefits are then judged by summing these marginal assessments of personal value and comparing them with the costs of this intervention for the economy. In this way the individual is asked to assess some intervention solely in terms of how it improves his or her chances of living. In practice, though, a public policy intervention may be valued by an

[6]Need to say something about all A's here no? And then put more details in the appendix?

individual not just for its effects on that individual's own life chances, it may also be valued because it affects everyone's life chances. This is especially the case for a communicable disease like COVID-19 where a person may value avoiding COVID-19 not only because their own life chances improve but also because they are less likely to transmit it to others and the person has other regarding preferences. This is quite unlike, for example, an air pollution intervention where there are no such spillover effects because pollution related respiratory diseases are not transmitted from one person to another. As a result, we ask a different social planner-like question regarding preferences over aggregate combinations of health and wealth outcomes. These can be answered in the same way as the more usual VSL questions if a person is only concerned with the effect on their own life chances exclusively by interpolating the individual marginal effects on life chances from the aggregate data. But, in addition, with our questions, those who believe that there are externalities involved or have other regarding preferences (or both), can take these effects into account when answering (i.e. our elicitation mechanism internalizes possible externalities and other regarding preferences).[7]

III RESULTS

PREFERENCES AT THE BASELINE AND THREAT PERCEPTIONS

We start by presenting peoples' health-wealth preferences at the baseline (before the provision of information) as well as their situational awareness and threat perceptions regarding COVID-19 –and any possible heterogeneity patters this might exhibit. First, in figure X

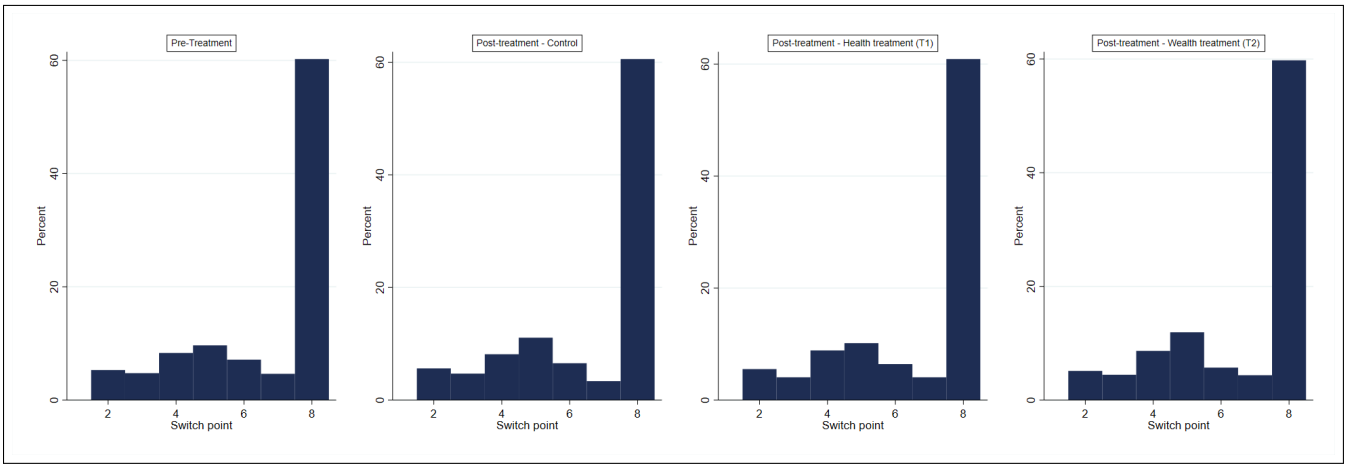
[7]In a subsequent study, we have elicited the VSL using the more usual individual focused question regarding "how much you would pay for an intervention that changed your chances of death form COVID-19 by 1 in x". We find that the conventional question yields VSL estimates that are extremely high (e.g. close £12m in the UK and $17m in the US, many orders of magnitude higher than off-the-shelf estimates currently used). Although this is not as high as the figure we report later for our social-planner form of question (based on the Holt and Laury (2002) elicitation), we also find that when the conventional question is modified to refer to an intervention that works to produce the same change in the chances of death by reducing transmission rates (i.e. it works through reducing spillover effects and so will trigger other regarding preferences if subjects have them), the VSL jumps by almost 50% and becomes very similar to the figure we report here. Thus, our subsequent conclusions and analysis are *not sensitive* to the specific VSL elicitation mechanism used.

