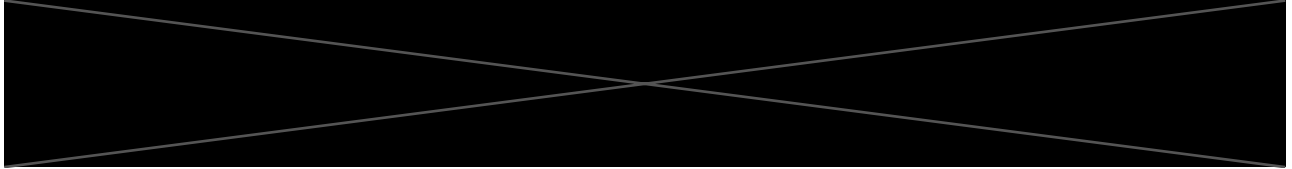


Fighting Poverty: International Attitudes Toward Poverty Reduction Programs



Date of latest draft:

October 1, 2026

Abstract

This study will provide the first large-scale, cross-country evidence on public support for different types of poverty reduction programs, including unconditional and conditional cash transfers, in-kind transfers, public works programs, asset transfers, and personalized tax refunds. Using a randomized survey experiment of 16,000 individuals across eight middle-income countries, home to 2.5 billion people, it examines the factors shaping this support. We also quantify the effects of educational videos on attitudes toward both the level of spending on poverty reduction programs and its allocation across program types. In particular, we examine whether information about the conditions faced by poor households and empirical evidence contradicting common misconceptions about poverty reduction programs increase support for cash transfers, which governments may not prioritize due to concerns about public opposition. We also explore whether these treatments can lower support for consumption tax exemptions, a ubiquitous but poorly targeted policy motivated on distributional grounds.

Keywords: social assistance programs, poverty, policy preferences, randomized experiments, tax policy.

JEL codes: C83, D72, D83, H53, I38.

Study pre-registration: AEARCTR-18825

Proposed timeline

No data collection occurred before our original submission to the JDE, except for small-scale piloting of the survey in Kenya and Nigeria in the first half of 2026 (and in December 2025). Due to a funding opportunity with a limited window for spending, we began data collection for the first phase of this study in Summer 2026, covering Brazil, India, Kenya, Nigeria, and South Africa. To preserve the pre-results nature of the Registered Report review, we will treat any data collected in this first phase as part of an extended pilot for the purposes of the JDE submission. We will commence the main data collection rounds after incorporating feedback from reviewers. At the time of writing, the remainder of the fieldwork is expected to take place in 2027, and we expect the study to be completed by March 2028.

Reporting checklist for Stage 1 submissions

Table 1: Reporting Checklist

Section	Item	Description and details to report	Reported?
Cover page (required)	Title	Informative title specifying the study design, population, and interventions	✓
	Date of latest draft	Date when the prospective review article was last edited	✓
	Study pre-registration status	Link, registration identifier and registry name (or intended registry if not yet registered)	✓
	Keywords	Up to six keywords for indexing	✓
	JEL codes	Up to six codes	✓
Abstract (required)	Abstract	Summarize research question, outcome variables, methodological framework and contribution in less than 150 words	✓
Timeline (required)	Expected completion date	Expected date for completion of the pre-specified research design	✓
Introduction	Background and relevance of the study	Overview of previous research and relevance of the research question(s) for economic development	✓
	Research question(s)		✓
Research design	Basic methodological framework	Outline of identification strategy (experimental/non-experimental)	✓
	Hypotheses	Pre-specified hypotheses to be tested and reported as primary findings	✓
	Outcome variable(s)	Definition of main outcome variables and secondary outcomes	✓
		Specification of how outcomes will be constructed from the dataset	✓
	Intervention(s)	Details of the intervention (when, where, how, by whom)	✓
		Number of treatment arms and whether they are exclusive or overlapping	✓
		Randomization strategy	✓
		Blinding strategy (if applicable)	✓

Section	Item	Description and details to report	Reported?
	Theory of change	Instructions and supporting materials for administering the intervention	✓
		Source(s) of exogenous variation	✓
		How and why the intervention is predicted to generate effects	✓
	Sample	Unit of analysis (individuals, organizations, countries, etc.)	✓
	Variations from intended sample	Data sources	✓
		Projected sample size and statistical power calculations	✓
		Degree of attrition that may threaten robustness	✓
	Data collection and processing	Strategies to deal with attrition and non-compliance	✓
		Data type, collection method, and timeline	✓
		Rule for terminating data collection / stopping rule	✓
		Data management plan	✓
		Pilot data and preparatory experiments	✓
Empirical analysis	Statistical method(s)	Main evaluation methods and assumptions	✓
		Rules for handling missing values	✓
		Definition and treatment of outliers	✓
	Multiple hypothesis testing	Strategies to prevent false positives	✓
	Heterogeneous effects	Anticipated heterogeneous effects and theoretical justification	✓
	Statistical model	Functional form of the causal model	✓
		Specification if regression model is linear, generalized linear, or other	✓
		Method for calculating standard errors	✓
Limitations and challenges	Implementation challenges	Potential circumstances that might jeopardize the study design	✓
Administrative information (required)	Ethics approval	Statement confirming ethics approvals	✓
	Funding	Funding sources	✓
	Acknowledgments	Individuals who contributed to the research	✓
Bibliography	Bibliography	References in any consistent format	✓
Other items	Appendices	Tables and figures	✓

1 Introduction

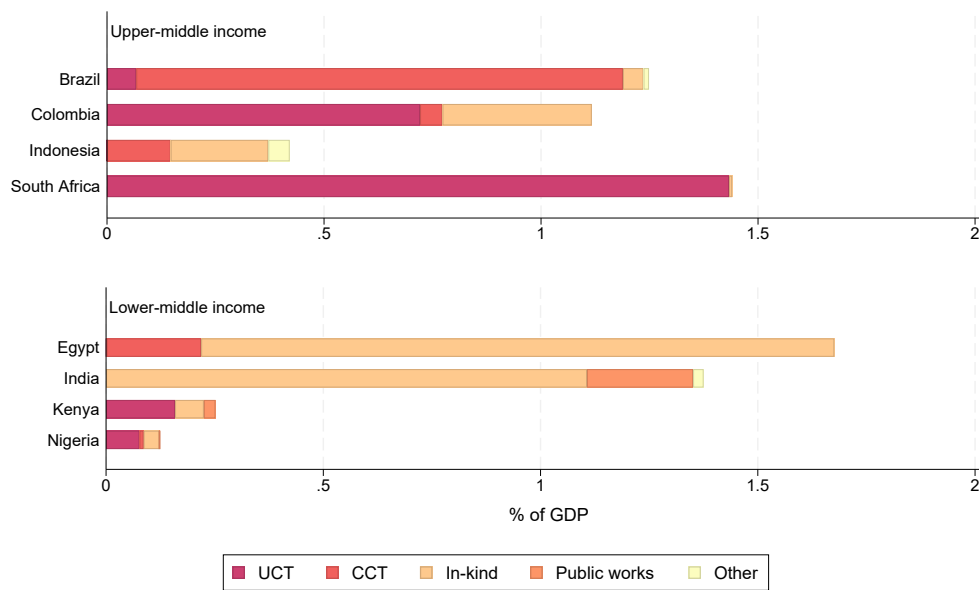
Progress in eliminating global poverty has slowed considerably in recent years, and this is projected to continue (World Bank, 2024), even though policymakers have an array of poverty reduction programs at their disposal, including cash transfers, in-kind transfers, and public works programs. A large body of literature has shown that many of these programs substantially reduce poverty and are cost-effective (Banerjee et al., 2024; Hanna and Olken, eds, 2026). Yet poverty reduction programs remain limited in many countries and are sometimes discontinued even when proven effective. A prominent example is *Prospera* in Mexico (formerly known as *Progres*a and *Oportunidades*), one of the most rigorously evaluated cash transfer programs in the world, which was discontinued despite overwhelming evidence of positive impacts (Parker and Todd, 2017). At the same time, total revenue foregone through reductions in consumption taxes on essential goods (“consumption tax exemptions” hereafter), which are generally motivated on distributional grounds, commonly exceeds spending on poverty reduction programs (Redonda et al., 2025), even though such exemptions are known to primarily benefit richer households (Inchauste and Lustig, 2017; Warwick et al., 2022; Bachas et al., 2024b).

One reason for this gap between evidence and policy may be that governments are constrained, or believe themselves to be constrained, by public preferences. Public works programs and in-kind transfers are often perceived as more publicly acceptable than cash transfers, since they impose ordeals or conditions that require beneficiary effort or restrict how benefits can be used. Similarly, policymakers may believe that universal consumption tax exemptions are more politically palatable than targeted social assistance programs. However, there is very little empirical evidence on the extent of public support for poverty reduction programs, how this varies across program types, and what drives these preferences.

We will fill this knowledge gap by conducting the first large-scale, harmonized survey experiment across eight middle-income countries on attitudes toward poverty reduction policies. These countries—Brazil, Colombia, Egypt, India, Indonesia, Kenya, Nigeria, and South Africa—are home to around 2.5 billion people, spread across Latin America, Sub-Saharan Africa, Asia, and the Middle East. The level of funding for poverty reduction programs in these countries varies from 0.1 to almost 2 percent of GDP, and interestingly, the relative composition of programs differs considerably (see Figure 1). The evidence we collect will shed light on how people reason about poverty reduction policies and will identify what drives support for one policy over another.

We elicit policy preferences across “traditional” social assistance program types (unconditional and conditional cash transfers, as well as in-kind transfers and public works programs) and two more novel program types: asset transfers and personalized tax refunds. Asset transfers are a central component of the “graduation approach”, which has been shown to be effective at reducing persistent poverty (Banerjee et al., 2015, 2021) but has seen limited government adoption. Personalized tax refunds have received increasing attention (de la Feria and Swistak, 2024; Kotlikoff et al., 2025), as they could provide a middle-ground between popular but poorly-targeted consumption tax exemptions and well-targeted but potentially less popular cash transfers. We also separately elicit respondents’ views on consumption tax exemptions.

Figure 1: Spending on social assistance programs as a share of GDP (selected countries)



Notes: Latest data for each country from the World Bank's ASPIRE Database, available at: <https://www.worldbank.org/en/data/datatopics/aspire>. "In-kind" includes social housing programs in Brazil. "Other" includes social care services in Brazil, a need-based scholarship program (PIP) and an indigenous empowerment program in Indonesia, and infrastructure and education schemes in India.

Our study is designed to answer the following research questions:

1. Primary questions

- 1.1 What types of poverty reduction programs do respondents support?
- 1.2 Which beliefs and characteristics best predict support for different types of programs?
- 1.3 Does information provision increase overall support for poverty reduction programs?
- 1.4 Does information provision change the types of programs respondents support?

2. Secondary questions

- 2.1 Do respondents support greater spending on the programs they prefer?
- 2.2 How does support for social assistance programs relate to the programs actually implemented across countries?
- 2.3 How do respondents perceive beneficiaries' relative willingness to pay for programs?
- 2.4 Which subgroups of respondents are the most likely to increase their support for poverty reduction programs in response to information provision?

2.5 Does respondents' preferred level of spending on poverty reduction programs depend on how it is funded?

We distinguish between three main factors that determine support for poverty reduction programs. The first is self-interest, which has been central to the seminal work on the political economy of redistribution, such as the median voter framework of Meltzer and Richard (1981), later extended by Benabou and Ok (2001) to account for (upward and downward) social mobility expectations. The second is a respondent's preferred role for government in reducing poverty, which incorporates beliefs about the depth of poverty and the deservingness of the poor. This relates to a theoretical literature on fairness views, ideology, and motivated reasoning (Alesina and Angeletos, 2005; Benabou and Tirole, 2006; Alesina and Giuliano, 2011). The third factor is the perceived effectiveness of different types of programs. This relates to the public finance literature on program design, which emphasizes that the form of transfers matters for both efficiency and political sustainability. For example, Nichols and Zeckhauser (1982) and Besley and Coate (1992) show that ordeals and conditionality can serve as screening devices to improve targeting, while Currie and Gahvari (2008) argue that in-kind transfers persist in part because voters are paternalistic about how the poor spend money. This broad framework guides the design of our survey and experimental treatments, allowing us to quantify the relative importance of self-interest, general attitudes toward government spending on poverty reduction, and beliefs about program effectiveness in shaping support. We show that our framework naturally maps to concepts in the welfare analysis of government policies, following Hendren and Sprung-Keyser (2020).

To better understand the causal links in people's reasoning, we will randomly vary respondents' exposure to two educational videos and quantify their impacts on attitudes toward both the level of spending on poverty reduction programs and its allocation across program types. In particular, directly motivated by our conceptual framework, we will evaluate the effects of providing accurate information about (i) the *conditions* poor households face and (ii) empirical evidence contradicting common misconceptions about the *effectiveness* of poverty reduction programs (e.g., the myth of the lazy welfare recipient).¹ This information is sourced from recent academic surveys, such as Banerjee et al. (2024) and Hanna and Olken, eds (2026), and the references contained therein. In the *effectiveness* video, we explicitly cite Evans and Popova (2017) showing that cash transfer beneficiaries do not increase "harmful" spending on drugs and gambling; Banerjee et al. (2017) showing that cash transfers do not reduce labor supply; and Warwick et al. (2022) and Bachas et al. (2024a) (building on Bachas et al., 2024b) showing that consumption tax exemptions primarily benefit richer households. The videos are brief, entertaining, simple, and tailored to the country context (similar to Dechezleprêtre et al., 2025). Collectively, the impact of these treatments speaks directly to what influences public support for poverty reduction programs and whether governments can use information to increase support for cost-effective programs that place fewer constraints on beneficiaries, such as unconditional cash transfers.

An important consideration is the representativeness of our sample, given that we plan to survey respondents online. Respondents are recruited from an online panel rather than randomly sampled, so we do not describe the sample as representative of a national population. As in other

¹Our small-scale pilot surveys confirm that these misconceptions are common in our study population (see Footnote 11).

large-scale online survey experiments that include middle-income countries (Dechezleprêtre et al., 2025; Hoy, 2025; Hoy et al., 2026), we expect it to be broadly representative of the internet-using population (at least in terms of age and gender), and to be younger, more educated, more urban, more likely to be male, and higher-income than the general population. The internet-using population in middle-income countries is large and growing rapidly, exceeding two-thirds of adults in most of the countries we will survey, and disproportionately includes voters and those who engage in political discourse (e.g., those using social media). Their preferences are therefore particularly relevant for understanding the political constraints governments face when designing poverty reduction programs. We will nonetheless report detailed comparisons between our sample demographics and nationally representative benchmarks for each country, allowing readers to assess the direction and magnitude of any selection. Our main results are reported for the achieved sample; results weighted toward the general population are reported in an appendix.

