

BIASED POLICY PROFESSIONALS

Sheheryar Banuri

Stefan Dercon

Varun Gauri

Biased Policy Professionals

Sheheryar Banuri; Stefan Dercon; Varun Gauri

Abstract: Although the decisions of policy professionals are often more consequential than those of individuals in their private capacity, there is a dearth of studies on the biases of policy professionals: those who prepare and implement policy on behalf of elected politicians. Experiments conducted on a novel subject pool of development policy professionals (public servants of the World Bank and the Department for International Development in the UK) show that policy professionals are indeed subject to decision making traps, including the effects of framing outcomes as losses or gains, and most strikingly, confirmation bias driven by ideological predisposition, despite having an explicit mission to promote evidence-informed and impartial decision making. These findings should worry policy professionals and their principals in governments and large organizations, as well as citizens themselves. A further experiment, in which policy professionals engage in discussion, shows that deliberation may be able to mitigate the effects of some of these biases.

JEL classification codes: C90, H83, Z18

Keywords: Biases, decision making, policy professionals, framing, confirmation bias, behavioural economics

Sheheryar Banuri (corresponding author) is an assistant professor at University of East Anglia, Norwich, UK; his email address is: s.banuri@uea.ac.uk. Stefan Dercon is professor at University of Oxford; his email address is: stefan.dercon@bsg.ox.ac.uk. Varun Gauri is senior economist at

the World Bank; his email address is vgauri@worldbank.org. The authors have no relevant or material financial interests that relate to the research described in this paper. The findings, interpretations, and conclusions expressed in this paper are entirely those of the authors and do not necessarily represent the views of the World Bank, its Executive Directors, or the countries they represent. The authors are grateful to Mark Lowcock, Permanent Secretary; the Chief Economist Office; and Research and Evidence Division of the Department for International Development (UK); and the core team of the World Development Report, 2015 of the World Bank; for support and guidance. The authors also thank the participants of the London Experimental Week conference 2015; Behavioral models of Politics conference 2016; the Symposium for Economic Experiments in Developing Countries 2017; and participants in the seminar series at the University of Virginia; Lancaster University; Université Catholique de Louvain (Louvain-la-Neuve); Oxford University; Inter-American Development Bank; the Department of Premier and Cabinet, Victoria, Australia; the associate editor, and three anonymous referees.

1. Introduction

Public policy professionals play an essential role in the design and implementation of policies, programs, and projects across the world. While key decisions are typically taken by elected officials or political appointees, they depend on policy professionals, often civil servants, for policy preparation and advice. These professionals play a central role in translating data and research into policy options and in guiding decision-making. In this paper, we use experiments to study cognitive biases in interpreting data for the purpose of providing advice to decision-makers. The subject pool is novel. We study policy professionals working in the area of international development – in particular, UK civil servants and international civil servants working for the World Bank. The experiments are adapted to the development context from a number of classic studies on biases in decision-making.

The objective and accurate use of data on the part of policy professionals is important for at least two reasons. First, it is almost axiomatic that policies will be more effective when those who design and implement them are able to form accurate beliefs about how the world works and conduct accurate assessments of the costs and effectiveness of policy initiatives. Even if policy professionals sometimes utilize effective heuristics or employ an “ecological rationality,” rather than cost-benefit analysis or expected utility theory or a related set of decision rules, biases in assessing information and calculating value remain important. After all, it is the policy professionals themselves who must adopt and use decision rules, as well as evaluate their effectiveness. Bureaucrats also confront novel cases or situations in which prevailing decision rules do not apply. For these reasons, it is important for policy professionals *to be able* to evaluate data and assess value accurately, even if not every decision requires that capacity.

Second, objectivity, impartiality, and accuracy are often legal or official requirements. Those values are elements of modern bureaucracy, in Weber's (1946) sense. Many governments and organizations maintain the Weberian idea that a bureaucracy is perfected "the more completely it succeeds in eliminating from official business love, hatred, and all purely personal, irrational, and emotional elements which escape calculation. This is the specific nature of bureaucracy and it is appraised as its special virtue." (Weber 1946: 216)

For instance, the UK Constitutional Reform and Governance Act of 2010 sets out a civil service code that provides guidance on its four "core values" for UK public sector employees and that should guide part of our subject pool: "integrity, honesty, objectivity, and impartiality."¹ World Bank staff, who form the other part of our sample pool, are required to act with "integrity, independence and impartiality" in line with their status as employees of an international organization.² The World Bank established its Independent Evaluation Group (IEG) in order to "provid[e] impartial evidence based assessments and lessons on drivers of success and failure."³

To promote the impartial and proper use of evidence, large public organizations recruit and rely on the judgments of well-qualified economists, epidemiologists, environmental scientists, engineers, and other professionals. Having been trained in the natural or social sciences, these professionals are, in theory, well-equipped to conduct and use evidence-based assessments of policies. But there are strong reasons to suspect that professionals in public sector organizations, like most individuals, are subject to and exhibit substantial biases in information processing, assessments of value, and decision making. The psychological literature suggests that

¹ Accessed January 2016: <https://www.gov.uk/government/publications/civil-service-code/the-civil-service-code>

² Staff Manual, accessed August 2016:

http://siteresources.worldbank.org/INTSTAFFMANUAL/Resources/StaffManual_WB_web.pdf, p. 3.

³ Taken from the 2015 Independent Evaluation Group annual report (accessed May 2016): <http://ieg.worldbankgroup.org/Data/ar2015-full.pdf>

biases in decision-making, both on the part of the general public and experts, are widespread (DellaVigna and Pope, 2016; Malmendier and Tate, 2008; Cooke, 1991; Shanteau, 1992; Shanteau, 1988; Englich, Mussweiler, and Strack, 2006; Tversky and Kahneman, 1981; Langfeldt, 2004; Stewart and Stasser, 1995; Bero and Jadad, 1997; Herek, Janis, and Huth, 1987; Calvert, 1985; among others). If bureaucrats and policy professionals are as biased as the general population, meritocracy and good recruiting processes will not protect against partiality and subjectivity.

Large public organizations implement a number of procedural safeguards to further promote objective and impartial decision making. Often, they encourage and sometimes require peer review and deliberation, cost-benefit analysis and other kinds of ex-ante policy scrutiny, and the ex-post evaluation of projects and programs on the basis of randomized controlled trials and other methods. It is unclear how effective these procedural safeguards are. In practice, policy professionals are often adept at identifying sympathetic peer reviewers and decision makers, influencing the direction of a deliberative meeting by setting the agenda in particular fashion, burying contestable assumptions of cost-benefit calculations in a thicket of footnotes and appendices, and carefully curating evaluations most relevant to a particular project or decision. More generally, because most principals, including elected politicians in charge of final decision-making, have little idea how biases affect decision making in their own organizations, they are not able to match procedural safeguards to the most significant or prevalent cognitive biases, let alone measure the effectiveness of those procedures.

This paper presents the results of a survey designed to identify decision making biases within a sample of development professionals in two prominent organizations: The World Bank and Department for International Development in the UK (the UK government department

responsible for development policy and spending the foreign aid budget, “DfID” hereafter). The survey used a series of experiments adapted to the development context and consistent with the kinds of decisions staff in these organizations are asked to make. The survey focused on confirmation bias and the effects of framing on risk aversion. This paper is, to our knowledge, the first examination of these cognitive biases in large public sector organizations. The results show that staff engage in biased decision-making, including apparent bias correlated with ideological priors.

The next section introduces the survey and procedures used. The succeeding sections assess, in turn, confirmation bias and the effect of framing on risk aversion. The final section presents the results of an experiment designed to assess the extent to which deliberation can mitigate some of these biases, and also presents a brief summary and concluding remarks.