[WE NEED TO ADD HERE THE DIFFERENCES IN PERCEPTIONS HERE]

These differences in threat perceptions and the severity of the situation also map directly into differential (stated) intentions of compliance with governments' regulations, guidelines and restrictions.

Figure 2 reports on the distribution of switch points in the UK and the US for those respondents switching only once in both the first and the second round (disaggregated by treatment and control). The majority in both countries switch at Decision 8 in both rounds, indicating a very high relative prioritization of health over wealth.

Figure 2: Percentage of those who switch once, by decision switch point



INFORMATION EFFECTS ON HEALTH-WEALTH PREFERENCES

Given that the current valuations of life saved are very high and that they likely impact on policy compliance, it is important to know how they might evolve as information on COVID-19 deaths and income losses accumulates. This matters for policy for an additional reason: as Hall et al. (2020) show people's location on the trade-off between COVID-19 deaths and consumption critically depends, among other things, on the relative prioritization of health (i.e. lives saved) over economic outcomes. Our information treatments are designed with this in mind. We compare the likelihood of person changing their switch point between the two

rounds in our information treatments as compared with the control. Such changes in switch point are, importantly, within-subject and so we can identify the information treatment effects by treatment dummies in a regression on what is a between-subjects comparison in the treatment and the control groups in this respect (see also SM A, section 1b-c).

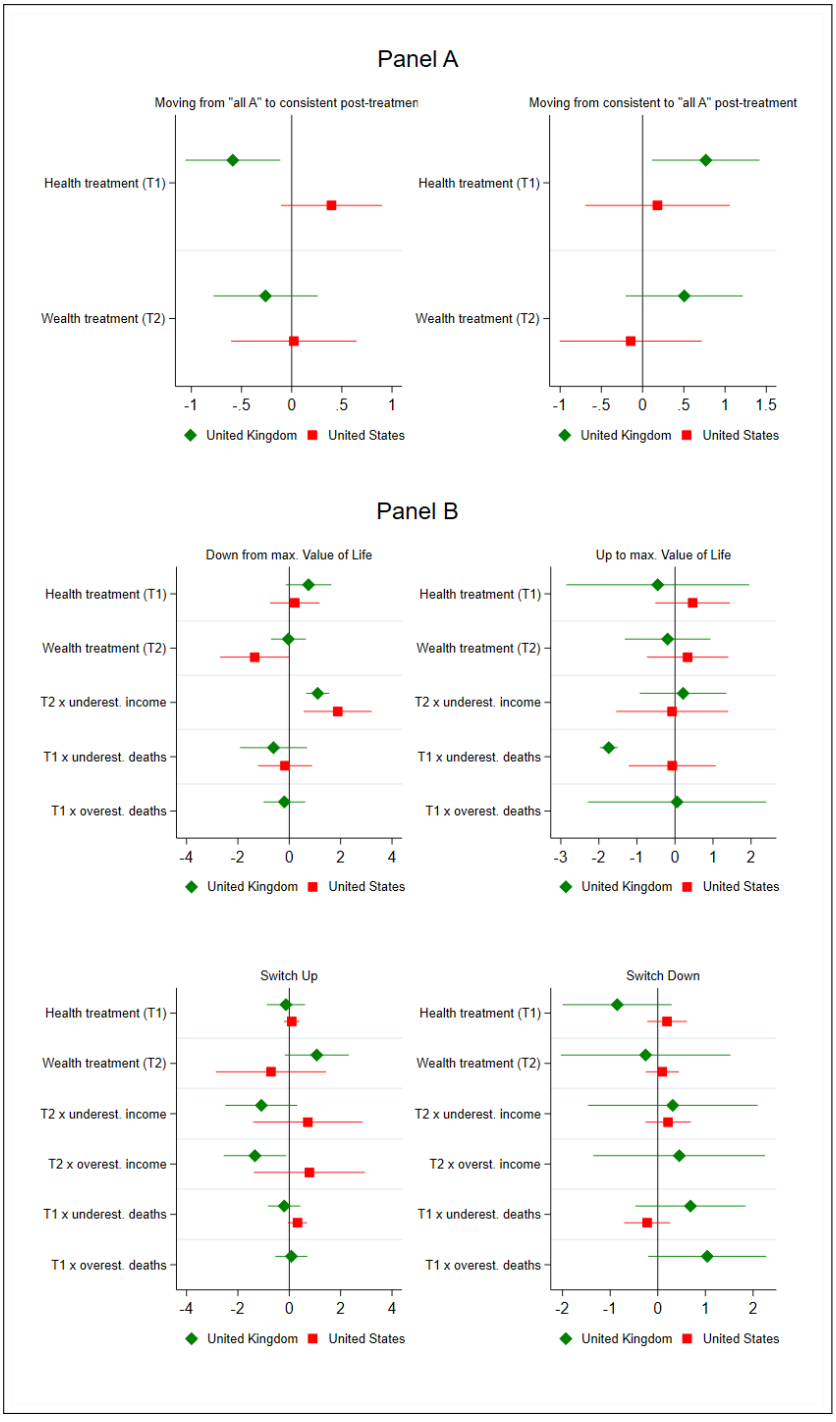
We have three treatment effects to report. First, an unconditional COVID-19 death information treatment effect in the UK: there is a significant increase in the number of subjects who switch from having a single switch point on the first occasion to always choosing option A (see Panel A in Figure 3). That is, they shift from having a preference ordering over health and wealth to a strict preference for health, whatever its wealth cost. This effect is not driven by any of the subgroups, which suggests that simply being sensitized to the health costs of the crisis increases the likelihood of strictly prioritizing health over wealth.