This study will make five main contributions to the literature on public attitudes toward poverty reduction programs. First, we provide the first large-scale, cross-country evidence on attitudes to poverty reduction programs in middle-income countries. Much of the existing literature focuses on support for redistribution in general, typically measured through questions about whether the government should reduce inequality (Alesina and Angeletos, 2005; Alesina and Giuliano, 2011; Kuziemko et al., 2015). Influential studies from middle-income countries in this vein, such as Cruces et al. (2013) in Argentina and Pellicer et al. (2019) in South Africa, have generated important insights but do not examine attitudes to poverty reduction programs specifically.² Yet governments face choices not only about how much to spend on poverty reduction but also about *how* to spend it, and a large body of evidence documents that the type of program likely matters enormously for political sustainability (Banerjee et al., 2024). This study will complement a growing body of work that uses large-scale, cross-country online surveys to study attitudes toward economic policies, including Stantcheva (2021) on tax policies, Dechezleprêtre et al. (2025) on climate policies, and Stantcheva (2022) on trade. We integrate the latest advances in methodology developed in this literature for eliciting perceptions, attitudes, and preferences (see Stantcheva, 2023) and conducting information experiments (see Haaland et al., 2023).

Second, our survey design allows us to quantify which beliefs can explain support for poverty reduction programs in general and for specific programs: self-interested motivations; general attitudes toward government spending on poverty reduction; and beliefs about program-specific effectiveness, including perceived waste and corruption, targeting accuracy, benefits for beneficiaries, and distortions and spillovers. By measuring all of these channels within the same instrument, we can assess their relative explanatory power and identify which beliefs are most strongly associated with differential support across program types, a question that prior work has been unable to address (Banerjee et al., 2017). Following Liscow and Pershing (2022), we measure respondents' preferred total spending on poverty reduction programs under different hypothetical allocations of spending to different programs, allowing us to characterize the potential trade-offs between program restrictiveness and overall spending on poverty reduction. We also capture respondents'

²De La O (2013) and Bursztyn (2016) study support for poverty reduction programs in middle-income countries but focus on *beneficiaries'* attitudes to cash transfers and how they translate into voting behavior. Liscow and Pershing (2022) study attitudes toward cash and in-kind transfers among different social groups in the United States, but attitudes in middle-income countries may be very different.

attitudes toward new types of poverty reduction programs, specifically asset transfers and personalized tax refunds, and compare them with traditional programs. This is particularly important in the context of the growing body of evidence showing that graduation programs, which rely on asset transfers, are highly effective at reducing poverty but have not been adopted by governments.

Third, our experimental design provides causal evidence on whether information provision shifts the level of support for social assistance and the types of programs that respondents prefer. The information experiment literature has examined the effects of correcting misperceptions about poverty and inequality (Cruces et al., 2013; Kuziemko et al., 2015; Hoy and Mager, 2021), but these studies focus almost exclusively on general redistributive preferences. A key policy question—whether evidence on program effectiveness can shift overall public support or support for specific programs—has not been experimentally addressed. Our video treatments, which separately present information about the effectiveness of different programs and the conditions the poor face, allow us to test whether preferences are impacted more through an “effectiveness” channel (shifting beliefs about how well programs work) or through a “conditions” channel (shifting perceptions of the conditions the poor face), and whether these channels differentially affect support for specific program types.

Fourth, we make a methodological contribution by measuring respondents’ perceptions of beneficiaries’ relative willingness to pay for different programs, which would be a central element of any formal welfare analysis of spending on different programs (Hendren and Sprung-Keyser, 2020). Using a series of pairwise comparisons, we elicit respondents’ beliefs about how beneficiaries would trade off between program types, and the level of additional financial support required to make them indifferent between an unconditional cash transfer and an alternative program. This approach allows us to measure the perceived monetary value of the non-monetary program features (conditionality, ordeals, in-kind provision). As a result, we are able to disentangle whether respondents support more restrictive programs despite the costs they impose on beneficiaries or because they believe these features are actually valued by beneficiaries themselves.

Finally, our design enables us to examine the relationship between each country’s existing social assistance policies and public attitudes toward them. A strand of the literature has documented that redistributive preferences are shaped by the institutional and policy environment in which individuals live. For example, Luttmer and Singhal (2011) find that migrants’ distributive preferences are correlated with average preferences in their birth countries, suggesting a role for cultural transmission, and Almås et al. (2020) demonstrate that cross-country differences in fairness norms mirror differences in social assistance policies. By surveying respondents across countries with widely varying levels and compositions of spending on poverty reduction programs, we can examine whether people’s preferences are anchored to the status quo in their country, and how the relationship between beliefs and support varies across policy environments.

The remainder of this paper is organized as follows. Section 2 presents our conceptual framework and hypotheses. Section 3 describes the research design, including the survey instrument, the video interventions, and power calculations. Section 4 presents our empirical strategy, Section 5 presents how we plan to interpret the results, and Section 6 discusses the study’s limitations and potential challenges. The Appendix includes a theoretical result on the ambiguous relationship

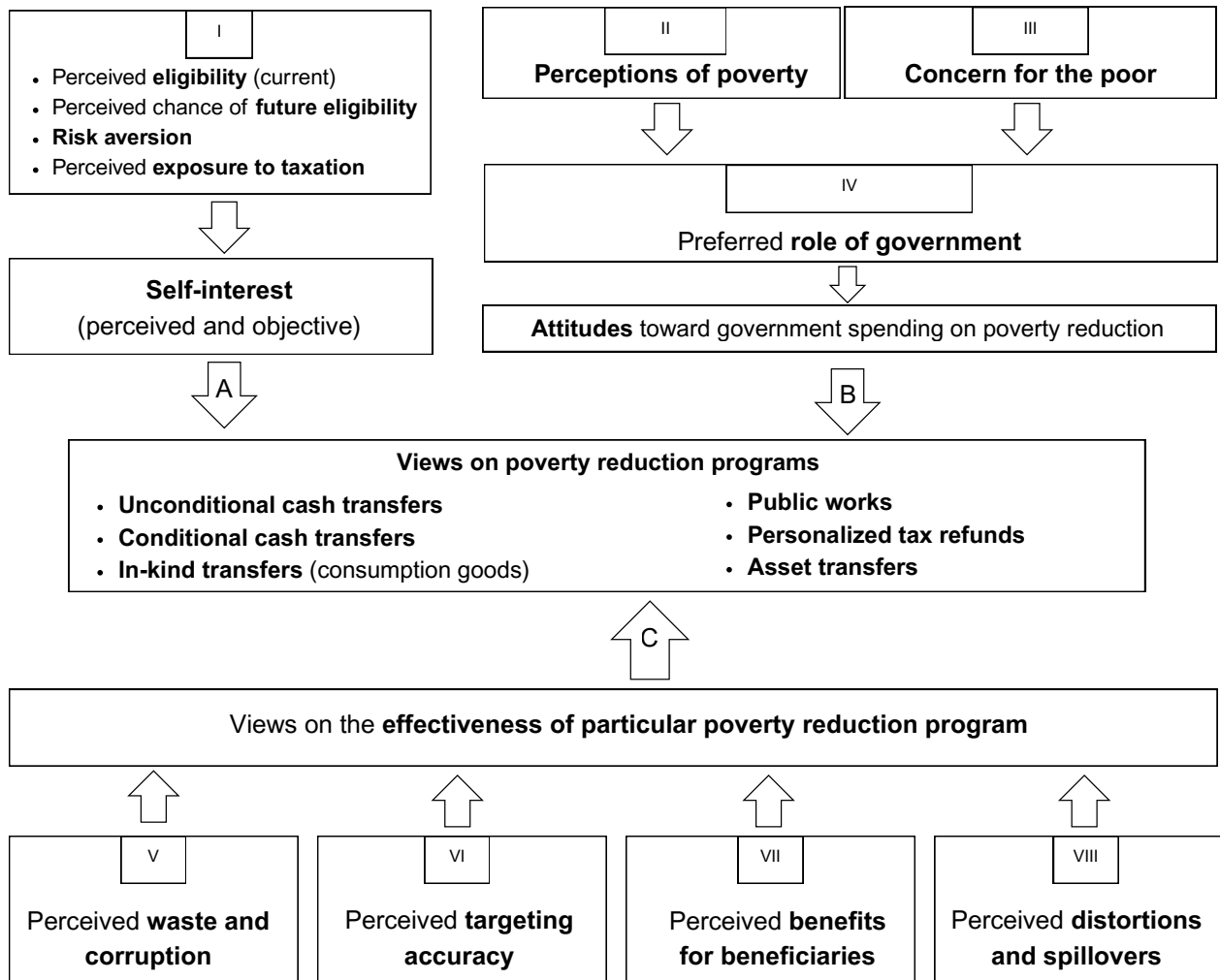
between allocation efficiency and overall support for poverty reduction programs, key findings from our small-scale pilot surveys, and the full survey instrument.

2 Conceptual Framework and Hypotheses

2.1 Conceptual framework

Figure 2 is a graphical representation of our conceptual framework. We use this framework to design the survey, considering each of the factors that may contribute to an individual's support for different types of poverty reduction programs. The figure is an organizing device for the survey rather than a causal model that we test. The boxes are not assumed to be causally independent. The arrows indicate the channels the survey is built to measure; the descriptive analysis in Section 4.2 documents how the measured beliefs co-vary with support. The experimental treatments are the only source of causal variation in the design.

Figure 2: Conceptual framework: factors shaping views on poverty reduction



Notes: The figure summarizes the conceptual framework for this study. Boxes I–IV capture general determinants of support for poverty reduction programs: self-interest, perceptions of poverty, concern for the poor, and views about the appropriate role of government. Boxes V–VIII capture program-specific beliefs about effectiveness: perceived waste and corruption, targeting accuracy, benefits for beneficiaries, and distortions and spillovers.

Box I outlines key factors that could motivate an individual's self-interested support for poverty reduction policies. Most straightforwardly, respondents will usually find it in their self-interest to support a program if they—or members of their family—are currently eligible for its benefits or expect to become eligible in the future. The higher the chance of becoming eligible in the future and the higher an individual's level of risk-aversion, the more likely they are to have a self-interested motivation for supporting a poverty reduction program. Conversely, respondents who perceive themselves as likely to be taxed more to fund a poverty reduction program will have a self-interested motivation to oppose it.³ We represent the impact of self-interest on attitudes toward poverty reduction programs as Arrow A.

Boxes II-III categorize general attitudes toward poverty that may generate support for poverty reduction programs, even among people without a self-interested motivation. Perceptions of the prevalence and the severity of poverty in a society (Box II) and the degree of concern for poor people (Box III) will influence the extent to which poverty is seen as a problem that needs to be addressed. Respondents may have various reasons for concern about the poor, including normative beliefs about universal entitlements (e.g., to food, housing and healthcare). Respondents' level of concern for the poor may also depend on their views about the causes of poverty: those who believe it arises from bad luck or systemic disadvantage may be more likely to support poverty reduction programs than those who perceive it as a result of low effort or self-discipline.⁴

Views on the appropriate role of government in reducing poverty (Box IV) will be influenced by people's perceptions of poverty and their level of concern for the poor, and will determine how those views translate into support for poverty reduction programs. For instance, some respondents might not believe it is the responsibility of the government to support the poor, or at least insofar as they are able-bodied adults without children. Or they may agree that poverty should be reduced but do not believe the government is capable of doing anything about it. The effect of the resulting general attitudes toward poverty reduction programs is represented by Arrow B.

Boxes V-VIII capture views on the effectiveness of particular poverty reduction policies. In contrast to most of the views in Boxes I-IV, these are in essence views on empirical questions (i.e., matters of fact). Box V captures the extent to which respondents expect a given program to be subject to corruption and waste. Respondents might be concerned that funds for poverty reduction programs would be misappropriated through corruption and embezzlement. Even if funds were used for legitimate purposes, respondents might worry that a large share of the programs' costs is spent on government administration (Banerjee et al., 2024).

Box VI captures concerns about targeting: given that poverty is not easily measured by governments in low- and middle-income countries, eligibility for poverty reduction programs can only be determined through some combination of proxy means tests and *ordeals*—barriers to accessing poverty reduction programs that encourage self-selection (Banerjee et al., 2024). Respondents may not support policies if they believe the targeting mechanism is weak, leading to a non-trivial share

³This list is not exhaustive. Other factors that might affect self-interested motivation to support a poverty reduction program include someone's perceived chance of moving out of poverty if they are poor and their discount rate.

⁴Andersen et al. (2023) find that individuals who receive an exogenous wealth shock become *less* likely to attribute poverty to luck, indicating that beliefs about the reasons for poverty are themselves shaped by self-interest. However, in a different context Cerkez et al. (2025) find no effect, suggesting that this result may be context-specific.

of program spending going to families that are not poor.

Box VII captures the perceived benefits of poverty reduction programs for the poor. Beliefs about the effectiveness of these programs at reducing poverty and/or providing some form of insurance for poor people are likely to influence support. However, respondents might be concerned that the value of poverty reduction programs to beneficiaries could be diminished by costs such as *ordeals*, conditionality (e.g., work requirements), or the stigma attached to receiving assistance, and some might worry that the poor would not use assistance in their own best interest.

Box VIII contains the perceived costs and benefits of social assistance programs for wider society. On the cost side, poverty reduction programs might reduce labor supply or raise prices, negatively impacting non-poor households. Perceived benefits include stimulating the local economy or raising wages (Banerjee et al., 2017). Taken together, the effect of these views about the effectiveness of particular poverty reduction programs on attitudes toward them is represented by Arrow C.