2. Survey design and data collection

The surveys were conducted entirely online, through email invitations via SurveyMonkey.com. The underlying population was all World Bank and DfID full-time staff of professional grade levels⁴, both in headquarters (DC and London, respectively) and in country offices for each organization. A random sample for each organization was selected in the following manner: the employee roster for World Bank and DfID staff (including names, email addresses, staff identification numbers, and location) were obtained and, after the population was split into two groups (headquarters and country offices), representative samples were drawn from each group.⁵

⁴The specific professional grade levels are GE through GL for the World Bank, and B2 through SCS for DfID.

⁵ The samples were drawn based on a 95% level of confidence, and a ± 4 confidence interval.

In order to increase the motivation of World Bank staff to participate in the survey, all respondents were offered a free coffee mug. No such incentive was provided to DfID staff. Recruitment was conducted in three waves, with each wave lasting one week. The survey was fairly extensive and took 30-40 minutes to complete. Invitations were sent to 4,724 World Bank and 1,148 DfID staff. At the end of the study period, 2,053 responses were received from the World Bank (response rate 43%) and 825 responses from DfID (response rate 72%).⁶ This yielded an overall sample of 2,878 respondents across the two organizations. The survey adapts classic behavioral experiments, as described in the corresponding sections below. Each treatment in a given experiment was randomly and independently assigned from the other experiments. Table 1 provides summary statistics for our sample, distinguishing between DfID and World Bank respondents.

[TABLE 1 HERE]

3. Confirmation bias

A longstanding finding from social psychology is that the general public uses evidence partially, interpreting findings in light of the symbols or metaphors they invoke, or in accord with the views of respected opinion leaders (Lord, Ross, and Lepper 1979). This is a specific, and socially contextualized, instance of confirmation bias, in which individuals selectively seek, remember, and prefer information in a manner that confirms prior views. The social problems arising from confirmation bias are of particular concern in public discussions of scientific evidence.

⁶ We define a valid response as the respondent completed at least one question in the survey. Since different vignettes came at different points of the survey, the total number of observations fluctuate across experiments.

Jelveh, Kogut, and Naidu (2015) find a striking correlation between the political ideology of economists and the policy-relevant parameter estimates that the economists report. The authors extract political preferences from economists' academic writings using machine learning techniques. The measure is correlated with political campaign contributions (i.e. Democrat or Republican donations of more than \$200 over the 1979-2012 period) and political petition signing activity (based on 35 petitions signed mainly by Economists as reported by Hedengren et al., 2010), which suggests the validity of the measure. The authors find that economists with certain political preferences cluster in particular fields and departments. Using all policy-relevant papers that estimate elasticities (i.e. papers reporting taxable income elasticities, labour supply elasticities, fiscal multipliers, labor demand elasticities, and intergenerational income elasticities), the authors find that political ideology, estimated using all the authors papers published *before* the policy-relevant paper was published, predicts the magnitude of an elasticity (i.e. reverse causality is mitigated by using only the papers published prior to the policy relevant paper). In other words, political preferences appear to affect published, scientific findings. Sunstein (2006) find that the political composition of judicial panels significantly influences judges' opinions. The same judge, whether appointed by a Democrat or a Republican, exhibits more ideological voting patterns when sitting on a panel with politically homogeneous judges than with ideologically mixed panels. Together, these studies show that technical expertise does not by itself resolve the problem of ideologically motivated or socially influenced confirmation bias.

In order to test the presence of confirmation bias among our sample of policy professionals, we designed an experiment in which we asked them to assess data from the evaluation of an intervention. The experiment we conducted was an extension of Kahan et al.

(2017). In the original experiment, respondents were asked to evaluate the outcome reported by a study that generated the following frequency table:

[TABLE 2 HERE]

Kahan et al. (2017) conduct four main treatments, all of which presented these exact data but changed the framing and outcomes of the study. Two treatments addressed the effectiveness of a skin rash cream (low prior biases); while the remaining two treatments addressed the impact of gun control laws on crime (high prior biases). Respondents were asked to interpret the conclusion of the study they saw, based on the data presented in the table. The numbers used in the table are constructed so that relying on the most accessible shortcuts (i.e. using only the first row or only the first column of the table, rather than the full table) leads to the incorrect answer.⁷ Kahan et al. (2017) found that respondent accuracy was higher with the skin cream framing treatments than with the gun control framing treatments. Furthermore, the authors found that responses were biased in favour of ideological priors.

Our study used a similar approach but substituted treatments framing the study in terms of minimum wage laws, in place of gun control laws. We did this as attitudes toward the minimum wage are likely to exhibit more ideological variation among development practitioners than attitudes toward gun control. The comparison frame in our study focused, like Kahan et al. (2017), on skin cream.⁸ In each frame, the numbers presented to respondents were identical, but the labels were changed to reflect the framing. Following Kahan et al. (2017), within each frame

⁷ According to Wasserman et al. (1990) and Kahan et al. (2017), most individuals employ one of two heuristics to arrive at the (incorrect) answer. The first is to compare the numbers in the top left cell with the top right cell. The second is to compare the numbers in the top left cell with the bottom left cell. The former neglects information from the control, while the latter neglects information needed to disentangle the effect from confounds.

⁸ We replicate their skin cream treatment, but recognize that gun control laws would work well for a random sample of US citizens, but less well with development professionals, as they are more liberal. We use a “minimum wage” treatment, where we utilize a minimum wage frame.

we implemented two treatments, randomly switching the labels for “Good Outcome” and “Bad Outcome” so that the data supported either one or the other policy conclusion (or statement regarding clinical outcomes). The presence of confirmation bias (to selectively “seek data that are likely to be compatible with the beliefs [people] currently hold” – Kahneman, 2011) would predict that our respondents would be more likely to make errors in the minimum wage treatments than in the skin cream treatment, as our sample are more likely to have biases in favor of, or against, minimum wages.⁹ The use of System 1 (a la Kahneman, 2011) favors the use of heuristics leading to an interpretation of the data corresponding to the respondents’ prior beliefs. Moreover, we would expect individuals with stronger ideological priors to be more likely to make mistakes when the data support the opposing conclusion.

Respondents in the skin cream frame saw the following prompt:

Medical researchers have developed a new cream for treating skin rashes. New treatments often work but sometimes make rashes worse. Even when treatments don't work, skin rashes sometimes get better and sometimes get worse on their own. As a result, it is necessary to test any new treatment in an experiment to see whether it makes the skin condition of those who use it better or worse than if they had not used it. Researchers have conducted an experiment on patients with skin rashes. In the experiment, one group of patients used the new cream for two weeks, and a second group did not use the cream. In each group, the number of people whose skin condition got better and the number whose condition got worse is recorded in the table below. Because patients do not always complete the studies, the total number of patients in each

⁹ Adding pecuniary incentives would be expected to increase accuracy rates overall, and potentially even reduce differences between frames, but it is not clear whether this difference would be eliminated as a result of incentives, particularly if one frame is more likely to trigger the heuristic.

two groups is not exactly the same, but this does not prevent the assessment of the results. Please consider two statements about this study:

(a) People who used the skin cream were more likely to GET BETTER than those who didn't

(b) People who used the skin cream were more likely to GET WORSE than those who didn't

Which statement (above) is the study most consistent with?

The labels on the rows (in table 2) were replaced with “Patients who *did* [did not] use the new skin cream” and the labels on the columns are replaced with “Rash got *better* [worse]”.