Second, Panel B in Figure 3 shows a significant conditional treatment effect in both the US and the UK. Those in the income loss information treatment group who learned that they underestimated the predicted income loss are significantly more likely to move down from the Decision 8 switch point. 60% of the population switch at Decision 8 and 28% underestimated the income loss. These effects are very robust: they are supported by the between subject analysis when comparing across the two treatment groups and the control (we report ATEs in SM A, section 2d). The last two plots in Panel B suggest that this treatment effect, however, does not occur throughout the range of possible switch points (i.e., for the other 40% of this group).

In the UK, this second treatment effect increases with age (i.e. vulnerability to Covid-19). The older respondents who underestimate income losses are, the more likely they are to move down from Decision 8 after receiving income loss information. In the US, this is not the case. Instead, the more fearful respondents are of COVID-19, the more likely they are to reveal this treatment effect (SM B, section 2a).

The third treatment effect can be seen in the second plot of Panel B in figure 3. It is conditional and qualifies the first effect for those who underestimate the predicted COVID-

Figure 3: Treatment effects



Notes: Figures are based on logistic regressions. 95% confidence intervals are shown. The outcome variable of the coefficient plot on the left side in Panel A is a binary variables equal to 1 if the respondent moved from choosing option A in all decisions pre-treatment to a consistent preference ordering post-treatment. The binary outcome variable on the right side of Panel A indicates the reverse movement.'Down from max. VoL' is a binary variable equal to 1 if the respondent moved from switching from option A to B at decision 8 to an earlier switchpoint post-treatment. 'Up to max. VoL' captures the reverse movement. 'Switch Up' is a binary variable equal to 1 if the respondent switched from option A to B at an earlier decision post-treatment and 'Switch Down' captures a later switchpoint post-treatment.