These factors map onto familiar concepts in the welfare analysis of government policies. The exposition here—following the framework of Hendren and Sprung-Keyser (2020)—is not intended to be structural but (as e.g. in Stantcheva, 2021) to help frame the survey design and guide interpretation of the empirical findings. Suppose respondents form their views as if maximizing the social welfare function W , which is a weighted sum of individual utilities $W = U_{i'} + \sum_{i \neq i'} \psi_i U_i$, where $U_{i'}$ is a respondent's own utility and ψ_i is the subjective social welfare weight for individual i (the welfare weight on the respondent's own utility is normalized to $\psi_{i'} = 1$). Then the change in perceived social welfare from a marginal increase in spending on policy j is

$$\frac{dW}{dp_j} = \frac{dU_{i'}}{dp_j} + \sum_{i \neq i'} \psi_i \frac{dU_i}{dp_j} = \lambda_{i'} WTP_{i'}^j + \sum_{i \neq i'} \psi_i \lambda_i WTP_i^j \quad (1)$$

where λ_i is individual i 's marginal utility of income and WTP_i^j is i 's willingness to pay for a marginal increase in spending on program j (as perceived by the respondent). WTP_i^j can be decomposed further; one useful decomposition is $WTP_i^j = \alpha_i^j \delta_i^j + \epsilon_i^j$, where α_i^j is the share of spending on j that is of direct benefit to i , δ_i^j is i 's willingness to pay for this direct benefit, and ϵ_i^j captures the indirect costs or benefits (in cash terms) that i experiences as a result of higher spending on j . For an unconditional cash transfer, $\delta_i^j = 1$ by definition; by contrast, we would have $\delta_i^j < 1$ for in-kind transfers if respondents believe that beneficiaries value them below their monetary value.

The total effect on perceived social welfare per dollar of net government expenditure on social assistance policy j is given by

$$\frac{\frac{dW}{dp_j}}{\frac{dR}{dp_j}} = \frac{\lambda_{i'} (\alpha_{i'}^j \delta_{i'}^j + \epsilon_{i'}^j) + \sum_{i \neq i'} \psi_i \lambda_i (\alpha_i^j \delta_i^j + \epsilon_i^j)}{1 + FE_j}, \quad (2)$$

which additionally takes into account fiscal externalities FE_j ; $FE_j < 0$ means that there are positive spillovers to the government budget, reducing the effective cost of the policy. Then the extent to which respondents are concerned about the poor (Box III) relative to themselves (Box I) maps onto others' relative welfare weights ψ_i . Others' perceived marginal utility of income λ_i relates to the respondent's perceptions of poverty (Box II). The perceived waste and corruption (Box V) and the

perceived targeting accuracy (Box VI) of a policy j affect the absolute value and the distribution of the α_i^j s. Beneficiaries' perceived willingness to pay for the direct benefits of poverty reduction programs is reflected in δ_i^j (Box VII).⁵ Distortions and spillovers (Box VIII) feed into FE_j if they go through the government budget and into the ϵ_i^j s (including those of the non-poor) otherwise.⁶

2.2 Hypotheses

We outline below our primary and secondary (both descriptive and experimental) hypotheses.

2.2.1 Primary hypotheses

Correlates of baseline support (descriptive analysis)

H1a: Overall support for poverty reduction programs will be higher among respondents who currently benefit, or expect to benefit in the future, from such programs.

This relates to Box I in our conceptual framework and our research question 1.2.

H1b: Overall support for poverty reduction programs will be higher among respondents who perceive poverty to be more severe.

This relates to Box II in our conceptual framework and our research question 1.2.

H1c: Support for each type of poverty reduction program will be higher among respondents who believe that type of program to be more effective.

This relates to Boxes V–VIII in our conceptual framework and our research question 1.2.

Experimental outcomes

H2a: A video providing respondents with accurate information about the conditions poor households face will increase their overall support for poverty reduction programs.⁷

This relates to Box II and Arrow B in our conceptual framework and our research question 1.3.

H2b: A video providing respondents with accurate information about the effectiveness of poverty reduction programs will increase their overall support for poverty reduction programs.

This relates to Boxes V–VIII and Arrow C in our conceptual framework and our research question 1.3.

H2c: A video providing respondents with accurate information about the effectiveness of poverty

⁵Paternalistic concerns could be captured in this framework as δ_i^j for $i \neq i'$ differing from the respondent's best judgment of what beneficiaries would actually be willing to pay for the direct benefit of the policy.

⁶Hendren and Sprung-Keyser (2020) emphasize a variant of equation (2), namely $\frac{dW}{dp_j} / \frac{dR}{dp_j} = \bar{\eta}_j MVPF_j$, where $\bar{\eta}_j = \sum_i \frac{WTP_i^j}{\sum_i WTP_i^j} \psi_i \lambda_i$ is the weighted average of the product of individuals' marginal utilities and welfare weights (weighted by their willingness to pay) and $MVPF_j = \frac{\sum_i WTP_i^j}{1+FE_j}$ is policy j 's Marginal Value of Public Funds (MVPF). $MVPF_j$ has an intuitive interpretation as policy j 's value for money.

⁷A decrease in support is also theoretically possible if respondents interpret the information as evidence that poverty is intractable or that assistance is unlikely to be effective.

reduction programs will increase their relative support for cash transfers.

This relates to Boxes V–VIII and Arrow C in our conceptual framework and our research question 1.4.

H2d: A video providing respondents with accurate information about the effectiveness of poverty reduction programs will decrease their support for consumption tax exemptions as an alternative to targeted social assistance.

This relates to Boxes V–VIII and Arrow C in our conceptual framework and our research question 1.4.

2.2.2 Secondary hypotheses

Additional descriptive hypotheses

H3a: Respondents will support higher poverty reduction spending when the allocation of resources across program types is closer to their preferred allocation.⁸

This relates to our research question 2.1.

H3b: Respondents will express greater support for poverty reduction programs that are more closely aligned with the programs currently implemented in their country.

This relates to our research question 2.2.

H3c: Respondents will believe that beneficiaries would need greater compensation to prefer programs with conditions/ordeals.

This relates to our research question 2.3.

Heterogeneous treatment effects

H4a: A video providing respondents with accurate information about the conditions poor households face will *increase* support for poverty reduction programs more among respondents who, prior to the treatment, perceived poverty as less severe than among those who perceived it as more severe.

This relates to Box II in our conceptual framework and our research question 2.4.

H4b: Providing respondents with either video treatment will *increase* support for poverty reduction programs more among respondents who, prior to treatment, believed that the government has the responsibility to reduce poverty than among those who did not.

This relates to Box IV in our conceptual framework and our research question 2.4.

H4c: A video providing respondents with accurate information about the effectiveness of poverty reduction programs will *increase* support for these programs more among respondents who, prior to the treatment, did not believe these programs to be effective than among those who did.

⁸As shown in Appendix A, this may or may not be the case for a rational social planner depending on the curvature of household utility functions.

This relates to Boxes V-VIII and Arrow C in our conceptual framework and our research question 2.4.

H4d: A video providing respondents with accurate information about the effectiveness of poverty reduction programs will *decrease* support for consumption tax exemptions more among respondents who believe they are personally less affected by changes to consumption taxes on essential goods than among those who believe they are more affected.

This relates to Box I in our conceptual framework and our research question 2.4.

Financing of poverty reduction spending

H5a: Respondents will give a higher preferred level of overall spending on poverty reduction programs when told that any additional spending would be funded by foreign aid than when told it would be funded by higher taxes or by lower government spending on other priorities.

This relates to Box I in our conceptual framework and our research question 2.5.

3 Research Design

3.1 Respondent sample

We will conduct an online randomized survey experiment with respondents from eight middle-income countries: Brazil, Colombia, Egypt, India, Indonesia, Kenya, Nigeria, and South Africa. We selected these countries because they have large online populations, have been successfully included in other cross-country randomized survey experiments, and have implemented a variety of different social assistance programs. Respondents will be recruited through a global survey firm, YouGov, which regularly runs online surveys with a pre-registered pool of people in these countries who already have experience taking part in surveys on different topics for compensation, such as phone credit. YouGov ensures that respondents have a track record of being attentive during surveys, excluding those who regularly rush through surveys or do not complete them in full. Recruitment is not subject to demographic quotas; we only set an overall target for the number of completed responses in each country.

Our target sample size is 2,000 respondents per country, yielding a total sample of 16,000 respondents. We expect the sample to be broadly representative of the internet-using adult population (at least in terms of age and gender), as has been documented for another survey conducted through the same YouGov panels (Hoy et al., 2026). We will report the age and gender composition of our sample compared to that of the internet-using and the adult populations in each country. Because respondents are drawn from a pool of people who take surveys for compensation, they are likely to be more literate and more practiced at answering survey questions than internet users in general; relative to the adult population, the sample is expected to be younger, more educated, more urban, more likely to be male, and higher-income.

Our main results are reported for the achieved sample and are therefore sample average treatment effects. As a robustness exercise, we also report results weighted to national census margins (Section 4.8.3). These target the adult population as a whole. Recovering a population average

treatment effect from them requires an ignorability assumption: that treatment effects among sampled respondents match those in the population, conditional on the weighting variables.

Table 2 provides the share of the population using the internet by age, gender and location in each country. These figures substantially understate the internet coverage of the *adult* population as children are included in the figures. This distinction is particularly important for the lower-income settings in this study, such as Kenya (37% of the population aged 0-14) and Nigeria (41% of the population aged 0-14).

Table 2: Online share of population and subgroups

	% of Population	% of Men	% of 15–24 Year Olds	% of Urban Population
Brazil	84.5	85.2	95.6	85.7
Colombia	79.3	78.6	92.3	84.6
Egypt	74.6	79.3	92.7	83.6
India	70.0	76.7	84.7*	79.9
Indonesia	72.8	75.7	95.5	79.1
Kenya	35.0	37.8	46.6	56.6
Nigeria	41.2	44.6*	49.9*	48.8*
South Africa	78.4	80.7*	94.9*	83.5*

Source: International Telecommunications Union. Available at: <https://datahub.itu.int/> (Accessed: 20 May 2026).

Note: * indicates figures are estimated based on regional averages.

3.2 Why an online sample is informative in this setting

A natural concern with our design is that an online sample is not representative of the full adult population in each country. We therefore interpret our main estimates as evidence on the attitudes and information responsiveness of internet-using adults, rather than as a direct estimate of population-wide preferences. Internet-using adults are a politically relevant population in their own right. In middle-income countries, public debate about social assistance, taxation, and government spending increasingly takes place through digital media, social media, and other online channels. Individuals with internet access are more likely to be exposed to government communication campaigns, advocacy campaigns, media coverage, and political messaging about poverty reduction programs. As a result, their views are especially relevant for understanding the public-opinion constraints that governments face when considering reforms to social assistance policy. This is the same logic underlying recent large-scale online survey experiments on policy attitudes in middle-income countries, including Hoy et al. (2026), who study attitudes toward fossil fuel subsidy reform using an online survey experiment across twelve middle-income countries.

Online implementation enables the fielding of a complex, standardized survey instrument across multiple countries within a short period of time. The survey includes video treatments, knowledge checks, budget-allocation exercises, and randomized program comparisons, all of which would be substantially more difficult, costly, and less comparable to implement through in-person surveys across eight countries. Online implementation allows us to deliver identical information treatments, randomize multiple survey components, monitor attention, and collect a large sample within a comparable fieldwork window. This improves the comparability of treatment ef-

fects across countries and enables a design that would be difficult to replicate using traditional household-survey methods. Online administration also helps mitigate experimenter effects and social desirability bias relative to more direct modes of interviewing (Kreuter et al., 2008; Haaland et al., 2023). Recent work in economics has used similar designs to study policy preferences and information effects in large online samples (Dechezleprêtre et al., 2025; Kuziemko et al., 2015; Alesina et al., 2018).

We nevertheless take the limits of online sampling seriously. In the main analysis, we will describe the demographic composition of the sample in each country and compare it with available nationally representative benchmarks. The main results are reported for the unweighted online sample. Survey-weighted results, adjusting for observable differences between the online sample and the adult population at a minimum by age and gender, will be reported in an appendix. We will also examine whether baseline support and treatment effects vary by observable characteristics that are correlated with internet access, including age, gender, education, income, and urban residence. These analyses will help assess whether the main findings are concentrated among particular subgroups or are broadly similar across the observable characteristics of the sample.