Respondents in the minimum wage frame saw the following prompt:

A decentralization reform gave local jurisdictions in an upper-middle income country authority over the minimum wage. Some raised the minimum wage, and others left it unchanged. Because of natural barriers dividing the localities, there was little population mobility in response to the changes. Capital, however, was mobile. Some believe that increasing the minimum wage tends to raise the income of the poorest 40%. Others think that raising the minimum wage slows business growth so much that the incomes of the poorest 40% tend to fall. To examine this question, researchers at a major university measured the number of jurisdictions in which the incomes of the poorest 40% rose, and the number in which the incomes fell, four years after the reform. Please consider two statements about this study:

(a) The income of the poorest 40% of the population FALLS when the minimum wage is increased

(b) The income of the poorest 40% of the population RISES when the minimum wage is increased

Which statement (above) is the study most consistent with?

As in the original study, respondents were asked to choose which statement best reflected the study findings. The labels on the rows (in table 2) were replaced with “Localities that *did* [did not] increase the minimum wage” and the labels on the columns were replaced with “Income of poorest 40% *rose* [fell]”. Respondents were randomly assigned to either the skin cream or the minimum wage frame, and whether the data supported statement A or statement B.¹⁰

Respondents were also asked to assess the quality of the study (on a 10-point scale). In addition to this, the study collected data on the respondents’ ideological orientation with regard to redistribution (adapted from the World Values Survey): *On a scale of 1 to 10, where do your views fall: 1 = “Incomes should be made more equal”; 10 = “We need larger income differences as incentives for individual effort”*. This question measures respondents’ ideological orientation with respect to redistribution, a key predictor of support for/against minimum wage laws.¹¹

If respondents evaluate data objectively and independent of prior beliefs regarding redistribution, they should have offered equally accurate assessments of the minimum wage study and skin cream study frames. Alternatively, if respondents are influenced by their

¹⁰ Since we were primarily interested in the responses to the minimum wage frame, we assigned 20% of respondents to the skin cream frame, while 80% were assigned to the minimum wage frame.

¹¹ Ideological orientation was measured towards the end of the extensive survey. In the survey, respondents were exposed to the vignettes first, then socio-demographic questions were recorded, followed by political orientation. As mentioned, while the treatments within experiments were randomized, the order of the questions was not.

ideologies or values, accuracy in the minimum wage frame should have been lower than in the skin cream frame. In fact, respondents were significantly less accurate in the two minimum wage treatments (45% responded with the correct answer) relative to the two skin cream treatments (65% responded with the correct answer: two sample proportions test: $p < 0.01$), and relative to random guesses ($p < 0.01$) – see figure 1. This suggests that development professionals exhibit a bias when interpreting data on ideologically charged interventions.¹²

[FIGURE 1 HERE]

Table 3 displays the results of probit regressions estimating the likelihood of providing a correct response. Model 1 includes a dummy variable equal to 1 if the respondent was exposed to the minimum wage frame, while model 2 includes controls for mathematical ability using the Cognitive Reflection Test (Frederick, 2005) and respondent assessment of the study (to control for the methods used in the two study frames, one being experimental and the other observational). Model 3 includes controls for socio-demographic variables (age and gender). Model 4 includes organizational controls (dummy variable for whether the respondent works for the World Bank or for DfID, whether the respondent was posted at HQ or a country office, and the salary grade of the respondent). Finally, model 5 includes respondent expertise relevant for the minimum wage frame (whether the subject of the respondents' highest degree was economics or development¹³, and whether the respondent currently works on poverty¹⁴).

¹² Recall that within each frame (skin cream or minimum wage), the data support either income (rash) improving, or income (rash) getting worse. 41% of the respondents report the correct answer when the data support income improving (58% accuracy for the rash); while 48% of the respondents report the correct answer when the data support income declining (72% accuracy for the rash getting worse).

¹³ In this and the subsequent experiments we report, the substantive results remain unchanged when all degree options are included, and not only degrees in economics or development.

¹⁴ The poverty expert variable was constructed based on responses to the question: "What best describes the focus of your work?". The variable takes on a value of 1 if the subject responded with poverty-relevant responses: "Poverty"; "Social development" or "Social protection" (World Bank), or "Economics Advisor"; "Humanitarian Advisor"; "Livelihoods Advisor"; or "Social Development Advisor" (DfID).

[TABLE 3 HERE]

Table 3 shows support for the finding that ideologically charged questions reduce accuracy in the interpretation of data among development professionals. Respondents were 20% less likely to provide the correct response when exposed to the minimum wage frame ($p < 0.01$), across all specifications (please see table A.1 in appendix A for marginal effects). As expected, respondents with higher scores on the Cognitive Reflection Test (related to higher mathematical ability) were significantly more likely to provide the correct response ($p < 0.01$). In addition, economics majors were significantly more likely to provide the correct answer in both study frames, suggesting that social science training helps professionals interpret data accurately ($p < 0.05$). Furthermore, respondents that assessed this study to be of a higher quality were also more likely to provide the correct response, though this variable was not robust to changes in model specification. Finally, there were no differences between DfID and World Bank respondents ($p = 0.66$) and between duty stations (HQ vs. country office: $p = 0.94$).¹⁵

Recall that the minimum wage frame question had two treatments with two different correct answers (depending on the labelling of the columns). Half of the respondents were provided with data that support the conclusion that the income of the poorest *falls* when minimum wages are increased (“Income Falls”), while the other half were provided with data supporting the conclusion that the income of the poorest *rises* when minimum wages are increased (“Income Rises”). This structure was identical for the skin cream frame, as well.

¹⁵ In addition to the above, we also test for whether the treatment effects vary by training (economics), or by area of expertise (poverty experts). No significant differences are found, indicating that economists are significantly more accurate overall, and that poverty experts show no significant differences across treatments.

Combining these treatments with the question about equality preference permits a direct test of the influence of ideology on data interpretation.

Table 4 tests confirmation bias. It presents the results of probit regressions for each of the four treatments. The dependent variable takes a value of 1 if the respondent provided the correct interpretation of the data. Models 1 and 2 correspond to the ideological frame (minimum wage), with model 1 reflecting the treatment where the data support the bad outcome (income falling) and model 2 reflecting the treatment where the data support the good outcome (income rising). There are corresponding good and bad outcomes in the skin cream treatments (the skin cream eliminates rash and does not eliminate rash). The estimations include controls for inequality preferences, mathematical ability, the respondent's assessment of the study, age, gender, and organization. The main relationship of interest is the effect of inequality preferences on getting the correct answer. If respondents rely on their priors, rather than evaluating the frequency table carefully, one should expect to see that the higher the respondents' preferences for inequality (and hence, their anticipated opposition to minimum wage laws), the *more likely* they are to respond correctly when the data support *income falling*, and the *less likely* they are to respond correctly when the data support *income rising*. Since this preference is irrelevant for evaluating the effectiveness of skin creams, one would expect to see no relationship in both treatments using the skin cream frame.

This is precisely what we find. Model 1 corresponds to the treatment where the data support the finding that minimum wage laws lead to the income of the poorest falling. In that model, there is a positive and significant relationship between inequality preferences and the correct interpretation of the data ($p < 0.05$). Respondents with extreme ideological orientations (Preference for inequality = 1) were 6.3 percentage points less likely to provide the correct

interpretation of data relative to respondents with neutral views (Preference for inequality = 5.5).¹⁶ Similarly, in model 2, where the data support the finding that minimum wage laws increase the income of the poorest, there is a negative and significant relationship between inequality preferences and correct interpretation of the data ($p < 0.05$). In this frame, respondents with extreme ideological orientations (Preference for inequality = 10) were 7.2 percentage points less likely to provide the correct interpretation of data relative to respondents with neutral views (Preference for inequality = 5.5). This means that when the priors and the results match, respondents were more likely to interpret the data correctly. By contrast, there is no evidence of this relationship in the skin cream frame. In models 3 and 4, both coefficients are not significantly different from 0 ($p = 0.85$ and $p = 0.77$ respectively). The results are similar across the two institutions across all four treatments ($p > 0.20$).¹⁷

[TABLE 4 HERE]

4. Framing and Risk

The framing of information affects the perceptions of risk, as well as decisions to take risky actions on behalf of others. The latter phenomenon appears linked to social preferences (Eckel and Grossman 2002; 2008; Song, 2008; Chakravarty et al 2011; Anderson et al 2012; Bradler 2009).