19 deaths in the death information treatment of the UK. Those who underestimate go in the opposite direction to the general treatment effect: they are less likely to move up to the maximum relative valuation of health over wealth than in the control group. Examining the possible reasons for this, we find that it is associated with individual respect for authority (SM B, section 2b). Those who have less respect for authority are, it seems, more likely to react perversely to the death information by becoming less likely to value health relative to wealth so highly. It is 'as if' they respond to information about the death toll being worse than anticipated by 'refusing' to update and decide instead that lives matter less: an informational backlash. This treatment effect – together with the first one – has the important implication that unexpected deaths will polarize the UK public: death information generally increases the valuation of health, but the reverse is true for those who underestimate the deaths. The polarizing effects of information on this particular instance also connect to the existing body of evidence on the public's divergent perception of and reaction to common facts (Alesina et al. 2020) and the politicization of epidemics and other public health crises (Adida et al. 2020).

WHO VALUES HEALTH OVER WEALTH

Table 1 presents the likelihood of having the highest possible valuation of life (a switch point at Decision 8) or a strict preference for health (choosing Option A throughout) given respondents' age, economic vulnerability and various subjective characteristics. These are the outcome variables for which we observe treatment effects and so understanding who expresses these valuations is important. We measure generalized trust, risk, patience and altruism preferences (Falk et al. 2016, 2018) of respondents as well as their fear of COVID-19. Our index of fear is based on polychoric principal component analysis. Details of the individual components of the index can be found in the supplementary material (SM A, section 1f). Fear predicts a high valuation of health over wealth in both countries, while trust and risk aversion additionally explain such a valuation in the UK. These are plausible

subjective correlates for the valuation of a COVID-19 life saved, adding to our confidence in the high reported estimates for the value of life saved.

Table 1: Characteristics of respondents with high valuation of life

	Highest Value of Life		Strict Prioritization of Life	
	UK	US	UK	US
Age	0.110	-0.084	0.336**	0.060
	(0.071)	(0.084)	(0.135)	(0.082)
Economic vulnerability	-0.024	-0.090**	-0.176***	0.003
	(0.055)	(0.043)	(0.050)	(0.069)
Trust	0.421***	0.263	-0.375	-0.149
	(0.140)	(0.169)	(0.293)	(0.167)
Risk preference	-0.057**	0.016	-0.018	-0.088**
	(0.029)	(0.032)	(0.050)	(0.038)
Patience	0.046	-0.006	-0.014	-0.050
	(0.031)	(0.041)	(0.055)	(0.046)
Altruism	0.000	0.000	0.000	0.000
	(0.000)	(0.000)	(0.001)	(0.001)
Fear	0.147***	0.137***	0.063	0.015
	(0.048)	(0.045)	(0.057)	(0.048)
Constant	0.915	0.314	-3.933***	-0.469
	(0.726)	(0.611)	(0.932)	(0.929)
Controls	✓	✓	✓	✓
Regional clustering	✓	✓	✓	✓
Observations	1,079	1,257	1,092	1,257
Pseudo R-squared	0.090	0.069	0.107	0.069

Notes: Estimates come from a logistic regression. Economic vulnerability is a self-reported measure based on the respondents' ability to cover household expenses in the following month. Trust is measured as a binary variable with a value of 1 indicating a general trust in other people. Risk preferences are self-reported on a scale from 0 to 10 with 10 being the most risk-seeking option. Patience is equally measured from 0 to 10 with 10 being the most patient option. Altruism is measured from 0 to 1000 based on hypothetical charity contributions. Regional clustering is done based on either the 12 regions of the UK, as defined by the ONS or the 50 states of the United States, as defined by the Census Bureau. Clustered standard errors are presented in parentheses. *** p<0.01, ** p<0.05 , * p<0.1.

In the UK, we observe a polarization between age and economic vulnerability on the likelihood of having a strict prioritization of health over wealth. Age is positively correlated with this valuation; economic vulnerability negatively. Our first treatment effect therefore suggests that while information about the income loss due to the crisis does not affect this polarization between the old and the economically vulnerable, information about deaths may further amplify it.

We do not observe this polarization between age and economic vulnerability in the US. As we also have no treatment effect on the strict prioritization of life in the US, this suggests that the conflict between the old and economically vulnerable is more prominent in the UK than the US.