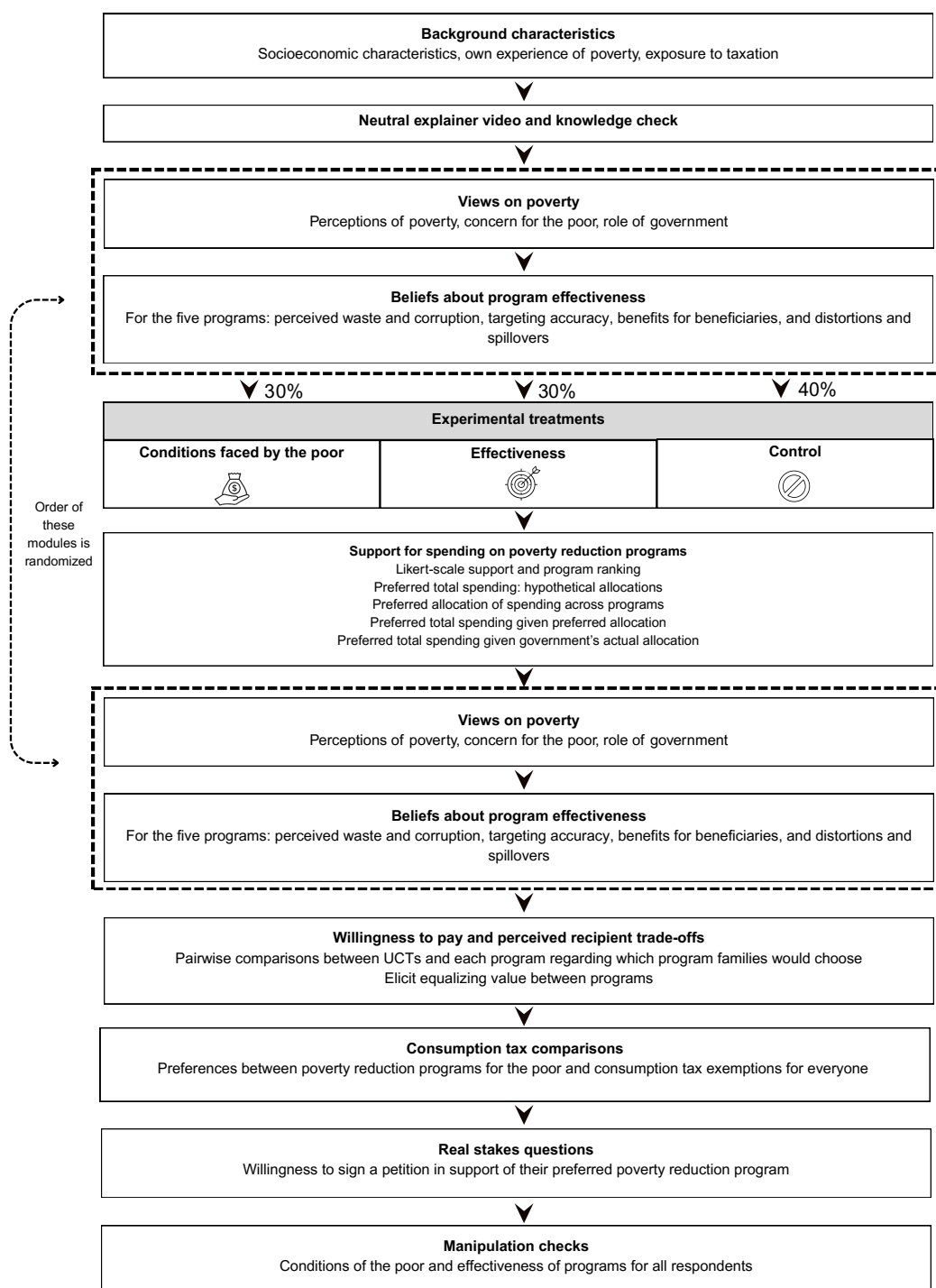
3.3 Survey instrument

The questionnaire is standardized across countries while allowing local adaptation of language, currency amounts and country-specific information, as detailed below. Surveys in India, Kenya, Nigeria, and South Africa are administered in English. Surveys in the remaining four countries—Brazil, Colombia, Egypt, Indonesia—are translated from English into the local language and reviewed by multiple native speakers. Currency amounts and other country-specific items—such as reference spending levels and program names—are adapted for each country as described further below. Video treatments are adapted for local context where appropriate (e.g., using country-specific imagery) while keeping the substantive content identical across countries. Video voiceovers are recorded by professional voice actors from each country, including native speakers of Indian, Kenyan, Nigerian, and South African English; all voices are female for consistency. All videos will be made available online. To monitor country-specific comprehension issues, we will report knowledge-check pass rates, attention-check performance, and completion times separately by survey language in the Stage 2 report.

As shown in Figure 3, the instrument is organized into ten major sections. As indicated by the dashed lines, half of respondents will be randomly allocated to completing the **Views on poverty** and **Beliefs about program effectiveness** sections after the main treatments and questions about support for different types of poverty reduction programs. This means the survey can be used to analyze both first-stage effects of the main treatments on views of effectiveness (for those who complete effectiveness questions after the treatment) and heterogeneous treatment effects based on prior beliefs about program effectiveness (for those who complete effectiveness questions before the treatment). Section 3.8 explains in more detail why we randomize the placement of these two modules and what this implies for statistical power.

The content of the sections is:

Figure 3: Survey outline



Notes: The figure summarizes the sequence of survey modules. All respondents first complete background questions, questions on views about poverty, and a neutral information video with a knowledge check. Respondents are then randomly assigned to the Conditions faced by the Poor video, the Effectiveness video, or the control group, with assignment probabilities of 30 percent, 30 percent, and 40 percent, respectively. The views on poverty and effectiveness modules are randomly placed either before or after the treatment stage: when asked before treatment, these measures are used to capture baseline beliefs and to define pre-specified heterogeneity; when asked after treatment, they are used to measure updated beliefs. After treatment, respondents complete the main support, willingness-to-pay, consumption-tax, and real-stakes modules. The figure abstracts from additional within-module randomizations, including the randomly assigned fifth program type and the consumption-tax comparator.

1. Background characteristics. We first collect demographic and socioeconomic information, including age, gender, income, area of residence, household composition, education, and employment status, as well as the respondent’s group membership along the social cleavage we expect to be most salient for attitudes toward redistribution in their country. This is ethnicity in Brazil, Colombia, Kenya and South Africa, and religion in Egypt, India, Indonesia and Nigeria. We measure one cleavage per country, in order to keep the instrument to a manageable length. We also elicit views on the current government, risk preferences, altruism, knowledge of existing poverty reduction programs in a given country, and perceived exposure to a consumption tax. Respondents’ answers to these questions are primarily used for heterogeneity analysis. We also ask respondents about their current subjective poverty and receipt of benefits from existing poverty-reduction programs, as well as their expectations of being poor or benefiting from these programs in the future; these questions directly relate to the self-interest component (Box I) of the conceptual framework.

Table 3: Treatment and explainer video links, by country

Country	Links to Videos
Kenya	Explainer 1, Explainer 2, Conditions faced by the Poor, Effectiveness
Nigeria	Explainer 1, Explainer 2, Conditions faced by the Poor, Effectiveness
Brazil	Explainer 1, Explainer 2, Conditions faced by the Poor, Effectiveness
South Africa	Explainer 1, Explainer 2, Conditions faced by the Poor, Effectiveness
India	Explainer 1, Explainer 2, Conditions faced by the Poor, Effectiveness

Notes: Each link points to the country-specific version of the video. *Explainer 1* and *Explainer 2* are the two neutral program-explainer videos shown to all respondents before treatment; *Conditions faced by the Poor* and *Effectiveness* are the two treatment videos. Videos are adapted to local context while keeping the substantive content identical across countries.

2. Neutral explainer video and knowledge check. All respondents will then watch a two-minute neutral explainer video on four “core” types of poverty reduction programs—unconditional cash transfers, conditional cash transfers, in-kind transfers, and public works programs—and one randomly assigned “new” type of program, either asset transfer or personalized tax refund. The information presented in the videos is shown in Table 4; Figure 4 shows the video storyboard (Kenya version with “asset transfer” as the fifth program). Links to the full explainer videos for Kenya, Nigeria, Brazil, South Africa, and India are given in Table 3 (both versions).

This video is intended to provide all respondents with a common baseline understanding of the policy instruments and does not present information about program effectiveness. After the video, respondents will complete a ten-point knowledge check quiz (including knowledge of the fifth program). Those scoring below 7 out of 10 are shown the correct answers and are asked to retake the quiz. Respondents who fail a second time are excluded from the analysis sample. As this exclusion occurs for all respondents and before treatment assignment, it does not impact the internal validity of our experiment. As a robustness check, we will also report results using alternative exclusion thresholds of 6 out of 10 and 8 out of 10.

Table 4: Program explainer video: content

Program	Receives	Requires	Regularity
UCT	Money to spend however they like	Nothing	Regular (monthly)
CCT	Money to spend however they like	Send their kids to school or get them vaccinated	Regular (monthly)
In-kind	Everyday essential goods, such as food	Nothing	Regular (monthly)
Public works	Money to spend however they like	Work on government projects	Only when they work
Asset transfer	An income generating asset, such as a cow	Nothing	One-off
Tax refund	A refund of the tax they have paid on essential goods	Nothing	Whenever they buy eligible goods

Notes: The table summarizes the information presented in the neutral program explainer video shown to all respondents before treatment assignment. The video introduces the basic features of each program type, including what beneficiaries receive, whether any conditions or requirements apply, and whether support is regular or one-off. All respondents are shown the four core program types: unconditional cash transfers, conditional cash transfers, in-kind transfers, and public works programs. Respondents are also randomly assigned to see one additional program type, either asset transfers or personalized tax refunds. The explainer video is intended to create a common baseline understanding of the policy instruments and does not present evidence on program effectiveness.

3. Views on poverty. Respondents are then asked questions corresponding to Boxes II–IV of the conceptual framework. These questions cover perceived poverty severity and its causes, beliefs about the possibility of escaping poverty, and views on the role of government. *Half of respondents will be randomly allocated to completing this section later (see Figure 3).*

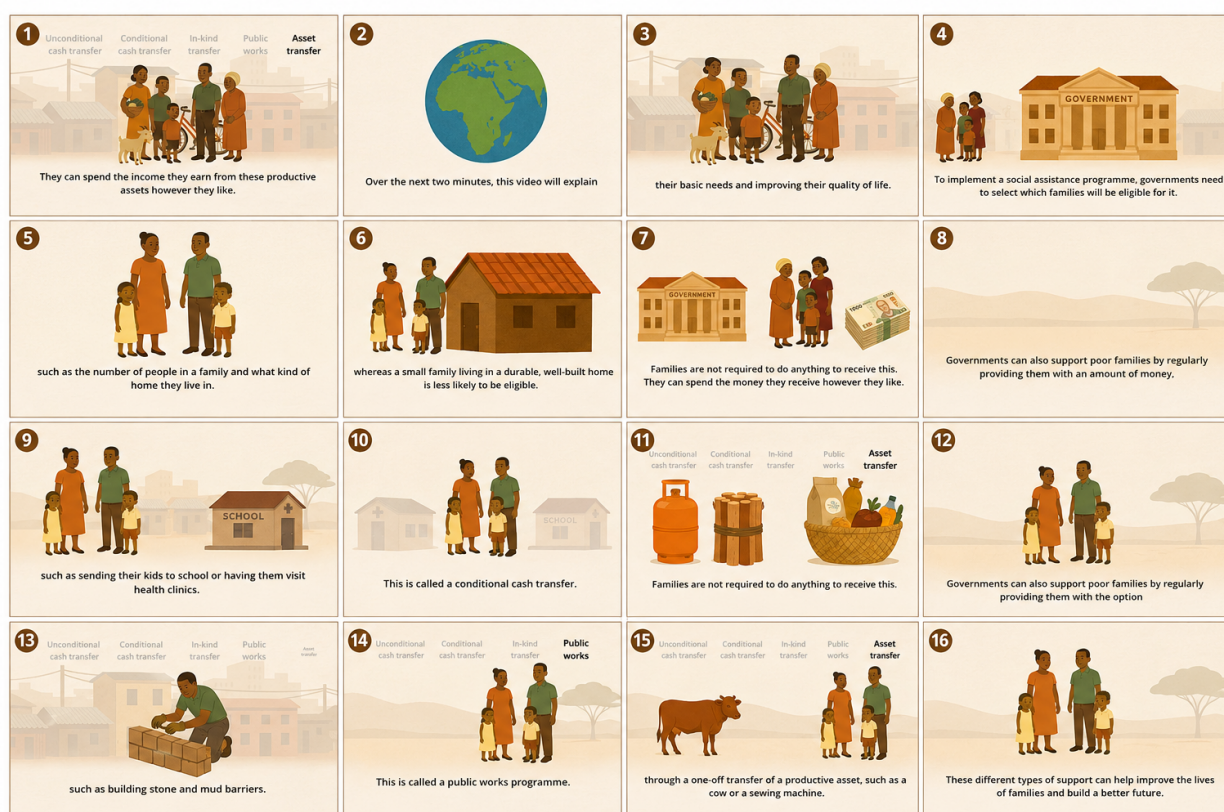
4. Beliefs about program effectiveness. Respondents then complete modules measuring beliefs corresponding to Boxes V–VIII of the conceptual framework. This includes a “leaky bucket” exercise, where respondents are asked what shares of spending on social assistance they believe goes to intended beneficiaries, unintended beneficiaries, the cost of administration, and corruption. Respondents also choose whether they think various statements about program effectiveness apply to the four “core” programs and the additional “new” program they were shown in the explainer video. *Half of respondents will be randomly allocated to completing this section later (see Figure 3).*

5. Experimental treatments. Respondents in the treatment groups receive one of two video treatments. Details are provided in Section 3.5.

6. Support for spending on poverty reduction programs. We first elicit support for social assistance using five-point Likert-scale questions on general support for poverty reduction programs and on each of the four core program types. These questions are followed by a question asking respondents to rank program types in order of preference.

This is followed by a series of budget-allocation tasks. Respondents state their preferred level of overall spending on poverty reduction programs under different hypothetical budget allocations across program types. These are

Figure 4: Program explainer video: storyboard



- an equal split across program types,
- all spending allocated to a single program type,
- the respondent's own preferred allocation, and
- the actual allocation in their country as sourced from the World Bank's Atlas of Social Protection Indicators of Resilience and Equity (ASPIRE).⁹

We elicit the respondent's own preferred allocation before telling them about the actual allocation in their country. We deliberately anchor the preferred level of overall spending by informing respondents of a typical level of spending in countries at a similar income level, expressed as spending per family member per month in local currency terms; this is to allow respondents to express a genuine preference even though they might have no prior knowledge of what would be a realistic amount.¹⁰

We explicitly highlight the government budget constraint. Before stating a preferred level, respondents are asked to assume that additional spending will be paid for by (i) foreign donor countries, (ii) higher taxes for the general population, or (iii) lower government spending on other things.

⁹The ASPIRE database is freely available at: <https://www.worldbank.org/en/data/datatopics/aspire>.

¹⁰We calculate typical monthly spending per beneficiary in countries at a similar income level as 5% of GDP per capita per month. This is based on typical spending on poverty reduction programs in middle-income countries of around 0.5% of GDP and these programs typically covering around 10% of the population (approximate figures based on the World Bank's ASPIRE database).

Which constraint is presented is randomized across respondents (see Section 3.4).

Finally, respondents are asked to repeat the ranking exercise when the “new” program type they have been told about in the explainer video—either asset transfer or personalized tax refund—is added to the choice set.

7. Willingness to pay and perceived recipient trade-offs. Respondents complete questions on a set of pairwise vignettes that elicit beliefs about poor households’ program preferences. For each of three pairs—CCT versus UCT, in-kind transfer versus UCT, and public works versus UCT—respondents are first asked which program they believe most poor families would choose, or whether families would be evenly split between the two. UCT serves as the common benchmark across all three comparisons. For respondents who identify one program as preferred, a follow-up question elicits an equalizing-value measure: how much would the transfer amount of the less-preferred program need to increase—holding the preferred program fixed at a reference transfer value—for poor families to be indifferent between the two? Responses are recorded on a discrete ordered scale, with an open-ended option for larger values. Because the vignettes ask what poor families would choose rather than what respondents themselves prefer, these items capture perceived recipient welfare, which is conceptually distinct from respondents’ own support for programs. The questions in this section complement qualitative information on perceived benefits for beneficiaries (Box VII), elicited in the previous section, with a quantitative measure.

8. Consumption tax comparisons. In the final section, respondents are asked about consumption tax exemptions as an alternative to targeted social assistance. We elicit beliefs about exposure to tax changes and present comparisons between a reduction in consumption taxes and a randomly assigned comparator program (either unconditional cash transfers or public works programs). In addition, we ask respondents a series of questions about their views on this policy, including who benefits the most from these exemptions, the fairness of this policy, and how likely it is to be subject to corruption and waste.