Several of these studies show that an individual's risk tolerance increases when taking risk on behalf

¹⁶ The changes in the probability of a mistaken judgement were calculated using marginal effects, and taking the difference in predicted probability between respondents with extreme views (at the tails of the distribution) and respondents with neutral views (in the middle of the distribution). Under the Income falls ideological frame, respondents were 1.4 percentage points *more* likely to provide the correct interpretation of the data as their ideological orientation moved toward preferring income inequality. Under the Income rises ideological frame, respondents were 1.6 percentage points *less* likely to provide the correct interpretation of the data as their ideological orientation moved toward preferring income inequality.

¹⁷ One concern that might bias our results is that of multicollinearity. However, none of the correlations between the independent variables in table 4 are particularly high, and the highest variance inflation factor of the linear probability model is 1.12. Nevertheless, to test the robustness of our findings we reran models 1 and 2 in table 4 by (1) using bivariate regressions with preference for inequality as independent variable, and (2) dropping socio-demographics (age, gender, and organization). We find that the coefficients on preference for inequality are stable and robust to changes in model specification.

of others, in comparison to the willingness to take risk when making decisions for oneself. In other words, individuals become more risk seeking when they are playing with house money, rather than their own. But these findings seem contrary to the common perception that public sector institutions, and the bureaucrats and professionals in them, are unusually risk averse, and more risk averse than the individuals who compose them.

To explore this issue, we conducted an experiment that examined the effect of framing on risk preferences. This was a replication, in our novel subject pool, of a classic experiment from Tversky and Kahneman (1981) that examined whether loss or gain framing affects willingness to take on risk, but importantly, with a health context relevant for policy professionals during the Ebola virus epidemic (2014). We replicated the original Tversky and Kahneman (1981) experiment (with one difference, see fn. 20), and asked respondents the following question:

Suppose your country is preparing for a new disease that is expected to infect 12,000 people. Scientists have come up with two treatments – let's call them Treatment 1 and Treatment 2. Here is the internationally validated scientific evidence on the effectiveness of the treatments.

[Gain frame]

If people take Treatment 1, then 4,000 people will be saved. [Safe choice]

If people take Treatment 2, then there is 1/3 probability that 12,000 people will be saved and 2/3 probability that no one will be saved. [Risky choice]

Which treatment do you think the health authorities should implement?

[Loss frame]

If people take Treatment 1, then 8,000 people will die. **[Safe choice]**

If people take Treatment 2, there is 1/3 probability that no one will die and 2/3 probability that 12,000 people will die. **[Risky choice]**

Which treatment do you think the health authorities should implement?

This vignette asked respondents to decide between two alternative medical treatments with the same expected number of lives saved, but one was safe and the other entailed risk. The experiment randomly assigned respondents to either a frame emphasizing gains (“will be saved”)¹⁸ or losses (“will die”). If framing does not matter, then respondents’ choices are guided by their own risk preferences, and one would expect no difference in the proportion of respondents selecting the risky policy. However, if respondents are prone to biases arising from framing (and prospect theory more generally), one would expect respondents to be more risk-seeking in losses relative to gains (Tversky and Kahneman, 1981).¹⁹

Figure 2 reports the results of the experiment across the two samples (pooled). 22% of the respondents assigned to the gains frame (n=1,314) selected the risky policy option while 65% of the respondents assigned to the losses frame (n=1,277) did the same. This difference in proportion is significant (two sample proportions test: $p < 0.01$).²⁰

¹⁸ An important difference between our vignette and the one used by Kahneman and Tversky (1981) is that the gains frame we used does not specify that those people infected will die. Hence, the vignette could be interpreted as those that are not saved will be infected, but will recover, instead of dying for sure. If respondents interpret the gains frame in this fashion, then the respondents would be more likely to select a risky policy relative to the gains frame explicitly mentioning death (as the perceived gains in the infection framing will be higher). Nevertheless, we still observe a stark treatment effect of the frame.

¹⁹ While we conducted this experiment in both the World Bank and DfID, we only collected risk preferences in the DfID sample. Hence, we do not control for risk preferences in the subsequent analysis. Controlling for risk restricts our sample to DfID only, but our core result remains robust in the DfID sample alone.

²⁰ These results are even more striking when compared to the original results reported in Tversky and Kahneman (1981): They are extremely similar. In the original paper, 28% of the sample chose the risky policy in the gains

[FIGURE 2 HERE]

Table 5 displays the results of probit regressions estimating the likelihood that a respondent chose the risky policy option. Model 1 includes a dummy variable equal to 1 if the respondent was exposed to the losses frame, while model 2 includes controls for socio-demographics (age and gender). Model 3 includes organizational controls (dummy variable for whether the respondent works for the World Bank or for DfID, whether the respondent is posted at HQ or a country office, and the salary grade of the respondent). Model 4 includes respondent expertise relevant for this vignette (whether the subject of the respondents' highest degree was economics or development, and whether the respondent was working on health²¹).

[TABLE 5 HERE]

From table 5, we find strong evidence for the presence of biases arising as a result of framing. Using marginal effects (see table A.2 in appendix A), respondents were 45% more likely to select the risky option when the decision problem was framed as a loss. The result was robust to a number of additional controls, including socio-demographics, organization specific variables, and prior expertise. Respondents from the World Bank were marginally more likely to select the risky option ($p < 0.10$), but this was not robust to alternate specifications. In sum, framing dramatically affects how development professionals perceive risk. While our results are very similar to Tversky and Kahneman (1981), we do note that it is striking that policy professionals, who provide advice based on similar types of evidence, are just as susceptible to the framing of the crisis as the college student subjects in Tversky and Kahneman (1981). The

frame (22% of our sample did the same; $p = 0.12$); while 78% chose the risky policy in the losses frame (65% of our sample did the same; $p < 0.01$).

²¹ The health expert variable was constructed based on responses to the question: "What best describes the focus of your work?" The variable takes on a value of 1 if the subject responded with the health-relevant response: "Health" (World Bank), or "Health Advisor" (DfID).

period of our study overlaps with the Ebola virus epidemic of 2014, suggesting that the manner in which the crisis was framed could have had an impact on the decisions made by policy professionals (and hence, policymakers).

5. Summary and discussion

These experiments with World Bank and DfID staff are the first to systematically document the presence of biases among policy professionals. The data suggest that despite the non-political charter of many bureaucracies, such as the World Bank and DfID, and despite the fact that public institutions are designed to promote objective and impartial decision making, significant biases in decision-making are evident. This finding should worry policy professionals and their principals in governments and large organizations, as well as citizens themselves.

As mentioned at the outset, some bureaucracies implement procedures that attempt to address cognitive biases in decision making. It is unclear how effective these procedures are, in significant part because research on the biases of policy professionals and bureaucrats is scarce. Further exploration of the extent and magnitudes of these biases could inform the design of bureaucratic procedures to address them. Among the procedures that may go some way in mitigating biases among bureaucrats are “red teaming” major decisions (e.g., implementing mock adversarial arguments, as in playing the “devil’s advocate,” or war games to identify the strengths and weakness of different courses of action or points of view), “dogfooding” products and services (e.g., sampling the products and services that consumers or citizens use before rollout), prediction tournaments, and group deliberation (Tetlock and Gardner 2015, World Bank 2015).

