Appendix A. Recruitment, Research Design, and Additional Tables

Appendix A1: Recruiting Elites at NSCL and NLC

We tested the role of alternative incentives at the annual meetings of the National Conference of State Legislatures (NCSL) and the National League of Cities (NLC) in 2015. The NCSL meeting was held in Seattle, Washington and the NLC meeting in Nashville, Tennessee. At both conferences we rented a booth in the exhibition hall under the banner of Washington University in St. Louis.

During the hours that the exhibit hall was open, 2-3 members of our research team actively recruited attendees in the hall to stop and take a survey. Attendees were told that if they took the survey they would receive a free t-shirt. An example t-shirt was on display at the booth. After giving informed consent and agreeing to participate in the research study, officials sat down at either a computer or a tablet and completed the survey.[1]

The research participants we recruited came from a wide variety of states. At NCSL we recruited 225 officials. A disproportionate number of participants came from the conference's home state of Washington (15.6% of the sample). Still, there was good coverage from the rest of the United States with 48 states represented. At the NLC meeting we recruited 249 participants, 203 of whom were elected officials. Again, the largest group of participants came from the host state, Tennessee (10.0% of the sample), but overall the sample included officials from 43 different states. Table A1 gives the number of participants by state.

The officials who participated in our study were also diverse along other dimensions. At NLC, 28.3% of participants were female, the average age was 57 years old and 72.5% had a bachelor degree or higher. Table A2 summarizes the demographic characteristics of our samples.

[1] A copy of the survey questions and the informed consent are provided in Appendix B.

Table A1: Participants by State

State	NCSL	NLC	State	NCSL	NLC	State	NCSL	NLC	State	NCSL	NLC
AL	0	13	IL	1	6	MT	2	1	PR	1	0
AK	1	2	IN	3	9	NE	9	4	RI	1	0
AZ	2	17	IA	4	0	NV	2	1	SC	4	2
AR	5	3	KS	0	12	NH	3	0	SD	3	2
CA	4	11	KY	8	5	NJ	3	0	TN	5	24
CO	6	5	LA	8	6	NM	5	2	TX	9	11
CT	2	3	ME	3	0	NY	2	1	UT	8	1
DE	1	0	MD	4	3	NC	3	9	VT	2	0
DC	4	1	MA	0	0	ND	6	3	VA	0	5
FL	3	6	MI	1	3	OH	3	8	WA	35	4
GA	1	5	MN	3	3	OK	6	5	WV	1	2
HI	5	0	MS	9	9	OR	2	7	WI	6	2
ID	3	1	MO	6	22	PA	2	2	WY	3	5

Note: Entries are the number of participants by state.

Table A2: Demographic characteristics: NCSL and NLC

	NCSL	NLC	Together
Female (%)	38.3	28.3	33.0
Age (years)	49.7	56.9	53.5
Bachelor's degree or higher (%)	91.2	72.5	81.2
Elected (%)	41.3	81.5	62.4
Democratic (%)	44.9	48.2	46.6
Republican (%)	27.6	31.3	29.5
In majority (%)	46.8	-	-
Non-partisan elections (%)	-	75.9	-

Note: Entries are percentages or average values for the 225 and 249 participants at NCSL and NLC, respectively.

Appendix A2: Experimental Design

We test the effect of two types of incentives relative to a baseline of no incentives. A no incentive category is used as the baseline because one of the enduring debates about experimental methods is whether incentives should be used at all. While many psychologists eschew financial incentives because it can hinder subjects' intrinsic motivation, experimentalists in the economics tradition strongly argue that salient payoffs improve performance and can be used to induce participants to have the same value ordering as assumed by the theory being tested.[2]

We tested the effects of incentives by randomizing respondents into one of three groups. The randomization was done in Qualtrics, the survey platform used to administer the survey. Respondents had an equal probability of being assigned to each of the three groups. Table A3 provides the results of a randomization check. The pre-treatment covariates are balanced across experimental conditions.

The first group received no incentives. These individuals saw the following text before answering the other questions (the text is shown here as it appeared to participating officials):

> *In the next few minutes you will play a game. No money will be given based on your performance.*

A second group was told that they had a chance to earn money for themselves.

> *In the next few minutes you will play a game. Depending on your answers, you will earn money for yourself. After we finish the study, we will randomly select one participant that will receive the cash reward they earned in the game. Hence, the better you perform, the more you can earn.*

The final group played for a charity of their choice. They were presented with the following text:

[2] A recent review of this debate can be found in chapter 10 of Morton and Williams (2010).