9. Real stakes questions. As is standard in online surveys (Stantcheva, 2023; Dechezleprêtre et al., 2025), we end the survey with a “real stakes” question to check that their survey responses reflect their true opinions. We ask respondents whether they would be willing to support a petition to the head of state promoting their preferred social assistance program (while preserving respondent anonymity).

10. Manipulation checks. In this section, we ask *all* respondents the attention check questions asked after the treatment videos, with minimally modified phrasing to make them understandable for all respondents. This allows us to directly measure the treatment effect on the precise beliefs targeted by the videos. Placing these questions at the very end of the survey serves two purposes. First, it ensures that asking these questions in itself cannot confound any treatment effects of the videos. Second, we will be able to measure a more persistent effect than if we ask these questions earlier.

The full survey instrument (Kenya version) is reproduced in Appendix C.

3.4 Organization of the experiment

Respondents are randomly assigned to treatment and control conditions just after the “General program information video and attention check” module. Within country and gender and age strata, participants are randomized at the individual level to one of two video treatments with probabilities of 30% (*Conditions faced by the Poor* and *Effectiveness*, described in Section 3.5) or to the control group with a probability of 40%. Because treatment assignment is randomized, differences in outcomes across groups can be interpreted as causal effects under standard assumptions.

Recruitment quotas ensure that the overall sample includes approximately 2,000 respondents from each country. Within each country-gender-age block, randomization assigns participants to treatment conditions as per the 30%, 30%, 40% split described above. This yields approximately $N = 16,000$ respondents in total across the eight countries, with 6,400 assigned to control and 4,800 assigned to each treatment group. The rationale for this allocation is discussed in Section 3.7.

Exogenous variation arises from the random assignment to the video treatments. By design, treatment assignment is independent from pre-treatment characteristics within strata, therefore identifying intent-to-treat (ITT) effects under standard SUTVA assumptions. Since the survey is administered individually and online, spillovers across respondents are extremely unlikely.

Formal blinding of respondents to their treatment status is not feasible in this design: each respondent is necessarily exposed to their assigned video (or no video, in the control condition). However, respondents are not informed of the existence of other treatment conditions, so they cannot infer what they did or did not see relative to other participants. The pre-registration of the analysis plan prior to data collection ensures that all analytical choices are made without knowledge of the outcomes, preventing post-hoc specification of hypotheses or estimation strategies.

In addition to the main treatment assignment, we randomize six other elements of the survey. These are:

- whether respondents are asked about their views on poverty and the effectiveness of programs before or after the main treatments,
- the fifth program type introduced to respondents (asset transfers or personalized tax refunds),
- the way the budget constraint would be closed if more money was spent on poverty reduction (higher taxes, lower government spending, higher foreign aid),
- whether respondents are asked how much they would gain from lower consumption taxes or how much they would lose from higher consumption taxes,
- the comparator program (either unconditional cash transfer or public works) used in the final consumption tax comparison, and
- the order in which the “all spending on a specific program” budget allocation questions are presented.

In addition, the order of all items is randomized within a given question whenever there is no natural order of the items. All random elements of the survey are randomized independently. In all cases, all options have equal probability.

3.5 Interventions

To test our experimental hypotheses, we implement two video treatments designed to provide respondents with accurate information in an accessible format. The intervention design follows recent methodological guidance on information experiments (Haaland et al., 2023; Stantcheva, 2023). The videos are produced separately for each country rather than dubbed from a common original. The people depicted in them reflect the ethnic composition of the country and prevailing norms about dress, so that the households shown are recognizable to respondents as ordinary poor households in their own setting.

Conditions faced by the Poor. The first treatment video presents accurate information on the conditions faced by poor households. This treatment is designed to affect respondents’ perceptions of poverty, corresponding to Box II of the conceptual framework. Table 5 summarises each claim in the video, the information presented to support it, and—in the third column—the specific belief item in Table 9 that each claim is expected to shift and in which direction. By shifting beliefs about poverty, we expect the video to increase overall support for poverty reduction spending, without necessarily shifting relative support across program types. Links to the full video for Kenya, Nigeria, Brazil, South Africa, and India are given in Table 3.

Table 5: Conditions faced by the Poor, video content

Topic	Content	Expected effect on beliefs
Living conditions	Poor households typically live in very basic conditions.	<i>Lives of the poor are harder</i> (D) ↑
Limited margin for error	Poor households have very little room for error compared to an average family.	<i>Lives of the poor are harder</i> (D) ↑
Limited opportunities	Poor households face more restricted opportunities than average families.	<i>Most can escape if work hard</i> (D) ↓

Notes: Variable names and panel letters in the third column refer to Table 9 (variable definitions); ↑/↓ = expected increase/decrease. The table summarizes the content of the Conditions faced by the Poor treatment video. This video presents respondents with information about the living conditions, vulnerability, and limited opportunities faced by poor households. We expect this treatment primarily to increase overall support for poverty reduction spending, without necessarily shifting relative support across program types. By construction, all three claims are expected to raise the “Perception of poverty index” in Panel E.

Effectiveness. The second treatment video addresses three common but incorrect beliefs about the effectiveness of social assistance programs, each contradicted by peer-reviewed evidence.¹¹ Table 6 summarizes each claim in the video and its expected effect on beliefs, following the same

¹¹ Among respondents to the second and third small-scale pilots, 49% believed at baseline that UCTs are wasted by recipients, and 62% believed that they make people work less (see Figure A1 in Appendix B). Moreover, 86% agreed with the statement: “Reducing taxes on essential goods for everyone is a good way for the government to support poor families.”

structure as Table 5. These claims map directly to Boxes VII (Recipients waste cash transfers), VIII (Cash transfers reduce labor supply) and VI (Tax exemptions benefit richer households) of the conceptual framework. We expect the video to increase overall support for poverty reduction spending, to increase relative support for cash transfers, and to decrease support for consumption tax exemptions. Links to the full video for Kenya, Nigeria, Brazil, South Africa, and India are given in Table 3.

Table 6: Effectiveness, video content

Myth	Evidence	Expected effect on beliefs
Recipients waste cash transfers	Research shows recipients of cash transfers do not increase spending on alcohol, tobacco, or other harmful goods.	<i>Recipients waste help</i> (F), UCT/CCT/Public works ↓
Cash transfers reduce labor supply	Research shows recipients do not reduce labor supply.	<i>Reduces labor supply</i> (F), UCT/CCT ↓
Tax exemptions on essential goods benefit the poor	Tax exemptions primarily benefit richer households, who spend more on these goods in absolute terms and at formal stores.	<i>Cuts help the poor</i> (K) ↓; <i>Cuts > SA</i> (K) ↓; <i>Cuts are fairer</i> (K) ↓; <i>Progressive</i> (K) ↓; <i>Prefers tax exemption</i> (K) ↓

Notes: Variable names and panel letters in the third column refer to Table 9 (variable definitions); ↑/↓ = expected increase/decrease. The table summarizes the content of the Effectiveness treatment video. This video presents respondents with evidence addressing common concerns about social assistance programs: that recipients of cash transfers waste the money on harmful goods, that cash transfers reduce labor supply, and that consumption tax exemptions primarily benefit poor households. The treatment is designed to affect beliefs about program effectiveness, corresponding to Boxes V–VIII and Arrow C in the conceptual framework. We expect this treatment to increase overall support for poverty reduction spending and to increase relative support for cash transfer programs. The prediction for *Recipients waste help* extends to public works programs because participants also receive cash; the prediction for *Reduces labor supply* applies to UCT and CCT only, as the evidence presented in the video is specific to unconditional and conditional cash transfers. Variable names for Myth 3 are abbreviated from Table 9 Panel K: *Cuts help the poor* = *Consumption-tax cuts help the poor*; *Cuts > SA* = *Consumption-tax cuts > social assistance*; *Cuts are fairer* = *Consumption-tax cuts are fairer*; *Progressive* = indicator that poor households benefit most from tax cuts; *Prefers tax exemption* = indicator that respondent prefers tax cuts over social assistance.

The control group is not shown a second video after the neutral explainer is shown to all respondents. The treatment effects should therefore be interpreted as the effect of exposure to a short, realistic information intervention relative to a common baseline explanation of program types, noting that we do not hold video exposure time constant. This estimand is policy-relevant because governments, civil society organizations, and media campaigns typically communicate information through short, salient, and accessible messages, not through content-neutral placebo communications. The estimates may therefore combine the informational content of the videos with the salience and engagement generated by receiving additional communication. We regard this as part of the evaluated intervention. Because both treatment groups receive an additional educational video in a similar format, we will also compare the two treatment arms directly; differences between these arms provide evidence on whether the specific content of the information mattered, holding constant the broad salience and engagement associated with receiving a second video.

3.6 Outcome variables

Our primary outcomes fall into three categories: the extent of support respondents express for poverty reduction programs overall (overall support), their support for particular program types (relative support), and support for consumption tax exemptions (tax exemptions). Secondary outcomes capture the perceptions and beliefs corresponding to Boxes I–VIII of the conceptual framework. In designing these measures, we follow recent survey methods for eliciting economic policy preferences (Stantcheva, 2023). Full definitions of all variables, including question wording, response scales, and coding, are given in Table 9.

3.6.1 Main outcome variables

Overall support. We measure overall support for poverty reduction programs using Likert-scale items, in line with best practices described in Stantcheva (2023), and monetary support for given budget splits (see Table 9, Panels H and I). For these outcome variables, respondents are not constrained, so they could express equally strong support for all poverty reduction programs.

Relative support. We measure relative support for poverty reduction programs using Likert-scale items, program rankings and a budget-allocation exercise (Panels H and I). We are interested in respondents’ differential support across program types, and particularly in their relative support for cash transfers.

Tax exemptions. We are also interested in respondents’ views toward consumption tax exemptions, as governments prioritize these policies over spending for poverty reduction programs despite their substantial shortcomings. We measure, on qualitative scales, whether respondents believe tax exemptions are a good policy, fair, less prone to corruption, and progressive, and the extent to which they prefer them relative to unconditional cash transfers and public works programs (Panel K). We use public works as a second comparator as this was the most popular type of poverty reduction program in the small-scale pilots of the survey.

3.6.2 Constructing indices for primary outcomes

To summarize the multiple variables captured within each of our three primary outcome families (overall support, relative support, and tax exemptions), we construct a standardized index for each family. For each component item, we standardize responses into z-scores using the mean and standard deviation of the control group, pooled across all eight countries, so that the standardized variable has a mean of zero and a standard deviation of one in the pooled control group. Items are signed so that higher values indicate greater support for poverty reduction spending, greater relative support for cash transfer programs, and greater support for consumption tax exemptions, respectively. Each index is then computed as the unweighted average of the standardized component items, where the average is taken over non-missing items for each respondent. The resulting average is then re-standardized using the mean and standard deviation of the control group, again pooled across countries, so that each index has mean zero and standard deviation one in the pooled control group; this is the scale in which the minimum detectable effects in Section 3.7 are expressed.

This three-step procedure follows the standard approach to summary indices in the experimental literature (Kling et al., 2007; Anderson, 2008; Dechezleprêtre et al., 2025). We use a common scale throughout the paper rather than standardizing within each country for two reasons. A within-country scale would make the same underlying effect appear larger where responses happen to be less dispersed. More fundamentally, because the index is an average of item z-scores, the item-level standard deviations determine the implicit weight each item carries: standardizing within country would make the index a differently weighted combination of the underlying items in each country, so that the country-specific estimates would not be estimates of the same quantity.

The components of each index are as follows.

Overall support index. The Overall support family is split into two indices, capturing qualitative and quantitative measures of support, respectively. The qualitative index combines five-point Likert items measuring general support for poverty reduction spending and support for each of the four core program types (UCTs, CCTs, in-kind transfers, public works programs), each coded from 1 (strongly oppose) to 5 (strongly support), and the support item for the randomly assigned fifth program (asset transfers or personalized tax refunds). We report two versions of this index: Index (core), which excludes the randomized fifth-program item, and Index (all), which includes it. The quantitative index combines respondents' preferred levels of total social assistance spending (relative to the "typical" amount we present as an anchor) under the seven hypothetical budget allocations elicited in the budget-allocation module: an even split across programs, allocations devoted entirely to each of the four core programs, the respondent's own preferred allocation, and the current government allocation. Write-in values replace slider responses that are at the upper cap, ensuring that the measure is not censored at the slider maximum.

Relative support index. This index, calculated separately for each program type, combines three measures of relative support: the five-point Likert item measuring support for the program type; the rank of the program type (from the bottom, so higher values are more preferred) in the ranking exercise including the randomized fifth program type; and the share of the budget allocated to the program type in the budget-allocation module.¹² Because the *Effectiveness* treatment is specifically designed to correct beliefs that depress support for cash transfers (that recipients waste money and reduce labor supply), we create an additional "cash transfer support index" that captures support for both unconditional and conditional cash transfers (UCT and CCT).

Tax exemptions index. This index combines five measures of support for consumption tax exemptions relative to targeted social assistance: the three five-point agreement scales measuring whether respondents regard reducing consumption taxes as a good way to support poor families, as helping poor families more than the comparator program, and as fairer than the comparator; a dummy equal to one if respondents say poorer households benefit most from tax exemptions (Progressive); and a dummy equal to one if respondents choose a tax-exemption reduction over the randomly assigned comparator program (either UCTs or public works) when asked which they would prefer to fund. All five components directly capture beliefs about the incidence and distri-

¹²As the budget allocation is only among the four "core" program types, this element is omitted for the fifth program type.

butional fairness of tax exemptions, which is the mechanism the *Effectiveness* treatment is designed to affect. The sixth item in Panel K—whether respondents believe tax exemptions are less prone to waste and corruption than social assistance programs—is conceptually distinct (it captures beliefs about administrative integrity rather than distributional incidence) and is therefore reported as a standalone descriptive item rather than included in the index.

3.6.3 Secondary outcome variables

Our secondary outcomes capture the perceptions and beliefs represented in Boxes I–VIII of the conceptual framework (see Table 9, Panels C–G).