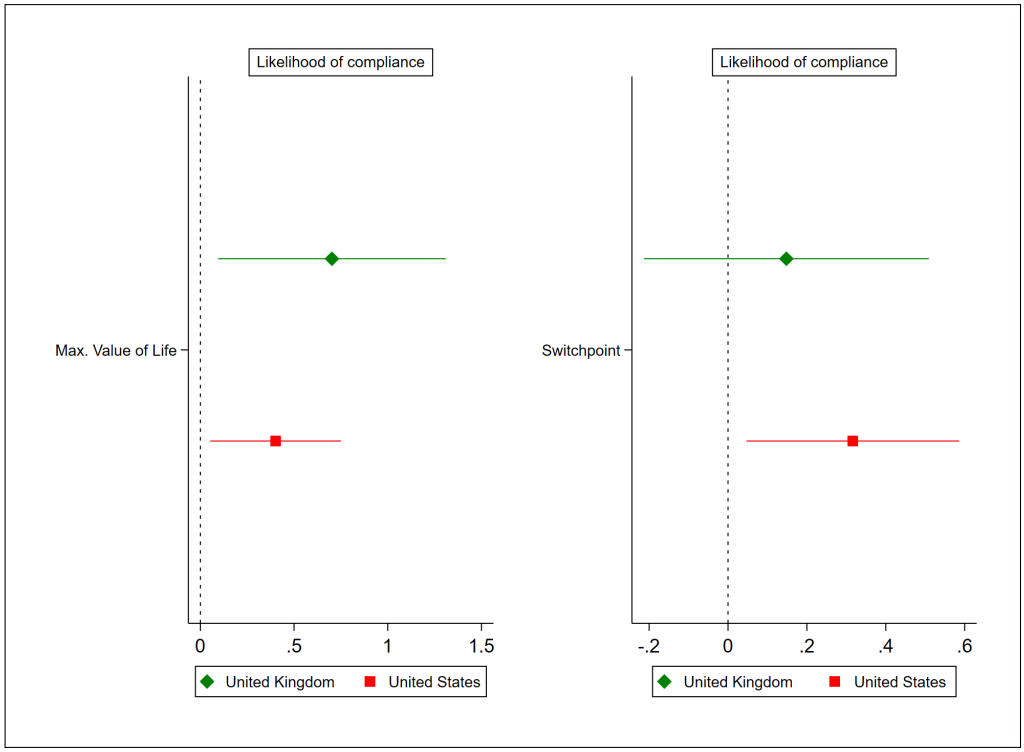
IMPLICATIONS FOR COMPLIANCE WITH GOVERNMENT GUIDELINES

In Figure 4, we present the regression results testing whether individual differences in the valuation of health versus wealth are likely to influence policy efficacy because they help predict differences in stated likelihood of compliance with the current government guidelines in both countries. They do. Those who choose the maximum valuation of health over wealth are twice as likely to state that they strictly comply with lockdown guidelines in the UK and 1.5 times as likely in the US compared to everyone else (see SM A, section 1d). A maximum valuation of health over wealth is the best predictor of compliance in both countries, together with being female and having a high level of trust in other people (see SM A, section 1d). Thus, policy makers must pay attention to the public's valuation of health over wealth, not only for electoral reasons but also for reasons of policy efficacy.

IV DISCUSSION

In the face of health-wealth trade-offs that are always involved in any decision regarding the introduction of pandemic behavioral interventions, policy makers need to take account of the value that citizens place on lives saved. This is not just for electoral considerations, it is also because we provide evidence that individual valuations map on to individual compliance with such behavioral policies. Thus these valuations also matter when judging the likely efficacy of any non-pharmaceutical behavioral intervention: people do not comply with policies that do not accord with their preferences. Of course, if these valuations could reasonably be expected to be close to those that are typically used in policy analysis, then there would be

Figure 4: Health over wealth preference and lockdown compliance



Notes: Figures are based on ordered logistic regressions. 95% confidence intervals are shown. The outcome variable of both coefficient plots is a categorical variable capturing respondents' stated likelihood of compliance with government guidelines on a scale from 0 to 4. Max. value of life is a binary variable equal to 1 if the respondent indicated the maximum value of life by switching from option A to B at decision 8 and 0 otherwise. Switchpoint is a categorical variable indicating the decision at which the respondent switched from option A to B.

no pressing need to check on them now. But it is known that these valuations are sensitive to context, and a global pandemic, such as COVID-19, is a life-defining event. When we check with this survey experiment what are the values attached to a COVID-19 life saved now, we find that they are very much larger than the range of figures used in more normal or routine policy analysis.