Hastie and Sunstein (2015) explore the effects of group deliberation and interaction on the quality of decision making. They argue that group interaction is helpful for “eureka” problems (when the right answer, once announced, is clear to all), and can mitigate the effects of availability heuristics and anchoring. On the other hand, group deliberation can increase polarization and reduce variance in the views that individuals hold (“groupthink”), largely as a result of informational and reputational cascades – individuals do not want to appear to be out of step with the views of others. Group deliberation also tends to increase the confidence, or overconfidence, of individuals in their own judgments and beliefs.

To explore the effects of one potential intervention, group deliberation, we conducted a follow-up experiment with a small sample of DfID policy professionals. The follow-up experiment was conducted during an annual retreat for DfID economists. We asked this sample to respond to a survey twice: first (prior to the morning session of the retreat), the policy professionals were exposed, individually, to the confirmation bias and loss versus gain framing experiments described above. Then (prior to the afternoon session of the retreat), they met in pairs and were asked, again, to provide answers following a brief period of deliberation. This experiment was conducted using pen and paper (rather than online as in the previous experiments reported above). Due to the small sample, we restricted the number of treatments: there were two versions of the survey, each containing a single sequence of treatments for the confirmation bias and framing experiments. Version 1 contained the gains frame and the confirmation bias vignette data supporting income falling as a result of the intervention. Version 2 contained the losses frame and confirmation bias vignette data supporting income rising as a result of the intervention. 81 participants were randomly allocated to either version 1 or 2 of the survey at the beginning of the retreat. In the afternoon, the participants that were given version 1 of the survey

were randomly paired and given version 1 again and asked to discuss before providing their answers (and the same for version 2). Hence, each individual completed the same version of the survey twice, once individually, and once with a partner. Three participants left the retreat early to meet other commitments, so there were a total of 39 groups in the partner phase of the survey.

[FIGURE 3 HERE]

Figure 3 presents the main results of the experiment for our two biases of interest. The confirmation bias result (left) pools both treatments together and finds an overall increase in accuracy of 12%.²² For the framing vignette (right), participants exposed to the losses frame were 17% more likely (relative to gains) to choose the risky policy option in the absence of deliberation. With deliberation, there is no reduction in this difference: in fact, it becomes slightly and insignificantly larger. Hence, we find that group deliberation shows some evidence of mitigating biases associated with confirmation bias. It did not, however, have any effect on the biases associated with framing.

The reason for this, we suspect, is that the confirmation bias experiment is a “eureka” problem. There is a “right” answer to the minimum wage experiment, in the sense that in each treatment the data are consistent with one interpretation of the data and inconsistent with the other. The experiment regarding the framing of losses and gains is different. The main finding relies on how different respondents (or pairs of respondents) answer differently under different frames, not on whether individuals or pairs provide the correct answer in a given frame. Since the responses then are simply based on preferences, deliberation may well aggregate the preferences within groups, but does not yield any bias mitigation.

²² This increase in accuracy, while modest, could also be explained by experience: respondents participate in the survey twice, and may have engaged in the calculation once again independent of any deliberation effects. Our design does not allow us to test for the effects of experience independently of deliberation, however.

To summarize, deliberation seems to be a more effective safeguard for some problems than others. In general, when organizations implement procedural safeguards to reduce biases and improve decisions among policy professionals, they need to be cognizant that the safeguards may be effective for some biases and ineffective (or even counterproductive) for others. Group deliberation might reduce confirmation bias, but have no effect on the framing of risk, and could even increase a policy professional's risk aversion if the group makes risk-aversion in the authorizing environment more salient.

Psychological processes and social contexts can have dramatic effects on the ways in which policy professionals use and interpret data, make project decisions, and take personal and organizational risks. Public sector organizations rarely study and collect data on the biases of their staff. As a result, they are often unaware of the extent to which these influences affect their organizations' choices and policies, and cannot develop evidence-based approaches to mitigate biases. In a different context, Pope, Price, and Waters (2018) report the findings from a natural experiment on in-group bias among Basketball referees, exploiting the years between which a study (Price and Waters, 2010) was published as a working paper, and when the findings were reported in the media. The authors found no evidence of ingroup bias following the media coverage. Pope, Price, and Waters (2018) find that once individuals were made aware of their biases, they corrected them. Efforts to study and understand the decision making processes of policy professionals could help governments and international organizations devise more effective policies and programs, and to live up to legal and professional ideals of objectivity, impartiality, and scientific accuracy.

Appendix

Table A.1.
Framing effects on the interpretation of data (marginal effects)

Dependent variable: correct response to vignette (= 1)					
	I	II	III	IV	V
Minimum wage treatments	-0.206*** (0.02)	-0.207*** (0.02)	-0.207*** (0.02)	-0.213*** (0.03)	-0.203*** (0.03)
Cognitive reflection test score		0.050*** (0.01)	0.047*** (0.01)	0.049*** (0.01)	0.051*** (0.01)
Study rating by respondent (10 = Extremely strong)		0.006 (0.01)	0.006 (0.01)	0.009 (0.01)	0.015** (0.01)
Age (in years)			0.000 (0.00)	-0.001 (0.00)	-0.001 (0.00)
Female			-0.040* (0.02)	-0.028 (0.02)	-0.008 (0.03)
Organization (0 = DfID; 1 = World Bank)				0.014 (0.03)	-0.019 (0.04)
Posting (0 = HQ; 1 = Country office)				0.002 (0.02)	0.002 (0.03)
Respondent grade in organization (1 = Junior; 9 = Senior)				0.020** (0.01)	0.013 (0.01)
Economics major (Subject of highest degree)					0.074** (0.03)
Development major (Subject of highest degree)					0.019 (0.05)
Poverty expert (Respondent works on poverty)					-0.020 (0.05)

Source: Authors' calculations based on survey data collected from World Bank and DfID

Note: Probit regressions. The dependent variable takes on a value of 1 if the respondent selected the response supported by the data, and 0 otherwise. Table reports marginal effects with standard errors in parentheses. The number of observations drops across specifications due to nonresponses on certain questions by respondents.

*** Significant at 1%. ** Significant at 5%. * Significant at 10%.

Table A.2.
Likelihood of respondent choosing risky policy option (marginal effects)

Dependent variable: Respondent chooses risky policy (=1)				
	I	II	III	IV
Losses frame (=1)	0.425*** (0.02)	0.426*** (0.02)	0.433*** (0.02)	0.449*** (0.02)
Age (in years)		0.001 (0.00)	0.000 (0.00)	0.002 (0.00)
Female		-0.018 (0.02)	-0.027 (0.02)	-0.018 (0.03)
Organization (0 = DfID; 1 = World Bank)			0.054* (0.03)	0.023 (0.04)
Posting (0 = HQ; 1 = Country office)			-0.020 (0.02)	-0.026 (0.03)
Respondent grade in organization (1 = Junior; 9 = Senior)			-0.004 (0.01)	-0.014 (0.01)
Economics major (Subject of highest degree)				-0.004 (0.03)
Development major (Subject of highest degree)				0.019 (0.05)
Health expert (Respondent works on health)				-0.040 (0.06)

Source: Authors' calculations based on survey data collected from World Bank and DfID

Note: Probit regressions. The dependent variable takes on a value of 1 if the respondent chose the risky option. Table reports marginal effects with standard errors in parentheses.

