E Screenshots

E.1 Main experiment

Respondents in the main experiment were randomly shown four of the five vignettes (in random order). We experimentally vary six features across vignettes (the communication of scientific findings, the statistical significance of the results, whether it includes a high or low or no expert forecast, seniority of the research team, university of the research team, and whether the journal is a general interest or field journal). Five features vary at the respondent-by-vignette level, and one feature varies at the respondent level (whether the main finding includes the p-value or the standard error associated with the main effect). The conditions shown in the following screenshots include a random draw of these six cross-randomized conditions.

E.1.1 Pre-treatment information

Introduction

We will now ask you about your views regarding **four** hypothetical studies. These studies are based on real studies whose details we modified for the purposes of this survey.

We will provide you with a short description of the study design and a summary of the main findings of each study.

→

E.1.2 Marginal effects of merit aid for low-income students

Marginal effects of merit aid for low-income students

Background and study design: 3 PhD students from the University of Illinois conducted an RCT in Texas in the years 2015–2019. The purpose of the RCT was to examine the effects of a randomly assigned $8,000 merit aid program for low-income students on the likelihood of completing a bachelor's degree.

The researchers worked with a sample of 1,188 high school graduates from low-income, minority, and first-generation college households. 594 of those students were randomly assigned to receive $8,000 in merit aid for one year, while the remainder of the students did not receive any additional aid.

Main result of the study: The treatment increased the completion rate of a 4-year bachelor's degree by 1.1 percentage points (p-value = 0.71) compared to a control mean of 17.0 percent.

Publishability

If this study was submitted to the Economic Journal, what do you think is the likelihood that the study would eventually be published there?

Very low likelihood									Very high likelihood	
0	10	20	30	40	50	60	70	80	90	100

Importance

On a scale from 0 to 100, where 0 indicates the "lowest possible importance" and 100 indicates the "highest possible importance," please indicate how **you** perceive the importance of this study.

Lowest possible importance								Highest possible importance		
0	10	20	30	40	50	60	70	80	90	100

Imagine that researchers in this field participated in an anonymous online survey and were asked to evaluate the importance of the study on the same 100-point scale as above (where 0 indicates the "lowest possible importance" and 100 indicates the "highest possible importance").

What importance rating would you expect **these researchers** to give to the study on average?

Lowest possible importance								Highest possible importance		
0	10	20	30	40	50	60	70	80	90	100

→

E.1.3 Long-term effects of equal land sharing

Long-term effects of equal land sharing

Background and study design: A team of 2 PhD students from Northwestern University studied the long-term effects of local changes in inheritance rules for land in Germany in the 19th century. The researchers were interested in whether introducing inheritance rules requiring equal division of land between siblings led to higher average incomes.

The authors use a geographic regression discontinuity design to study the effect of equal division of land on average county-level income. They use data on 387 counties that were at most 35 km away from the border which separated counties with equal versus unequal sharing rules. In 193 counties, inherited land was to be shared or divided equally among children (treatment group), while in the remaining 194 counties land was ruled to be indivisible and had to be passed on to a single heir (control group).

The authors provide evidence in support of the validity of the identifying assumptions: The change in inheritance rules led to a more equal division of land in treated counties. Furthermore, other potential drivers of growth are smooth at the boundary of the discontinuity.

Main result of the study: Average incomes in 2014 were 0.5 percent higher (standard error 2.4) in counties with equal division of land.

Expert prediction: 23 experts in this literature received the same background and study design information as shown to you above. The experts on average predicted a treatment effect of 1.7 percent. The standard deviation of the expert forecasts was 4.7.

Publishability

If this study was submitted to the Review of Economic Studies, what do you think is the likelihood that the study would eventually be published there?

Very low likelihood									Very high likelihood	
0	10	20	30	40	50	60	70	80	90	100

Quality

On a scale from 0 to 100, where 0 indicates the "lowest possible quality" and 100 indicates the "highest possible quality," please indicate how **you** perceive the quality of this study.