Self-interest (Box I, Panel C). Self-interested motivations are measured through questions on current and expected benefit receipt, own perception of position on the income scale and likelihood of being in poverty, and perceived exposure to tax changes. In line with hypothesis H1a, we focus on *Family currently/expects to benefit* (Panel C) as our primary measure of self-interest; we also show robustness to measuring self-interest using a *Self-interest index* comprising *Family currently/expects to benefit*, *Another family benefits*, *Subjectively poor*, and *Likelihood of becoming poor*. These items are primarily used as predictor variables in the correlates regressions since they are measured prior to the treatment.

General views on poverty and government (Boxes II–IV, Panels D–E). Perceptions of poverty (Box II) and concern for the poor (Box III) are measured through Likert items on the prevalence and severity of poverty and how much respondents care about the poor, together with beliefs about whether poor households can escape poverty and the causes of poverty. Values regarding equality and government responsibility correspond to Box IV. To summarize these items, we construct five standardized indices (Panel E) using the same three-step z-score procedure described in Section 3.6.2: a *Perception of poverty index*, a *Concern for the poor index*, an *Internal causes of poverty index* (poverty attributed to individual factors), an *External causes of poverty index* (poverty attributed to structural factors), and a *Role of government index*.

Program-specific effectiveness beliefs (Boxes V–VIII, Panels F–G). Beliefs about the effectiveness of different program types are measured using two types of instruments. First, respondents assign a series of attribute statements to program types—for example, whether a program makes people work less or helps households escape poverty in the long run. For each program $p \in \{\text{UCT, CCT, In-kind, Public works, Fifth program}\}$, we summarize these items in four standardized indices (Panel G): a *Waste and corruption index* (p), a *Targeting accuracy index* (p), a *Benefits-for-beneficiaries index* (p), and a *Distortions and spillovers index* (p). Second, respondents complete a leaky bucket exercise in which they allocate hypothetical government spending across intended beneficiaries, unintended beneficiaries, administrative costs, and corruption. Third, we ask respondents directly about their perception of different programs’ direct benefit for beneficiaries relative to unconditional cash transfers.

Correspondence with the welfare framework. Several of these measures correspond directly to objects in Section 2.1. The leaky bucket exercise elicits the perceived share of spending that reaches

intended beneficiaries (*Reaches the poor*, Panel F of Table 9), which is the respondent's perception of α . The pairwise vignettes elicit the implicit valuation of each program relative to an unconditional cash transfer (*Implicit relative valuation*, Panel J), which is the respondent's perception of beneficiaries' δ , normalized to one for cash by construction. The *Distortions and spillovers index* (Panel G) captures the indirect costs and benefits that enter ϵ and, where they operate through the public finances, FE . While we do not directly attempt to recover respondents' welfare weights ψ or how they perceive beneficiaries' marginal utilities of income λ , the questions on self-interest, perception of poverty, and concern for the poor target the same underlying preferences and beliefs.

Open-ended responses. Two open-ended questions ask respondents to explain the reasoning behind their choices: the main reason for preferring their top-ranked program, and the rationale for their budget allocation.

3.7 Sample and statistical power

Our target sample size is 2,000 respondents in each country, giving a total sample size of 16,000 across the eight study countries. This sample size was determined by examining the sample size of other cross-country randomized survey experiments in this field and conducting statistical power calculations.

Under the randomization described in Section 3.4, respondents are allocated to three mutually exclusive treatment cells. Two video treatments (Conditions faced by the Poor and Effectiveness) each account for 30% of the sample, while 40% of the sample is assigned to the control group. Hence, within each country, approximately 600 respondents will be assigned to each treatment group (1,200 in total) and 800 to the control group. The larger share allocated to control (40% versus 30% per treatment arm) reflects two considerations. First, the raw means of baseline support for poverty reduction programs and of attitudes toward consumption tax exemptions are computed on the control group only, since these items are elicited after the treatment stage for every respondent (Section 4.2); a larger control group improves the precision of these estimates. Second, the additional control-group observations increase statistical power for the pooled treatment-effect estimates, because in this specification both treatment arms are compared against the same control group. Haaland et al. (2023) summarize best practices in online randomized survey experiments providing information interventions, and suggest that having in the order of 700 respondents per treatment/control group should provide adequate power to detect a meaningful effect if one exists. As a result, the comparison between a single treatment (600 respondents) and the control group (800 respondents) in a single country is powered as recommended.

We consider three broad scenarios for specific power calculations. The first is to consider the minimum detectable effect (MDE) (with type one and type two errors set at probabilities of 0.05 and 0.2, respectively) from receiving any of the main treatments on a single binary outcome - for example, collapsing a 5 point Likert-scale into a dummy variable that takes the value of 1 for strongly support and support, and 0 otherwise. The survey experiment could detect an effect of roughly 6.4 percentage points, assuming 50 percent of the control group agrees with the statement in question.

The second is to consider the MDE (with the standard probability of type one and type two errors

of 0.05 and 0.2) from a single treatment on a single binary outcome (using the example of collapsing a 5-point scale into a dummy variable that takes the value of 1 for strongly support or support and 0 otherwise). This is similar to the first scenario, but instead of combining the treatment groups, they are treated separately. The survey experiment could detect an effect of about 7.6 percentage points, assuming 50 percent of the control group agrees with the statement in the question.

The third is to consider the MDE for a subset of respondents from a single treatment on a single binary outcome. In this case, the underlying sample size in each of the treatment and control groups will shrink based on the prior beliefs and preferences of respondents. Three examples are considered, the first covering 75 percent of all respondents, the second covering 50 percent of all respondents and the third covering 25 percent of all respondents. The effect size that could be detected with these smaller sample sizes are shown in Table 7, once again based on the conservative assumption of 50 percent of the control group agreeing with the statement in the question.

Table 7: How the minimum detectable effect in a single country varies based on sample size

	600/800 (100%)	450/600 (75%)	300/400 (50%)	150/200 (25%)
Respondents in the combined treatment/control group with a particular prior belief/preference	100%	75%	50%	25%
Minimum detectable effect (percentage points)	7.6	8.7	10.7	15.1

Note: This table shows how the minimum detectable effect in a single country varies based on the sample size.

These calculations imply that the study is well powered for analyses even within a single country and adequately powered for the principal treatment-control contrasts within countries, especially where treatment arms are combined in line with the main hypotheses. However, power will be more limited for highly disaggregated comparisons, treatment interactions, and subgroup analyses within individual countries. We therefore treat the main treatment-control contrasts as our primary focus and interpret more granular heterogeneity analyses with appropriate caution.

There are at least three reasons to believe the randomized survey experiment will potentially have much higher statistical power to detect effects than what is presented above. Firstly, in the instance whereby the mean in the control group for a binary outcome is higher or lower than 0.5, the MDE will be smaller. Secondly, by pooling data across countries (with country fixed effects as discussed in the following section), there will be substantially more power to detect an impact from the treatments. Yet, it is hard to quantify this ex-ante as it will depend on variation between countries. Thirdly, three main indices will be created across the outcome variables (see the following Section for details), and it will be easier to detect an effect on these indices than on a single outcome. However, the extent to which this is the case will vary depending on how consistent the treatment effects are across the outcomes.

Power for the primary outcome indices. The calculations above describe a single binary item in a single country, and are deliberately conservative. Our pre-specified primary outcomes, however, are the standardized indices defined in Section 3.6.2 — the qualitative and quantitative absolute-

support indices, the relative-support index for cash transfers (both UCT on its own and UCT + CCT combined), and the tax-exemptions index — *pooled across countries* with country fixed effects. We therefore also report MDEs for this estimand directly. Because each index is standardized to have a standard deviation of one in the control group, MDEs are naturally expressed in control-group standard-deviation (SD) units. Under the 40/30/30 allocation described in Section 3.4, the full sample yields approximately 6,400 control respondents, 9,600 in the two treatment arms combined, and 4,800 in each individual arm. Table 8 reports the resulting MDEs for the principal pooled contrasts.

Table 8: Minimum detectable effects for primary outcome indices, pooled across countries (SD units)

Contrast	Treatment N	Comparison N	MDE (SD)
Pooled treatment vs. control (Panel A)	9,600	6,400	0.045
Single treatment arm vs. control (Panel B)	4,800	6,400	0.054
Conditions faced by the Poor vs. Effectiveness	4,800	4,800	0.057
Financing assumption, pairwise (H5a)	5,333	5,333	0.054
Pooled treatment $\times$ median-split interaction	—	—	0.090
Single arm $\times$ median-split interaction	—	—	0.107

Note: Calculations assume $\alpha = 0.05$ (two-sided), power 0.80, and a control-group SD of one. Interaction rows give the MDE for the difference in treatment effects across a within-country median split of a pre-treatment moderator. These MDEs are conservative in three respects: they ignore the residual-variance reduction from the country fixed effects and pre-specified covariates; they treat each index as having unit residual SD even though aggregation of multiple standardized items into an index reduces idiosyncratic noise relative to any single component; and they assume equal treatment- and control-group variances.

These figures indicate that the study is well powered for its primary analyses: the pooled treatment–control contrast on each index can detect effects of approximately 0.045 SD, and the individual-arm contrasts approximately 0.054 SD. We regard an effect of 0.05 SD on a primary index as the smallest magnitude that is substantively meaningful for the policy questions motivating this study, and the pooled and single-arm contrasts are powered at or below this threshold.

The three financing assumptions in H5a are randomized independently of treatment assignment and with equal probability, so each is seen by approximately 5,333 respondents and a pairwise contrast between two assumptions is powered at approximately 0.054 SD—the same as a single treatment arm against control. Power against the conjunction H5a requires is somewhat lower, since it asks for both pairwise differences.

Consistent with the caveats above, the design is slightly less well powered for disaggregated analyses. The median-split heterogeneity tests in Equation (5) in Section 4.4 can detect interaction effects of approximately 0.09 SD (pooled treatment) and 0.11 SD (single arm) in the full sample; the H4a, H4b, and H4c analysis, which is restricted to the half of respondents who answered the views on poverty and poverty reduction in general and the effectiveness modules before treatment, is powered for interaction effects of roughly 0.15 SD. H4d and the six additional prespecified moderators are estimated on the full sample, and are therefore powered at the 0.09–0.11 SD figures above, as is the module-placement test described below. Country-specific index estimates are correspondingly imprecise—with per-country MDEs of approximately 0.13 SD (pooled) and 0.15 SD (single arm)—which is why we will treat the pooled estimates as our primary evidence and report

country-specific and granular heterogeneity results in an appendix.

3.8 Placement of the belief modules relative to treatment

The *Views on poverty* and *Beliefs about program effectiveness* (‘beliefs and perceptions’) modules are randomly placed either before or after the treatment stage (Section 3.3). This enables us to estimate “first-stage” treatment effects on beliefs and perceptions using the half of the sample that saw these modules *after* treatment (the first additional analysis in Section 4.5). The beliefs and perceptions modules include many beliefs the videos do not address, which is what allows us to ask whether a treatment moves the beliefs it targets and leaves others unchanged (Section 5.3). We can also test heterogeneous effects depending on beliefs and perceptions, as per hypotheses H4a–H4c, using the other half of the sample, which answered these questions *before* treatment. A reduction in the sample size by half causes the minimum detectable effect in these analyses to increase by a factor of $\sqrt{2}$.

The cost is limited to those two analyses. The primary experimental hypotheses H2a–H2d are tested using the full sample, since the primary outcome questions are asked after the treatment stage for all respondents. Similarly, the tests of the treatment effects on agreement with the six specific claims in the videos (‘manipulation checks’) use the full sample, as the relevant questions are asked at the end of the survey for all respondents (Panel M of Table 9). Tests for Hypothesis H4d and other hypotheses regarding heterogeneity by background characteristics are also unaffected, as all respondents are asked the relevant questions before the treatment.

We prefer this to either fixed-placement alternative. Administering the modules *before* treatment to all respondents would mean we could not measure effects on the full battery of “first-stage” beliefs and perceptions, beyond the six claims covered by the manipulation checks, depriving us of information about the mechanism by which the treatments affect support for different poverty reduction programs. Administering both modules *after* treatment to all respondents would leave no pre-treatment measure of the beliefs in Boxes II–VIII, so H4a–H4c could not be tested at all. We acknowledge that a reader who placed a different relative weight on the heterogeneity tests and the evidence on mechanisms would reasonably have chosen differently.

A third option—measuring views and perceptions *after* treatment for everyone, but moving a few items forward to partition respondents for the heterogeneity analysis—does not avoid this trade-off. Any item moved forward is no longer available for measuring “first-stage” effects. Moreover, the moderators for H4a–H4c are indices rather than single items, so either every constituent item would need to be moved forward or the partition would rest on a noisier proxy, which would cause attenuation bias. The difficulty is greatest for H4c, whose moderator comprises all fourteen program-characteristic items for each of the four core programs; moving them all forward would essentially amount to the first alternative above. Duplicating questions could avoid this problem but might confuse respondents (especially in the control group) and would likely increase survey fatigue, negatively affecting response quality throughout the survey.

The randomization brings two further advantages. First, we can test whether the placement of these modules affects the measured treatment effects (see Section 4.5); placing them before the

treatment might, for example, reduce treatment effects as respondents try to maintain consistency with their previous answers (Haaland et al., 2023). Second, within the control group, who see no treatment video under either placement, we can test whether position in the survey affects the answers to the belief questions themselves, which bears on the interpretation of descriptive means (Section 4.2).

3.9 Variation from the intended sample size

The two principal sources of deviation from the intended sample are partial completions and inattentive responding.

Partial completions. Respondents who exit before reaching the final page of the survey are not counted toward the target of 2,000 complete responses per country (see *Data collection and processing* below) and are excluded from the analytical sample. Dropping partial completions is standard practice in online survey experiments, and is consistent with the design: the survey platform requires a response to each item before the respondent can advance, so a partial completion reflects an exit from the survey rather than item non-response within a completed interview. In both small-scale pilots, completion rates among those who passed the initial screening were high, and we expect similar rates in the main wave.

Inattentive responding. The survey tests whether respondents are paying adequate attention through a program knowledge check administered immediately after the explainer video. Respondents are informed they will be quizzed after the video, and those scoring below 7 out of 10 are shown the correct answers and asked to retake the quiz; those who fail the second attempt are excluded from the analytical sample. Based on our small-scale pilot data, we expect around 90 percent of respondents to adequately complete the quiz. Because this exclusion occurs before randomization, it is a pre-randomization screen rather than attrition. To guard against under-recruitment after this exclusion, the survey firm will recruit a modest buffer of around ten percent above the 2,000-per-country target, consistent with the approach taken in the small-scale pilots. Importantly, respondents in the two treatment arms also complete a post-video comprehension question; however, those who answer incorrectly are not excluded from the analysis. Their score is recorded and used in the data-quality robustness checks described in Section 4.8.7.