We have probed these high valuations in a variety of ways. They appear robust both in the sense that most of our subjects in the control appear to have the same high valuation across the two rounds and in the sense that the individual differences in valuation are predicted by plausible objective and subjective individual characteristics.

Based on such high valuations, the original lockdown measures, that may have saved several hundred thousand lives in both countries at the loss of perhaps as much as 10%

of GDP (Bank of England 2020), were consistent with the public's preferences for health over wealth. This, in turn, fits with the high trust and approval ratings that governments enjoyed when the lockdown measures were introduced (Bol et al. 2020). However, if these high valuations remain (OECD 2012), policy makers have a daunting task in calibrating the right mix of any such behavioral intervention in the future. If looser restrictions are accompanied by relatively modest increases in deaths, then they will not be very popular, given these high valuations, with the majority of the population in both countries even if it were to restore income losses. Furthermore, since we find that such unpopularity will likely reduce compliance and so undermine the effectiveness of the relaxation in restoring income losses, the initial unpopularity could quickly spiral. The possibility of such a vicious circle developing points to caution in relaxing such interventions and the need for quick re-introducing restrictions if significant spikes in COVID-19 deaths occur after any relaxation.

This conclusion is reinforced by the experimental element of our survey. The information treatment effects that we find suggest that people's relative valuation of health over wealth is dynamic and will change in predictable ways as the experience of death and income loss unfolds. In particular, the longer the restrictions last in both countries, the bigger the income losses and the less likely are these losses to have been anticipated, leading to a reduction of the high relative valuation of health over wealth. The subgroup effects further suggest that as income losses accumulate, those who are most vulnerable to and fearful of COVID-19 will reduce their relative valuation of health over wealth the most. This has important policy implications for the current but also future pandemics and other public health threats. The public will likely become more willing to countenance increases in deaths as behavioral NPIs are relaxed, the later and the slower is their loosening. This message is reinforced in the UK where this valuation is likely to tilt in the opposite direction if COVID-19 deaths are salient; and this is more likely under an earlier relaxation of such policies.

REFERENCES

Adida, C. L., K. Y. Dionne, and M. R. Platas (2020): "Ebola, elections, and immigration: How politicizing an epidemic can shape public attitudes," *Politics, Groups, and Identities*, 8, 488–514.

Aksoy, C. G., B. Eichengreen, and O. Saka (2020): "The political scar of epidemics," *NBER Working Paper Series*, No. 27401.

Alesina, A., A. Miano, and S. Stantcheva (2020): "The polarization of reality," *AEA Papers and Proceedings*, 110, 324–328.

Bank of England (2020): "Monetary Policy Report, May 2020," https://www.bankofengland.co.uk/-/media/boe/files/monetary-policy-report/2020/may/monetary-policy-report-may-2020.

Barrera, O., S. Guriev, E. Henry, and E. Zhuravskaya (2020): "Facts, alternative facts, and fact checking in times of post-truth politics," *Journal of Public Economics*, 182, 104123.

Bateman, I. J., R. T. Carson, B. Day, M. Hanemann, N. Hanley, T. Hett, M. Jones-Lee, G. Loomes, S. Mourato, D. W. Pearce, et al. (2002): *Economic Valuation with Stated Preference Techniques: A Manual*, Cheltenham, UK: Edward Elgar.

Bol, D., M. Giani, A. Blais, and P. J. Loewen (2020): "The effect of COVID-19 lockdowns on political support: Some good news for democracy?" *European Journal of Political Research*.