Lowest possible quality									Highest possible quality	
0	10	20	30	40	50	60	70	80	90	100

Imagine that researchers in this field participated in an anonymous online survey and were asked to evaluate the quality of the study on the same 100-point scale as above (where 0 indicates the "lowest possible quality" and 100 indicates the "highest possible quality").

What quality rating would you expect **these researchers** to give to the study on average?

Lowest possible quality									Highest possible quality	
0	10	20	30	40	50	60	70	80	90	100

→

E.1.4 Female empowerment program

Female empowerment program

Background and study design: In 2018, a team of 4 PhD students from Columbia University conducted an RCT in Sierra Leone. The purpose of the RCT was to examine whether access to a female empowerment program increased women's labor supply.

In the RCT, 360 women were evenly randomized into a treatment group and a control group. Respondents in the treatment group were offered a female empowerment program, combining both psychosocial therapy and vocational skills training. The program was very intensive: participants attended meetings for up to 5 hours every day during a 12-month period.

Main result of the study: Treated respondents were 1.7 percentage points (standard error 5.0) more likely to take up a job offer compared to a control mean of 37.0 percent.

Expert prediction: 34 experts in this literature received the control mean and the same background and study design information as shown to you above. The experts on average predicted a treatment effect of 0.6 percentage points. The standard deviation of the expert forecasts was 7.6.

Publishability

If this study was submitted to the Journal of Development Economics, what do you think is the likelihood that the study would eventually be published there?

Very low likelihood								Very high likelihood	
0	10	20	30	40	50	60	70	80	90	100

Quality

On a scale from 0 to 100, where 0 indicates the "lowest possible quality" and 100 indicates the "highest possible quality," please indicate how **you** perceive the quality of this study.

Lowest possible quality								Highest possible quality	
0	10	20	30	40	50	60	70	80	90	100

Imagine that researchers in this field participated in an anonymous online survey and were asked to evaluate the quality of the study on the same 100-point scale as above (where 0 indicates the "lowest possible quality" and 100 indicates the "highest possible quality").

What quality rating would you expect **these researchers** to give to the study on average?

Lowest possible quality								Highest possible quality	
0	10	20	30	40	50	60	70	80	90	100

→

E.1.5   Financial literacy program

Financial literacy program

Background and study design: In 2019, a team of 3 PhD students from Ohio State University conducted an RCT in India. The purpose of the RCT was to examine whether access to a two-day financial literacy program affected savings among small business owners.

In the RCT, 780 small business owners were evenly randomized into a treatment group and a control group. Respondents randomly assigned to the treatment group were offered a two-day financial literacy program addressing personal and small business financial management and planning within five content areas: (i) Budgeting and record keeping, (ii) Savings, (iii) Debt management, (iv) Investment, (v) Money transfer.

All treated respondents completed the two-day program. After the two-day program, treated respondents had a 41.5 percent of a standard deviation higher financial literacy score.

Main result of the study: Treated respondents were 1.6 percentage points (standard error 3.8) more likely to have savings in their mobile money account compared to a control mean of 42.0 percent.

Expert prediction: 26 experts in this literature received the control mean and the same background and study design information as shown to you above. The experts on average predicted a treatment effect of 2.7 percentage points. The standard deviation of the expert forecasts was 5.8.

Publishability

If this study was submitted to the Review of Economics and Statistics, what do you think is the likelihood that the study would eventually be published there?

Very low likelihood									Very high likelihood	
0	10	20	30	40	50	60	70	80	90	100

Quality

On a scale from 0 to 100, where 0 indicates the "lowest possible quality" and 100 indicates the "highest possible quality," please indicate how **you** perceive the quality of this study.

Lowest possible quality								Highest possible quality		
0	10	20	30	40	50	60	70	80	90	100

Imagine that researchers in this field participated in an anonymous online survey and were asked to evaluate the quality of the study on the same 100-point scale as above (where 0 indicates the "lowest possible quality" and 100 indicates the "highest possible quality").