3.10 Data collection and processing

Primary data collection will be conducted through our online survey. **Stopping rule:** data collection in each country will cease once 2,200 complete responses have been recorded (defined as reaching the final page of the survey). No interim analyses will be conducted during fieldwork, and the stopping criterion is fixed ex ante and independent of any observed outcomes. We will run the survey once in each country, allowing individuals who participate in the survey to respond only once. Based on previous studies, we expect the data collection process to take only a matter of days in each country.

3.10.1 Data management

Because the survey collects no personally identifiable information, data management is straightforward. Respondent anonymity is protected by a structural decoupling between contact information and survey responses: the survey firm contacts respondents and distributes survey links but never receives their answers, while the survey platform (Alchemer) hosts the responses but holds no contact details. Neither the survey firm nor the research team has access to both a respondent's identity and their answers. In the unlikely event that personally identifiable information is inadvertently shared with the research team, it will be deleted immediately. The analytical dataset provided to the research team will therefore contain no information that could identify individual respondents. The anonymized dataset and replication code will be made publicly available upon publication.

3.10.2 Small-scale pilot data

We conducted three small-scale pilot surveys prior to the full-scale launch across countries, each targeting an analysis sample of around 200 respondents. The first (Kenya, December 2025) ran without any experimental treatments and had the following objectives:

Eliciting first-order concerns In line with previous survey experiments (see Stantcheva, 2020, and Ferrario and Stantcheva, 2022), we make use of open-ended questions to understand the topics that are at the top of respondents' minds.¹³ In particular, we asked respondents their views about what the government can do to reduce poverty, for their feedback on the survey and the logic behind the choices they made. These responses were used to inform the options provided in the closed-ended questions of the final survey.

Understanding of questions The first small-scale pilot was important for checking that respondents understand what we are asking in each question. This is particularly important for questions that require an understanding of mathematics or that ask about potentially complex economic topics. To check this, we include consistency checks throughout the survey, which either explicitly ask the respondent to confirm their understanding, or implicitly do so by asking a follow-up question that should align with their previous response. If responses consistently do not align, this would suggest that respondents do not fully understand the questions.

Engagement and fatigue The first small-scale pilot also allowed us to estimate respondent engagement and fatigue. We do this in a number of ways, including looking for patterns in responses, the time spent on a page, and the number of words used in open-ended questions. We also randomize the order of sections of the survey, allowing us to analyze different response patterns depending on where a question appears.

The second small-scale pilot (Kenya, May 2026) introduced the two video treatments. It also included a third arm using AVA, an AI chatbot developed by the World Bank, as an additional information treatment. This arm was subsequently dropped from the main study design because AVA did not reliably generate contextually appropriate responses for the target audience.

¹³See Haaland, Roth, Stantcheva and Wohlfart (2024a) for a review of the use of open-ended questions in surveys.

A third small-scale pilot (Nigeria, May 2026) replicated the two video treatment design and provided initial evidence on the cross-country generalizability of the survey instrument. As we show in Tables A1 to A4 in Appendix B, based on our pre-specified construction of outcomes, we already observe meaningful effect sizes from both video treatments in the second and third small-scale pilots. This provides considerable assurance that the main study will be adequately powered to detect meaningful effects if they exist.

Importantly, the small-scale pilot data will not be used for our main analysis, and we are also open to disregarding the full-scale wave of data collection from Summer 2026 if required to address reviewer feedback.

4 Empirical analysis

4.1 Coding of variables

Table 9: Variable definitions

Variable	Definition
<i>Panel A: Treatment assignment (indicator variables)</i>	
Treatment: Control	Respondent was randomized to see no information video, i.e. the control group (usually omitted category in regressions).
Treatment: Conditions faced by the Poor	Respondent was randomized to see the video describing the living conditions, vulnerabilities, and opportunities faced by poor households.
Treatment: Effectiveness	Respondent was randomized to see the video that addresses common but empirically incorrect beliefs about social assistance programs.
Fifth program	Categorical variable indicating which program the respondent saw as the (randomly assigned) fifth program in addition to UCT, CCT, in-kind, and public works. Values: <i>Asset transfers, Personalized tax refunds</i>
Budget-constraint: foreign donors	Indicator for the funding assumption in Panel I being “spending will be paid for by foreign donor countries”.
Budget-constraint: higher taxes	Indicator for the funding assumption in Panel I being “spending will result in higher taxes for the general population”.
Budget-constraint: spending cuts	Indicator for the funding assumption in Panel I being “spending will result in lower government spending on other things”.
Consumption-tax comparator	Categorical variable for which social-assistance program is used as the comparator in the consumption-tax vs. social-assistance module.
<i>Panel B: Socio-economic characteristics and background views (indicator variables unless noted)</i>	
Age 18–24	Respondent is 18–24 years old (usually omitted category).
Age 25–34	Respondent is 25–34 years old.
Age 35–49	Respondent is 35–49 years old.
Age 50+	Respondent is 50 or older.
Male	Respondent reports gender as male (usually omitted).
Female	Respondent reports gender as female.
Other / prefer not to say (gender)	Respondent reports gender as “other or prefer not to say”.
National	Respondent is a national of the surveyed country.
Resident	Respondent currently lives in the surveyed country.

Continued on next page

Table 9 – continued from previous page

Variable	Definition
Ethnicity	Categorical variable for the respondent's ethnic group, asked in Brazil, Colombia, Kenya and South Africa. Categories are country-specific (e.g. in Kenya: Kikuyu, Luhya, Luo, Kalenjin, Kamba, Other, Prefer not to say).
Religion	Categorical variable for the respondent's religious affiliation, asked in Egypt, India, Indonesia and Nigeria. Categories are country-specific.
Largest group	Indicator equal to 1 if the respondent reports belonging to the largest ethnic or religious group in their country, as recorded in the most recent census or equivalent national statistics; 0 otherwise; missing if "Prefer not to say".
Largest group among poor	Indicator equal to 1 if the respondent reports belonging to the ethnic or religious group that is the largest <i>among the poorest 20%</i> of people by equivalized household income in their country. Data on the largest ethnic or religious group among the poorest 20% will be sourced from nationally representative household surveys. The variable will be coded as missing where these data are not available.
Large city	Respondent lives in the largest urban areas of the country (e.g. Nairobi or Mombasa in Kenya).
Small city/town	Respondent lives in a town with fewer than one million but more than ten thousand people.
Rural	Respondent lives in a small town or rural area with fewer than ten thousand people.
Single	Respondent reports marital status as single (usually omitted category).
Married	Respondent reports marital status as married.
Divorced	Respondent reports marital status as divorced.
Widowed	Respondent reports marital status as widowed.
No education	Respondent has not completed any formal education (usually omitted category).
Primary education	Highest level completed is primary school.
Secondary education	Highest level completed is secondary school.
Post-secondary, non-university	Highest level completed is higher than secondary but not university (e.g. vocational, diploma).
University	Respondent has at least a university degree.
Employee with contract	Respondent is currently employed under a formal contract (usually omitted category).
Employee without contract	Respondent is currently employed without a formal contract.
Self-employed or business owner	Respondent is self-employed or owns a business.
Not working, looking for work	Respondent is not currently working but is looking for work.
Not working, not looking	Respondent is not currently working and is not looking for work.
Student	Respondent is a student.
Personal income (categorical)	Self-reported income from work in a typical month of 2025, in local currency. Six brackets, country-specific.
Household adults	Number of adults aged 18+ in the household. Coded 1, 2, 3, 4, 5+.
Household children	Number of children aged 0–17 in the household. Coded 0, 1, 2, 3, 4, 5+.
Per-adult household income (categorical)	Respondents report total household income in a typical month of 2025, in local currency, within six income brackets. The brackets are proportional to the number of adults in the household.
Risk tolerance	Self-reported willingness to take risks, 1–5 slider (1 = very unwilling, 5 = very willing).
Altruism	Self-reported willingness to give to others without expecting anything in return, 1–5 slider.

Continued on next page

Table 9 – continued from previous page

Variable	Definition
Supports current government	Indicator equal to 1 if the respondent answers “Yes” to “Do you support the current government?”; 0 if “No”; missing if “Prefer not to say”.
Votes on poverty programs	Respondent’s answer to “When choosing who to vote for in national elections, do you consider candidates’ views on government programs to help the poor?” Five ordered options coded 1–5: 1 = <i>No, not at all</i> ; 2 = <i>No, not much</i> ; 3 = <i>Neither yes nor no</i> ; 4 = <i>Yes, a bit</i> ; 5 = <i>Yes, a lot</i> .
Panel C: Self-interest (Box I of the conceptual framework)	
Heard of program x	Indicator variables, one per country-specific social-assistance program (sourced from ASPIRE), equal to 1 if the respondent indicates having heard of that program.
Family currently/ expects to benefit	Indicator equal to 1 if the respondent selects “Yes, my family” or “Both” in response to “Does your family, or a family you know, currently benefit from any of these programmes, or do you expect this to be the case in the future?” for at least one program.
Another family benefits	Indicator equal to 1 if the respondent selects “Yes, another family” or “Both” in response to the same question for at least one program.
Subjective income group	Respondent’s self-placement of their household among five equal-sized national income groups, 1 (poorest) to 5 (richest).
Subjectively poor	Indicator equal to 1 if the respondent placed their household in the bottom quintile (subjective income group = 1).
Likelihood of becoming poor	Self-reported likelihood of being in the poorest group at any point in the next ten years, 1–5 slider (1 = very unlikely, 5 = very likely).
Likelihood of escaping poverty	Self-reported likelihood of moving out of the poorest group within the next ten years, 1–5 slider. Conditional on subjective income group = 1, otherwise missing.
Exposure to consumption tax (loss)	Respondent’s answer to “If the government were to increase these [consumption taxes on essential items], how much would your family lose out financially compared to the average family in [country]?” 1–5 slider. <i>Half of respondents are asked the question in this framing, the other half are asked the question in the gain framing.</i>
Exposure to consumption tax (gain)	Respondent’s answer to “If the government were to decrease these taxes, how much would your family save financially compared to the average family in [country]?” 1–5 slider. <i>Half of respondents are asked the question in this framing, the other half are asked the question in the loss framing.</i>
Panel D: General views on poverty (Boxes II–IV)	
Lives of the poor are harder	Respondent’s answer to “How hard do you think the lives of families in the poorest group are compared to families in the middle group?” on a 1–5 scale (1 = not at all hard, 5 = extremely hard). When used as an indicator, equals 1 if the respondent answers 4 or 5.
Concern for the poor (raw)	Respondent’s answer to “How concerned would you say you are about the lives of families in the poorest group?” on a 1–5 scale.
Most can escape if work hard	Respondent’s answer to “What fraction of the families in the poorest group do you think will move out of the poorest group within the next ten years if they work hard?” Five labelled options: 1 = <i>Nobody</i> ; 2 = <i>Around a quarter</i> ; 3 = <i>Half</i> ; 4 = <i>Around three quarters</i> ; 5 = <i>Everyone</i> . When used as an indicator, equals 1 if the respondent answers 4 (<i>Around three quarters</i>) or 5 (<i>Everyone</i>).

Continued on next page

Table 9 – continued from previous page

Variable	Definition
<i>Causes of poverty (1–5 scale, 1 = not very important, 5 = very important):</i>	
Lack of ability or talent	Importance attached to lack of ability or talent as a cause of poverty.
Lack of effort	Importance attached to lack of effort.
Poor life choices	Importance attached to poor life choices.
Poor health	Importance attached to poor health.
Lack of access to jobs	Importance attached to lack of access to well-paying jobs.
Lack of money to invest / high credit costs	Importance attached to lack of money to invest in a business or too high borrowing costs.
Low education	Importance attached to low education.
<i>Normative views and role of government (1–5 agreement scale; indicator = 1 if respondent agrees or strongly agrees, i.e. scale value ≥ 4):</i>	
Right to basic goods	“Everyone has a right to basic food, housing and healthcare.”
Right to a fair-wage job	“Everyone should have access to a job that pays them a decent wage.”
Working-age should earn	“People of working age with access to jobs should earn their own living, provided they are not disabled or have to take care of children.”
Government should reduce poverty	“It is the responsibility of governments to reduce poverty.”
Government can reduce poverty	“Governments have the capacity to effectively reduce poverty.”
Prefers high taxes on wealthy	Respondent’s answer to “What would do more to reduce the gap between the rich and the poor in [country]?” Indicator equal to 1 if the respondent chooses “High taxes on the wealthy and corporations to fund programmes that help the poor” over “Low taxes on the wealthy and corporations to encourage investment and economic growth”.
Panel E: Summary indices on general views	
	Each index is constructed in three steps (following Kling et al., 2007; Dechezleprêtre et al., 2025): (i) each underlying variable is converted to a z-score using the control-group mean and standard deviation, with signs oriented consistently within domain; (ii) the z-scores are averaged; (iii) the resulting average is standardized again by subtracting its mean and dividing by its standard deviation.
Perception of poverty index	Index based on: <i>Lives of the poor are harder</i> , – <i>Most can escape if work hard</i> . Higher values indicate that the respondent perceives the lives of the poor as harder and poverty as more difficult to escape.
Concern for the poor index	Index based on: <i>Concern for the poor (raw)</i> , <i>Right to basic goods</i> , <i>Right to a fair-wage job</i> , – <i>Working-age should earn</i> . Higher values indicate stronger concern for the welfare of poor people and stronger normative commitment to supporting them.
Internal causes of poverty index	Index based on the <i>Causes of poverty</i> sub-variables that attribute poverty to individual characteristics: lack of ability, lack of effort, poor life choices, poor health. Higher values indicate stronger attribution of poverty to individual factors.