Table A3: Covariate balance across treatment groups

	NCSL			NLC		
	(C → S)	(C → Ch)	(S → Ch)	(C → S)	(C → Ch)	(S → Ch)
Female	−0.093	−0.277	−0.109	0.136	0.176	0.045
	(0.247)	(0.230)	(0.226)	(0.232)	(0.232)	(0.226)
Age	−0.014	−0.008	0.006	−0.014	−0.011	0.002
	(0.009)	(0.008)	(0.008)	(0.009)	(0.008)	(0.009)
Education	−0.088	−0.086	−0.086	0.019	0.025	−0.024
	(0.140)	(0.123)	(0.112)	(0.081)	(0.081)	(0.089)
Elected	0.096	−0.299	−0.357	−0.198	0.147	0.346
	(0.288)	(0.260)	(0.264)	(0.265)	(0.286)	(0.285)
Democratic	−0.439	−0.317	0.140	0.302	0.052	−0.340
	(0.332)	(0.300)	(0.280)	(0.285)	(0.272)	(0.297)
Republican	−0.450	−0.224	0.247	0.198	−0.200	−0.477
	(0.375)	(0.335)	(0.319)	(0.297)	(0.288)	(0.320)
Constant	1.520	2.055*	0.455	0.579	0.315	0.017
	(0.945)	(0.877)	(0.807)	(0.742)	(0.700)	(0.752)
Observations	125	150	149	159	160	153
AIC	181.99	209.93	213.71	228.98	231.42	222.90

Note: Entries are probit coefficients with SEs in parentheses. C - Control, S - Self, Ch - Charity. As an example, the first model estimates the probability among NCSL participants of being in the money-self condition rather than the control group, as a function of baseline characteristics. *p<0.1

In the next few minutes you will play a game. Depending on your answers, you will earn money for charity. We will randomly select one participant and pay the earned money to the charity of your choice. Hence, the better you perform, the more can be given to charity.

Which of the charities would you like to play for today?

__ *American Red Cross*

__ *Habitat for Humanity*

__ *A Foodbank in Your Area*

__ *A Group Supporting K-12 Education in Your Area*

__ *The General Scholarship Fund of your Choice*

After being told about their incentives, respondents were asked to answer a series of factual questions about politics. We chose to have the respondents answer factual questions because we wanted them to perform an exercise where we could measure how close they were to the correct answer. Our surveys used questions designed by Bullock et al. (2015). Respondents answered questions using a sliding scale. An example of this survey instrument is presented in Appendix B. One advantage of the sliding scale is that it gives us more variation because it allows us to measure how far respondents were from the correct answer. The eight questions administered in the survey and used in the analysis are presented in Table A4.[3] An answer was coded as correct if it fell within $\pm 5\%$ of the true answer. Each correct answer was worth 50 dollars for those assigned to the treatment groups. To limit the length of the survey we timed the responses, giving participants 20 seconds to answer each question. Before starting the timed portion of the survey, an example question was provided so respondents could understand how to use the sliders when answering the remaining items.

[3]Originally we had intended to ask nine questions but only eight of these nine questions were asked of participants at NLC because of a programming error when preparing that survey. For comparability across treatments we analyze the eight questions that were asked at both sites.

Table A4: Average Responses and Total Number of Respondents by Question

Question Wording	Range	Correct Answer	NCSL Average	NLC Average
From January 2001, when President Bush first took office, to January 2009, when President left office, how had the unemployment rate in the country changed?	-4 to 4 p.p.	3.6 p.p.	0.63 (205)	-0.49 (226)
For every dollar the federal government spent in fiscal year 2014, about how much went to the Department of Defense (U.S. Military)?	3 to 30 cents	17 cents	18.85 (211)	18.02 (227)
In the 2012 Presidential Election, Barack Obama defeated his Republican challenger Mitt Romney. In the nation as a whole, of all the votes cast for Obama and Romney, what percentage went to Obama?	50 to 62%	51.06%	53.20 (209)	53.48 (228)
Approximately 12 to 13% of the U.S. population is Black. What percentage of U.S. Soldiers killed in Iraq since the invasion in 2003 are Black?	9 to 21%	9.90%	16.35 (213)	14.95 (232)
Medicaid is a jointly funded, Federal-State health insurance program for low-income and needy people. For every dollar the federal government spent in fiscal year 2014, about how much went to Medicaid?	3 to 27 cents	7.9 cents	13.40 (213)	12.74 (232)
According to NASA, by how much did annual average global temperatures, in degrees Fahrenheit, differ in 2010 from the average annual global temperature between 1951 and 1980?	-2 to 2 Fahrenheit	1.1F	0.70 (213)	0.37 (236)
About how many U.S. soldiers were killed in Iraq between the invasion in 2003 and the withdrawal of troops in December 2011?	1,000 to 7,000 soldiers	4,486	3728.85 (212)	3571.02 (234)
According to the Census Bureau, in 2010 what percentage of the total population of the United States was born outside of the United States (foreign-born)?	0 to 100%	12.92%	20.18 (213)	24.15 (233)
From January 2009, when President Obama first took office, to May 2015, how had the unemployment rate in the country changed?	-4 to 4 p.p.	-2.3 p.p.	-0.52 (207)	-