Cooper, R., D. V. DeJong, R. Forsythe, and T. W. Ross (1992): "Communication in coordination games," *The Quarterly Journal of Economics*, 107, 739–771.

Ellingsen, T. and R. Östling (2010): "When does communication improve coordination?" *American Economic Review*, 100, 1695–1724.

Falk, A., A. Becker, T. Dohmen, B. Enke, D. Huffman, and U. Sunde (2018): "Global evidence on economic preferences," *The Quarterly Journal of Economics*, 133, 1645–1692.

Falk, A., A. Becker, T. J. Dohmen, D. Huffman, and U. Sunde (2016): "The preference survey module: A validated instrument for measuring risk, time, and social preferences," *IZA Discussion Papers, No. 9674*.

Fetzer, T., L. Hensel, J. Hermle, and C. Roth (2020): "Coronavirus perceptions and economic anxiety," *Review of Economics and Statistics*, 1–36.

Finkelstein, A., E. F. Luttmer, and M. J. Notowidigdo (2013): "What good is wealth without health? The effect of health on the marginal utility of consumption," *Journal of the European Economic Association*, 11, 221–258.

Flaxman, S., S. Mishra, A. Gandy, H. J. T. Unwin, T. A. Mellan, H. Coupland, C. Whittaker, H. Zhu, T. Berah, J. W. Eaton, et al. (2020): "Estimating the effects of non-pharmaceutical interventions on COVID-19 in Europe," *Nature*, 1–8, early view.

Geanakoplos, J. D. and H. M. Polemarchakis (1982): "We can't disagree forever," *Journal of Economic theory*, 28, 192–200.

Grossman, G., S. Kim, J. M. Rexer, and H. Thirumurthy (2020): "Political partisanship influences behavioral responses to governors' recommendations for COVID-19 prevention in the United States," *Proceedings of the National Academy of Sciences*, 117, 24144–24153.

Hall, R. E., C. I. Jones, and P. J. Klenow (2020): "Trading off consumption and COVID-19 deaths," *NBER Working Paper No. 27340*.

Holt, C. A. and S. K. Laury (2002): "Risk aversion and incentive effects," *American Economic Review*, 92, 1644–1655.

Hsiang, S., D. Allen, S. Annan-Phan, K. Bell, I. Bolliger, T. Chong, H. Druckenmiller, L. Y. Huang, A. Hultgren, E. Krasovich, et al. (2020): "The effect of large-scale anti-contagion policies on the COVID-19 pandemic," *Nature*, 1–9.

IHME (2020a): "COVID-19 Projections," https://covid19.healthdata.org/.

——— (2020b): "New COVID-19 Forecasts for Europe: Italy and Spain Have Passed the Peak of Their Epidemics; UK, Early in Its Epidemic, Faces a Fast-Mounting Death Toll," http://www.healthdata.org/news-release/new-covid-19-forecasts-europe-italy-spain-have-passed-peak-their-\epidemics-uk-early-its.

——— (2020c): "New IHME COVID-19 Forecasts Show State-By-State Peaks," http://www.healthdata.org/news-release/ihme-hold-media-briefing-4-pm-eastern-today-details-below.

IMF (2020): *World Economic Outlook, April 2020: The Great Lockdown*, Washington, DC: International Monetary Fund.

NRS (2011): "DC6107SC – Economic Activity by Sex by Age," https://www.scotlandscensus.gov.uk/ods-web/standard-outputs.html.

OECD (2012): *Mortality Risk Valuation in Environment, Health and Transport Policies*, Paris: Organisation for Economic Co-operation and Development.

ONS (2011): "*CT*0486_2011 Census - Sex by Age by Ethnic Group by Economic Activity – National to Region and London Boroughs," https://www.ons.gov.uk.

Tyler, T. R. (2006): *Why People Obey the Law*, Princeton, NJ: Princeton University

Press.

US Census Bureau (2018): "Current Population Survey," https://www.census.gov/cps/data/cpstablecreator.html, Accessed: 2020-04-16.