*** Significant at 1%. ** Significant at 5%. * Significant at 10%.

References

- Bero, L. A., and A. R. Jadad. 1997. "How consumers and policymakers can use systematic reviews for decision making." *Annals of Internal Medicine* 127(1): 37–42.
- Bradler, C. 2009. "Social preferences under risk-an experimental analysis." *Jena Economic Research Papers* 22: 1–35.
- Calvert, R. L. 1985. "The value of biased information: A rational choice model of political advice." *The Journal of Politics* 47(2): 530–555.
- Chakravarty, S., G. W. Harrison, E. E. Haruvy, and E. E. Rutström. 2011. "Are you risk averse over other people's money?" *Southern Economic Journal* 77(4): 901–913.
- Cooke, R. M. 1991. *Experts in Uncertainty: Opinion and Subjective Probability in Science*. New York: Oxford University Press.
- DellaVigna, S. and D. Pope. 2016. "Predicting Experimental Results: Who Knows What?" NBER Working Paper No. w22566, National Bureau of Economic Research, Cambridge, MA.
- Eckel, C. C., and P. J. Grossman. 2002. "Sex differences and statistical stereotyping in attitudes toward financial risk." *Evolution and Human Behavior* 23(4): 281–295.
- Eckel, C. C., and P. J. Grossman. 2008. "Men, women and risk aversion: Experimental evidence." In *Handbook of Experimental Economics Results*, edited by C. Plott and V. Smith, 1(7): 1061–1073.
- Englich, B., T. Mussweiler, and F. Strack. 2006. "Playing dice with criminal sentences: The influence of irrelevant anchors on experts' judicial decision making." *Personality and Social Psychology Bulletin* 32(2): 188–200.

- Frederick, S. 2005. "Cognitive reflection and decision making." *The Journal of Economic Perspectives* 19(4): 25–42.
- Hastie, R., and C. Sunstein. 2015. *Wiser: Getting Beyond Groupthink to Make Groups Smarter*. Boston, MA: Harvard Business Review Press.
- Herek, G. M., I. L. Janis, and P. Huth. 1987. "Decision making during international crises is quality of process related to outcome?" *Journal of Conflict Resolution* 31(2): 203–226.
- Jelveh, Z., B. Kogut, and S. Naidu. 2015. "Political language in economics." Working Paper 14-57, Columbia Business School.
- Kahan, D. M., E. Peters, E. C. Dawson, and P. Slovic. 2017. "Motivated numeracy and enlightened self-government." *Behavioural Public Policy* 1(1): 54-86.
- Kahneman, D. 2011. *Thinking, fast and slow*. New York: Farrar, Straus and Giroux.
- Langfeldt, L. 2004. "Expert panels evaluating research: decision-making and sources of bias." *Research Evaluation* 13(1): 51–62.
- Lord, C. G., L. Ross, and M. R. Lepper. 1979. "Biased assimilation and attitude polarization: The effects of prior theories on subsequently considered evidence." *Journal of Personality and Social Psychology* 37(11): 2098–2109.
- Malmendier, U. and G. Tate. 2008. "Who makes acquisitions? CEO overconfidence and the market's reaction." *Journal of financial Economics* 89(1): 20-43.
- Price, J. and J. Wolfers. 2010. "Racial Discrimination Among NBA Referees." *The Quarterly Journal of Economics* 125(4): 1859-1887.
- Pope, D. G., J. Price, and J. Wolfers. 2018. "Awareness reduces racial bias." *Management Science* (Forthcoming).

- Shanteau, J. 1988. "Psychological characteristics and strategies of expert decision makers." *Acta Psychologica* 68(1): 203–215.
- Shanteau, J. 1992. "Competence in experts: The role of task characteristics." *Organizational Behavior and Human Decision Processes* 53(2): 252–266.
- Song, F. 2008. "Trust and reciprocity behavior and behavioral forecasts: Individuals versus group-representatives." *Games and Economic Behavior* 62(2): 675–696.
- Stewart, D. D., and G. Stasser. 1995. "Expert role assignment and information sampling during collective recall and decision making." *Journal of Personality and Social Psychology* 69(4): 619–628.
- Sunstein, C. R. 2006. *Infotopia: How many minds produce knowledge*. New York: Oxford University Press.
- Tetlock, P. E., and D. Gardner. 2015. *Superforecasting: The Art and Science of Prediction*. New York: Broadway Books.
- Tversky, A., and D. Kahneman. 1981. "The framing of decisions and the psychology of choice." *Science* 211(4481): 453–58.
- Weber, M. 1946. *From Max Weber: Essays in Sociology*, edited by Hans H. Gerth and Charles W. Mills. New York: Oxford University Press.
- World Bank. 2015. *World Development Report 2015: Mind, Society, and Behavior*. Washington DC: World Bank.

FIGURES

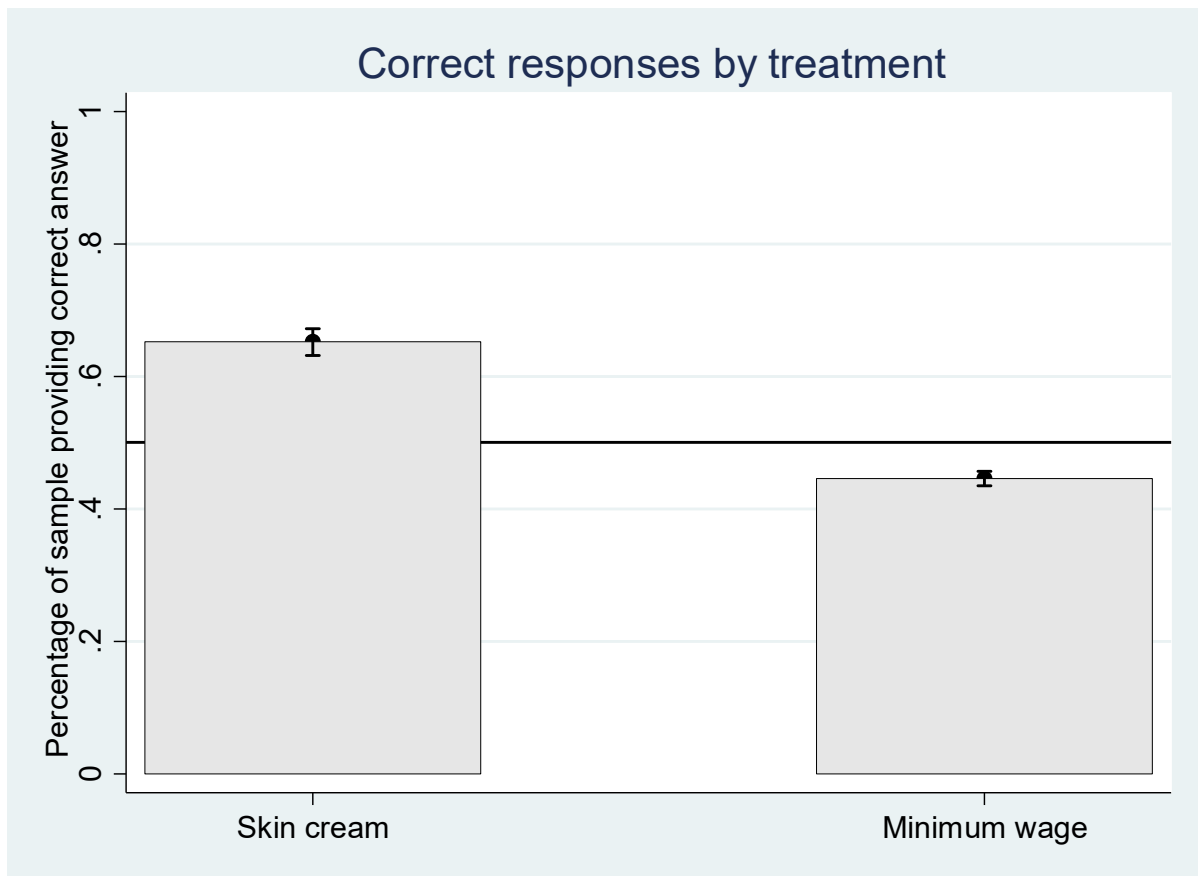


Figure 1. Percentage of respondents reporting correct answer

Source: Authors' calculations based on survey data collected from World Bank and DfID