Note: Number of valid responses in parentheses. The last question was not asked at NLC due to a programming error. For comparability across treatments we analyze the eight questions that were asked at both sites.

To measure the respondents' performance in answering the factual questions, we started by computing for each item the respondent's distance from the correct answer by using the following formula:

$$\text{Distance from Correct Answer} = \frac{|\text{Correct Answer} - \text{Response}|}{\text{Range of Scale}} \tag{1}$$

To illustrate how this calculation works, suppose that the range of the scale for a question went from 0 to 50 and the correct answer was 20. If the respondent answered 25, then the dependent variable would be equal to 0.10 ($\frac{|20-25|}{50} = 0.10$), indicating that the respondent missed the correct answer by a distance that was equal to 10 percent of the scale. Finally, to create a measure of accuracy by respondent, we took the average error rate across all questions. By using this metric we are able to compare respondents' performance in a meaningful way. On average, the dependent variable took a value of 0.29 (the median is also equal to 0.29).

At the NCSL meeting we also used the survey item designed by Eckel and Grossman (2002) to measure individuals' level of risk aversion.[4] Previous results show that risk aversion increases when subjects have more at stake (Holt and Laury 2005). Our expectation was that we would observe higher levels of risk aversion when individuals played for themselves or for charity. The survey item, which is given in Appendix B, asks individuals to choose among a series of different gambles. Each gamble involves a payoff for a "high" and a "low" roll. After they make their choice a die was rolled to determine the reward. The key fact is that the gambles with the highest expected payoff also have the highest variance. This survey instrument thus allows us to measure whether the type of incentive affects the level of risk aversion shown by the participants. As in the previous task, respondents assigned to the control group faced hypothetical payoffs, while those assigned to the incentives conditions could earn between 20 and 700 dollars.

[4]We did not use this item at NLC because the survey at NCSL was much longer than we had anticipated and we needed to shorten the survey for NLC.

Appendix A3. Additional Tables

Table A5: Total Number of Questions Answered

	(1)
Dependent Variable	# Qs Answered
Model	OLS
Incentive: Money for Self	0.054
	(0.120)
Incentive: Money for Charity	0.102
	(0.115)
Constant	7.55*
	(0.084)
Sample	NCSL & NLC
Observations	465

Note: Entries are OLS coefficients with SEs in parentheses. *p<0.05 (One-sided)

Table A6: The Effect of Incentives on Getting the Correct Answer, by quiz item

	(1)
Dependent variable	Distance from Correct Answer
Model	OLS
Incentive: Money for Self	−0.019*
	(0.010)
Incentive: Money for Charity	−0.019*
	(0.010)
Cluster on Respondent?	Yes
Includes Questions FE?	Yes
Test Sites	NCSL & NLC
Observations	3,537

Note: Entries are coefficients with SEs in parentheses.
*p<0.05 (One-sided)

References

Bullock, John G. Alan S. Gerber, Seth J. Hill and Gregory A. Huber. 2015. "Partisan Bias in Factual Beliefs about Politics." *Quarterly Journal of Political Science* 10 (4): 519-578.

Eckel, Catherine C., and Philip J. Grossman. 2002. "Sex Differences and Statistical Stereotyping in Attitudes Toward Financial Risk." *Evolution and Human Behavior* 23 (4): 281-295.

Holt, Charles A., and Susan K. Laury. 2005. "Risk Aversion and Incentive Effects: New Data without Order Effects." *American Economic Review* 95 (3): 902-904.

Morton, Rebecca B., and Kenneth C. Williams. 2010. *Experimental Political Science and the Study of Causality: From Nature to the Lab.* Cambridge University Press.