What quality rating would you expect **these researchers** to give to the study on average?

Lowest possible quality								Highest possible quality		
0	10	20	30	40	50	60	70	80	90	100

→

E.1.6   Salience of poverty and patience

Salience of poverty and patience

Background and study design: In 2021, a team of 2 PhD students from UC Berkeley conducted an experiment on an online survey platform. The purpose of the experiment was to examine whether financial anxieties increase people's inclination to make more impatient choices.

800 US respondents were evenly randomized into a treatment and control group. Respondents were asked to write a few sentences about how they would raise $5,000 (treatment group) or $50 (control group) to cover an unexpected expense. The main outcome of interest was whether respondents choose to receive $100 now or $110 in a week. The choices were implemented for 25% of respondents.

The treatment increased respondents' financial anxieties by 29.1 percent of a standard deviation.

Main result of the study: Treated respondents were 7.8 percentage points (standard error 3.5) more likely to choose money now compared to a control mean of 45.0 percent.

Publishability

If this study was submitted to the Proceedings of the National Academy of Sciences (PNAS), what do you think is the likelihood that the study would eventually be published there?

Very low likelihood									Very high likelihood	
0	10	20	30	40	50	60	70	80	90	100

Quality

On a scale from 0 to 100, where 0 indicates the "lowest possible quality" and 100 indicates the "highest possible quality," please indicate how **you** perceive the quality of this study.

Lowest possible quality									Highest possible quality	
0	10	20	30	40	50	60	70	80	90	100

Imagine that researchers in this field participated in an anonymous online survey and were asked to evaluate the quality of the study on the same 100-point scale as above (where 0 indicates the "lowest possible quality" and 100 indicates the "highest possible quality").

What quality rating would you expect **these researchers** to give to the study on average?

Lowest possible quality									Highest possible quality	
0	10	20	30	40	50	60	70	80	90	100

→

E.2 Mechanism experiment

The mechanism experiment was identical to the main experiment except that respondents were shown all five vignettes and that we asked about the precision of the study instead of its quality or importance. Since the wording of the vignettes was identical across the experiments, we only show screenshots of one of the vignettes for the mechanism experiment (the female empowerment program vignette).

E.2.1 Pre-treatment information

Introduction

We will now ask you about your views regarding **five** hypothetical studies. These studies are based on real studies whose details we modified for the purposes of this survey.

We will provide you with a short description of the study design and a summary of the main findings of each study.

→

E.2.2   Female empowerment program

Female empowerment program

Background and study design: In 2018, a team of 4 PhD students from the University of Pittsburgh conducted an RCT in Sierra Leone. The purpose of the RCT was to examine whether access to a female empowerment program increased women's labor supply.

In the RCT, 360 women were evenly randomized into a treatment group and a control group. Respondents in the treatment group were offered a female empowerment program, combining both psychosocial therapy and vocational skills training. The program was very intensive: participants attended meetings for up to 5 hours every day during a 12-month period.

Main result of the study: Treated respondents were 1.7 percentage points (p-value = 0.73) more likely to take up a job offer compared to a control mean of 37.0 percent.

Expert prediction: 34 experts in this literature received the control mean and the same background and study design information as shown to you above. The experts on average predicted a treatment effect of 0.6 percentage points. The standard deviation of the expert forecasts was 7.6.

Publishability

If this study was submitted to the Journal of Development Economics, what do you think is the likelihood that the study would eventually be published there?

| Very low likelihood | | | | | | | | Very high likelihood | |
| 0 | 10 | 20 | 30 | 40 | 50 | 60 | 70 | 80 | 90 | 100 |

Precision

How would you rate the statistical precision of the main result?

○ Very precisely estimated

○ Precisely estimated

○ Somewhat precisely estimated

○ Imprecisely estimated

○ Very imprecisely estimated

→