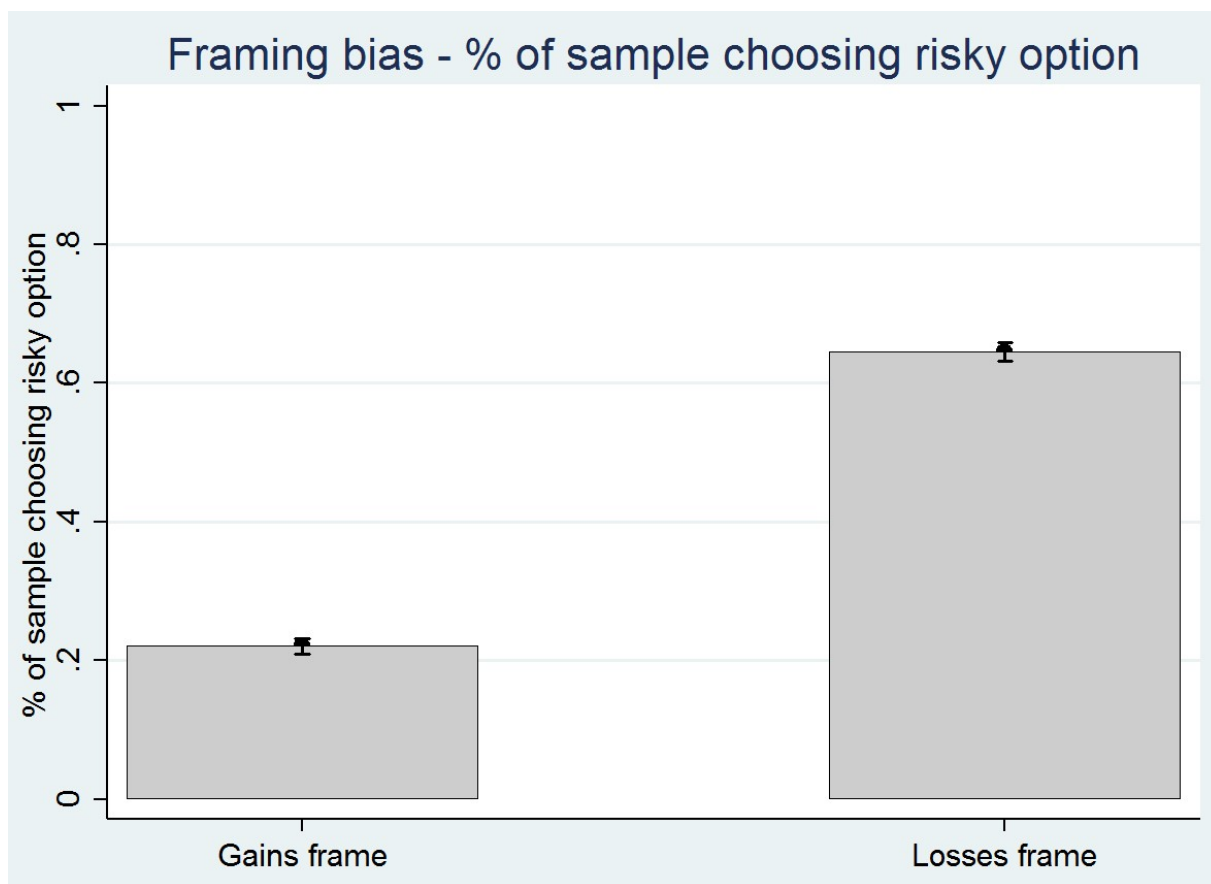


Figure 2. Percentage of respondents choosing risky policy option under different frames

Source: Authors' calculations based on survey data collected from World Bank and DfID

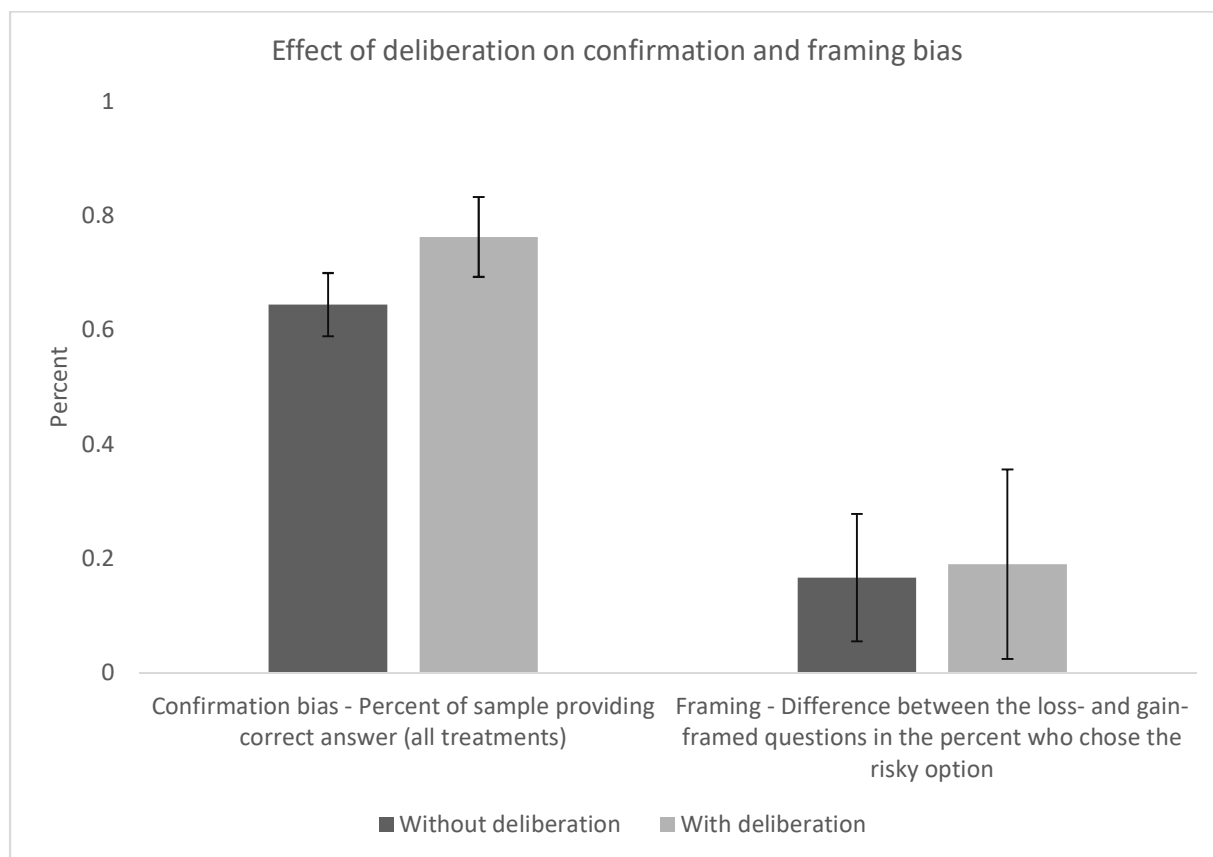


Figure 3. Effects of deliberation on confirmation and framing bias

Source: Authors' calculations based on survey data collected from DfID

TABLES

Table 1.
Summary statistics

	Department for International Development	World Bank
Observations	825	2,053
Age (mean)	42.87	43.71
Female (%)	51.62	41.34
Cognitive reflection test (mean, on scale 0 to 3)	1.76	1.75
Inequality preference (mean, on scale 1 to 10)	4.21	4.91
Risk preference (mean)	3.00	--
Salary Grade (median)	7 (A2)	4 (GG)
Posted at country office (%)	45.69	51.35
Degree		
Economists (%)	15.67	24.05
Development (%)	17.02	4.24
Focus of work		
Poverty and social development (%)	12.01	7.72
Environment and infrastructure (%)	3.91	19.27
Health (%)	3.77	4.16

Source: Authors' calculations based on survey data collected from World Bank and DfID

Note: Inequality preference was based on a 1-10 scale, where 1= "Incomes should be made more equal" and 10= "We need larger difference as incentives for individual effort." The Cognitive Reflection Test (Frederick, 2005) is a three item measure of the respondents' level of cognitive engagement (measured towards the end of the survey instrument). The risk preference measure is based on Eckel and Grossman (2008) with higher numbers indicating higher risk-seeking. The salary grades for DfID range from B2 (low) to SCS (high) while at the World Bank, range from GE (low) to GL (high). The median grade is in line with expectations. We highlight two subjects in which respondents hold their highest degrees and three areas of work that are contextually appropriate to include as controls in the analysis. The poverty expert variable was constructed based on responses to the question: "What best describes the focus of your work?". The variable takes on a value of 1 if the subject responded with poverty-relevant responses: "Poverty"; "Social development" or "Social protection" (World Bank), or "Economics Advisor"; "Humanitarian Advisor"; "Livelihoods Advisor"; or "Social Development Advisor" (DfID). The environment expert variable was constructed based on responses to the question: "What best describes the focus of your work?" The variable takes on a value of 1 if the subject responded with environment-relevant responses: "Environment" or "Infrastructure" (World Bank); or "Climate and Environment Advisor" or "Infrastructure Advisor" (DfID). The health expert variable was constructed based on responses to the question: "What best describes the focus of your work?" The variable takes on a value of 1 if the subject responded with the health-relevant response: "Health" (World Bank), or "Health Advisor" (DfID).

Table 2.
Data presentation template for confirmation bias vignette

	Good Outcome	Bad Outcome
Individuals taking action X	223	75
Individuals taking action Y	107	21

Source: Authors' calculations

Table 3.
Framing effects on the interpretation of data

Dependent variable: correct response to vignette (= 1)					
	I	II	III	IV	V
Minimum wage frame	-0.526*** (0.06)	-0.532*** (0.07)	-0.531*** (0.07)	-0.548*** (0.07)	-0.524*** (0.08)
Cognitive reflection test score		0.126*** (0.02)	0.117*** (0.02)	0.122*** (0.03)	0.128*** (0.03)
Study rating by respondent (10 = Extremely strong)		0.015 (0.01)	0.015 (0.01)	0.023 (0.01)	0.037** (0.02)
Age (in years)			0.000 (0.00)	-0.003 (0.00)	-0.003 (0.00)
Female			-0.100* (0.05)	-0.071 (0.06)	-0.020 (0.07)
Organization (0 = DfID; 1 = World Bank)				0.034 (0.08)	-0.048 (0.11)
Posting (0 = HQ; 1 = Country office)				0.005 (0.06)	0.005 (0.07)
Respondent grade (1 = Junior; 9 = Senior)				0.050** (0.02)	0.034 (0.03)
Economics major (Subject of highest degree)					0.188** (0.08)
Development major (Subject of highest degree)					0.047 (0.12)
Poverty expert (Respondent works on poverty)					-0.051 (0.11)
Constant	0.391*** (0.06)	0.156* (0.09)	0.202 (0.16)	0.058 (0.19)	0.043 (0.23)
Log likelihood	-1825.8	-1605.5	-1572.5	-1461.4	-1103.8
Pseudo R2	0.020	0.029	0.030	0.035	0.035
P	0.000	0.000	0.000	0.000	0.000
Observations	2689	2386	2339	2185	1651

Source: Authors' calculations based on survey data collected from World Bank and DfID

Note: Probit regressions. The dependent variable takes on a value of 1 if the respondent selected the response supported by the data, and 0 otherwise. Table reports coefficients with standard errors in parentheses. The number of observations drops across specifications due to nonresponses on certain questions by respondents.

*** Significant at 1%. ** Significant at 5%. * Significant at 10%.

Table 4.
The effect of ideology on the interpretation of data

Dependent variable: correct response to vignette (= 1)				
	Ideological frame		Neutral frame	
	Income falls	Income rises	Rash worsens	Rash improves
	I	II	III	IV
Preference for inequality (10 = Prefer income inequality)	0.036** (0.02)	-0.041** (0.02)	-0.007 (0.04)	-0.011 (0.04)
Cognitive reflection test score	0.047 (0.04)	0.172*** (0.04)	0.216*** (0.08)	0.117 (0.08)
Study rating by respondent (10 = Extremely strong)	-0.011 (0.02)	0.062*** (0.02)	-0.065 (0.04)	0.072 (0.05)
Age (in years)	-0.004 (0.00)	0.002 (0.00)	0.011 (0.01)	0.002 (0.01)
Female	-0.167* (0.09)	-0.098 (0.09)	0.230 (0.18)	-0.170 (0.18)
Organization (0 = DfID; 1 = World Bank)	-0.092 (0.09)	-0.174* (0.09)	0.094 (0.21)	0.090 (0.19)
Constant	0.090 (0.26)	-0.476* (0.26)	-0.472 (0.55)	0.111 (0.59)
Log likelihood	-644.6	-607.5	-150.3	-140.2
Pseudo R2	0.010	0.032	0.042	0.022
P	0.053	0.000	0.041	0.379
Observations	939	921	235	241

Source: Authors' calculations based on survey data collected from World Bank and DfID

Note: Probit regressions. The dependent variable takes on a value of 1 if the respondent selected the response corresponding to the data, and 0 otherwise. Models 1 and 2 correspond to the ideological (minimum wage) frame, while models 3 and 4 correspond to the neutral frame. Table reports coefficients with standard errors in parentheses. *** Significant at 1%. ** Significant at 5%. * Significant at 10%.

Table 5.
Likelihood of respondent choosing risky policy option

Dependent variable: Respondent chooses risky policy (=1)				
	I	II	III	IV
Losses frame (=1)	1.142*** (0.05)	1.146*** (0.05)	1.170*** (0.06)	1.212*** (0.06)
Age (in years)		0.002 (0.00)	0.000 (0.00)	0.006 (0.00)
Female		-0.047 (0.05)	-0.069 (0.06)	-0.046 (0.07)
Organization (0 = DfID; 1 = World Bank)			0.140* (0.08)	0.058 (0.11)
Posting (0 = HQ; 1 = Country office)			-0.052 (0.06)	-0.067 (0.07)
Respondent grade in organization (1 = Junior; 9 = Senior)			-0.010 (0.02)	-0.037 (0.03)
Economics major (Subject of highest degree)				-0.010 (0.08)
Development major (Subject of highest degree)				0.048 (0.12)
Health expert (Respondent works on health)				-0.103 (0.17)
Constant	-0.770*** (0.04)	-0.839*** (0.14)	-0.777*** (0.18)	-0.864*** (0.22)
Log likelihood	-1524.0	-1475.0	-1361.5	-1020.1
Pseudo R2	0.139	0.141	0.147	0.158
P	0.000	0.000	0.000	0.000
Observations	2591	2512	2337	1768

Source: Authors' calculations based on survey data collected from World Bank and DfID

Note: Probit regressions. The dependent variable takes on a value of 1 if the respondent chose the risky option. Table reports coefficients with standard errors in parentheses.

*** Significant at 1%. ** Significant at 5%. * Significant at 10%.