Continued on next page

Table 9 – continued from previous page

Variable	Definition
External causes of poverty index	Index based on the <i>Causes of poverty</i> sub-variables that attribute poverty to systemic factors: lack of access to well-paying jobs, lack of money / high credit costs, low education. Higher values indicate stronger attribution of poverty to systemic factors.
Role of government index	Index based on: <i>Government should reduce poverty, Government can reduce poverty, Prefers high taxes on wealthy</i> . Higher values indicate stronger belief that the government can and should play a role in reducing poverty.
Panel F: Perceived effectiveness of each program (Boxes V–VIII)	
	All perception variables below are elicited separately for each of the five programs $p \in \{\text{UCT, CCT, In-kind, Public works, Fifth program}\}$. The two “program characteristics” modules present each statement and ask the respondent to tick <i>all</i> programs to which it applies. The variable is coded 1 if the respondent ticked program p and 0 otherwise.
<i>Box V – Perceived waste and corruption</i>	
Wasted on bureaucracy (p)	“A lot of money [spent on this program] will be wasted on inefficient government bureaucracy.”
Stolen by corrupt politicians (p)	“A lot of money [spent on this program] will be stolen by corrupt politicians.”
<i>Box VI – Perceived targeting accuracy</i>	
Leaks to non-poor (p)	“A lot of money [spent on this program] will go to people who are not actually poor.” (Type II / inclusion error.)
Excludes the poor (p)	“Many people who are genuinely poor will be left out [of this program].” (Type I / exclusion error.)
<i>Box VII – Perceived benefits for beneficiaries</i>	
Recipients waste help (p)	“Recipients [of this program] waste the help they receive.”
Helps escape poverty (p)	“[This program] can help individuals escape poverty in the long run.”
Protects against price risk (p)	“[This program] can protect recipients against fluctuations in prices of local goods and services.”
Does not give what they want (p)	“[This program] doesn’t give the poor what they really want.”
Burdensome to qualify (p)	“It is too burdensome to qualify for this programme.”
Shameful for recipients (p)	“It is shameful for recipients to receive help from this programme.”
<i>Box VIII – Perceived distortions and spillovers</i>	
Stimulates local economy (p)	“[This program] can stimulate the local economy.”
Raises local prices (p)	“[This program] leads to higher prices for goods consumed by poor people.”
Reduces labor supply (p)	“[This program] makes people work less.”
Raises local wages (p)	“[This program] leads to higher wages for regular jobs carried out by poor people.”
<i>Box V (general) – Allocation task:</i>	
	“Suppose the government were to spend [amount] on social assistance programs. How much would actually reach beneficiaries and where would the rest go?” Four shares summing to [amount]:
Reaches the poor	Amount that goes to intended beneficiaries (poor families).
Leaks to non-poor	Amount that goes to unintended beneficiaries (non-poor families).

Continued on next page

Table 9 – continued from previous page

Variable	Definition
Administration	Amount absorbed by the cost of administering the program.
Corruption	Amount lost to corruption.
Panel G: Effectiveness indices (per program p)	
	Constructed using the same three-step z-score procedure as the general-view indices.
Waste and corruption index (p)	Index of <i>Wasted on bureaucracy</i> (p) and <i>Stolen by corrupt politicians</i> (p). Higher = more perceived waste and corruption.
Targeting accuracy index (p)	Index of <i>Leaks to non-poor</i> (p) and <i>Excludes the poor</i> (p). Higher = more perceived targeting accuracy.
Benefits-for-beneficiaries index (p)	Index of <i>Helps escape poverty</i> (p), <i>Protects against price risk</i> (p), <i>Recipients waste help</i> (p), <i>Does not give what they want</i> (p), <i>Shameful for recipients</i> (p), <i>Burdensome to qualify</i> (p). Higher = more perceived benefits to beneficiaries.
Distortions and spillovers index (p)	Index of <i>Stimulates local economy</i> (p), <i>Raises local wages</i> (p), <i>Raises local prices</i> (p), <i>Reduces labor supply</i> (p). Higher = more positive spillovers.
Panel H: Outcomes – support for social assistance	
Supports social assistance (general)	Indicator equal to 1 if the respondent answers “Support” or “Strongly support” to a question on support for the government using social assistance programs to support poor households.
Supports program p	For $p \in \{\text{UCT, CCT, In-kind, Public works, Fifth program}\}$: indicator equal to 1 if the respondent rates support for program p as 4 or 5 on the 1–5 Likert scale (1 = strongly oppose, 5 = strongly support).
Support score p (Likert)	Raw 1–5 Likert rating of support for program p , used as a continuous outcome.
Rank of program p	Integer rank from 1 (most preferred) to 4 (least preferred) elicited via a drag-and-drop ordering of the four core programs. The fifth-program question asks respondents to drag the fifth program into their existing ordering, yielding a 1–5 ranking by insertion rather than a fresh elicitation.
Most preferred program	Categorical variable indicating the program ranked first.
Panel I: Outcomes – preferred spending levels and allocation	
	For all questions on overall spending in this section, a randomly selected third of respondents is told to assume each of (a) “spending will be paid for by foreign donor countries”, (b) “spending will result in higher taxes for the general population”, or (c) “spending will result in lower government spending on other things”.
Spending: even split	Preferred monthly spending per family member, in local currency, when budget is split evenly across the four core programs. Slider range 0 to twice the country-specific anchor amount; an open-ended text box captures values above the slider maximum (capped at approximately five times the anchor). Anchor amounts are 5% of GDP per capita (rounded).
Spending: p only	For each program p : preferred monthly spending per family member if that program were the only social-assistance program the government funded.
Allocation share to p	For each program p : share of a hypothetical social-assistance budget the respondent allocates to program p . Shares sum to budget across the four programs.

Continued on next page

Table 9 – continued from previous page

Variable	Definition
Spending: preferred split	Preferred monthly spending per family member under the respondent's own allocation. Continuous, with the same open-ended top-up.
Spending: current (government) split	Preferred monthly spending per family member under the country's actual current allocation across programs (from ASPIRE). Respondents are shown the actual allocation as a diagram.
Panel J: Outcomes – pairwise comparisons / willingness-to-pay	
<i>For each pair q vs UCT, with $q \in \{\text{CCT, In-kind, Public works}\}$:</i>	
Most families would choose q	Indicator equal to 1 if the respondent thinks most poor families would choose program q over a UCT of equal nominal value (country-specific anchor amount A per family member per month, e.g. KSh 1,000 in Kenya); 0 if the respondent thinks most would choose UCT or that families would be evenly split.
Most families would choose UCT	Indicator equal to 1 if the respondent thinks most poor families would choose the UCT; 0 if the respondent thinks most would choose q or that families would be evenly split. Both indicators are 0 when the respondent selects <i>Families would be evenly split between the two programmes</i> ; the WTP follow-up is shown only to respondents who chose one program as preferred.
Implicit relative valuation of q	Implicit valuation of q relative to UCT elicited as described in the main text. Pre-set options at 1/2, 2/3, 4/5, 10/11, 1, 1.1, 1.25, 1.5, 2; open-ended box for values below 1/2 (minimum approximately 1/5) and above 2 (maximum approximately 5).
Panel K: Outcomes – consumption-tax comparison	
Consumption-tax cuts help the poor	Respondent's answer to "Reducing taxes on essential goods for everyone is a good way for the government to support poor families." 1–5 Likert (1 = strongly disagree, 5 = strongly agree); indicator = 1 if ≥ 4 .
Consumption-tax cuts > social assistance	Respondent's answer to "Reducing taxes on essential goods for everyone would help poor families more than [comparator program] would." 1–5 Likert; indicator = 1 if ≥ 4 .
Consumption-tax cuts are fairer	Respondent's answer to "Reducing taxes on essential goods for everyone is fairer than [comparator program] because everyone benefits." 1–5 Likert; indicator = 1 if ≥ 4 .
Who benefits most from tax cut	Respondent's answer to "Who would benefit the most from reducing taxes on essential goods for everyone?" Categorical: <i>Richer households, Poorer households, Rich and poor households would benefit the same, Don't know</i> . The variable <i>Progressive</i> is an indicator equal to 1 if the respondent selects <i>Poorer households</i> .
Less wasted: SA vs tax cut	Respondent's answer to "Which of the following do you think loses less money to waste and corruption?" Categorical: <i>[Comparator program] to poor families, Reducing taxes on essential goods for everyone, About the same, Don't know</i> . Indicator equal to 1 if respondent selects <i>Reducing taxes on essential goods for everyone</i> .

Continued on next page

Table 9 – continued from previous page

Variable	Definition
Preferred use of spending	Respondent's answer to "Suppose the government has [amount] to help poor families. Which of the following options would you prefer?" Categorical: <i>[Comparator program] to poor families, Reduce taxes on essential goods for everyone, Split the amount between the two, Neither, Don't know</i> . The variable <i>Prefers tax exemption</i> is an indicator equal to 1 if the respondent selects <i>Reduce taxes on essential goods for everyone</i> .
Panel L: Real-stakes	
Supports petition	"Real stakes" outcome: respondent's answer to "Would you be willing to support a petition promoting this type of programme? As soon as the survey is complete, we will send the results to the president's office, informing them what share of people who took this survey were willing to support this type of programme." Indicator = 1 for <i>Yes</i> ; 0 for <i>No</i> or <i>Maybe</i> .
Panel M: Manipulation checks	
	Respondents are shown two multiple-select questions and asked to tick <i>all</i> options that apply. Each item below is coded as an indicator equal to 1 if the respondent selected that option and 0 otherwise.
<i>Conditions faced by the Poor:</i>	
Basic living conditions	"Poor people live in very basic living conditions."
Little room for error	"Poor people live with very little room for error."
Limited opportunities	"Poor people face more limited opportunities."
<i>Effectiveness of social assistance programs:</i>	
Recipients waste money	"Poor people will waste the money they receive."
Recipients work less	"Poor people will work less when they receive money."
Tax cuts help more	"Reducing taxes on essential goods will benefit poor people more."
Panel N: Quality-control and survey-feedback variables	
Wants further information	Indicator equal to 1 if the respondent wrote a non-trivial response to "Is there any further information you would like to know about the programmes governments implement to support poor families?" (open-ended text box).
Knowledge-check score (programs)	Number of program-feature pairs the respondent matched correctly across the two pre-treatment knowledge-check tables (out of 10). Respondents scoring below 7 on their first attempt are shown correct answers and asked to retake; those scoring below 7 on both attempts are excluded from the analytical sample by the researchers (not automatically by the survey software).
Knowledge-check score (treatment video)	For respondents in the <i>Conditions faced by the Poor</i> or <i>Effectiveness</i> arms, number of correct options selected on the post-video attention-check (maximum 3 per arm: three factual claims from the video, with "None of the above" as the trap option).
Survey perceived as biased	Indicator equal to 1 if the respondent answers "Yes" to "Do you feel this survey was politically biased?"
Completion time	Time in minutes between consent and the last substantive question. Used to drop respondents in the bottom percentile (speeders).

4.2 Descriptive analysis

In this section, we describe beliefs and attitudes toward poverty reduction programs, examine how they are related, and test our descriptive hypotheses (H1a–H1c and H3a–H3c). These analyses describe correlations. They are not tests of the conceptual framework, and a positive correlation between a belief and support is not evidence that the belief causes support: the two may be related through characteristics we do not observe, or through reasoning that runs in the opposite direction. Because no causal claims are made, no multiple-testing corrections are applied to any analysis in this section.

We report these descriptive results for three reasons. First, the correlations are new in their own right: there is at present no cross-country descriptive evidence of how support for particular poverty reduction programs relates to beliefs about poverty and perceptions about program effectiveness. Second, they identify which beliefs are candidates for the channels through which information might shift support, and the experiment then supplies the causal variation with which to test them. Third, they provide the belief–support slopes used in the consistency check in Section 4.5, which asks whether the treatment effects we measure are consistent with the magnitude implied by the belief shifts the treatments produce; without the descriptive analysis, that comparison could not be made.

We report the correlations with and without adjustment for individual characteristics. Adjustment is not intended to deliver a causal interpretation but enables us to distinguish correlations that are explained by observable characteristics such as age, education or income from correlations that remain holding those characteristics fixed. All standard errors reported and used to construct confidence intervals will be HC1 heteroskedasticity-robust standard errors. Descriptive statistics are presented for the unweighted online sample. Survey-weighted descriptive statistics are reported in an appendix.

4.2.1 Raw means

Following Dechezleprêtre et al. (2025), we will present raw means of key variables in the form of heat-map tables. This is intended to give readers an overview of the commonalities and differences in attitudes toward poverty reduction programs and consumption tax exemptions in different countries. Raw means are based on the control group only (baseline support for poverty reduction programs, and attitudes towards consumption tax exemptions) or on the control group and the part of the treatment group that are treated after being asked the relevant questions (general views on poverty and program-specific effectiveness beliefs).¹⁴ This ensures that they are not contaminated by treatment-induced shifts in beliefs or support.

Baseline support for poverty reduction programs. We will report country means of support for each program type (*somewhat support* or *strongly support*) in a heat-map table with countries as columns and program types as rows. Means by country income group (lower middle income and upper middle income) will be shown as additional columns.

¹⁴We will check whether the different order in which the questions on general views on poverty, program-specific effectiveness beliefs, and support for poverty reduction programs are asked affects how they are answered by the control group.

General views on poverty. We will report country means for the items in Panel D of Table 9—poverty perceptions, raw concern for the poor, causes of poverty, normative views on poverty and inequality, and beliefs about the role of government—in a similar heat-map table with countries as columns and items as rows, with additional columns showing country income group means.

Program-specific effectiveness beliefs. Means for the program-specific effectiveness belief items in Panel F of Table 9—capturing views on waste and corruption, targeting accuracy, beneficiary benefits, and distortions and spillovers (Boxes V–VIII of the conceptual framework)—will likewise be presented in a heat-map table but with *program types* as columns and beliefs as rows, broken down by country income group (see Figure A1 in Appendix B for an example based on data from the small-scale pilots). Results for individual countries will be presented in an appendix).

Attitudes toward consumption tax exemptions. We will report country means of the items in Panel K of Table 9 in a heat-map table with countries (and country income groups) as columns and items as rows. The five components of the tax exemptions index (views on whether exemptions are a good policy, help the poor, are fairer than alternatives, and are progressive, plus the preference for exemptions over targeted social assistance) will be shown alongside the standalone item on perceived susceptibility to corruption, which is excluded from the index because it captures beliefs about administrative integrity rather than distributional incidence.

In addition, we will show raw means of measures that could affect self-interested motivation in an appendix. These will be based on the full sample, as all relevant questions are asked prior to the treatment.

4.2.2 Conditional correlations

To examine H1a–H1c—the roles of self-interest, perceptions of poverty, and effectiveness beliefs—we will use OLS regressions of the form

$$y_{ci} = \alpha_c + d'_{ci}\beta + x'_{ci}\gamma + z'_{ci}\delta + \epsilon_{ci}. \quad (3)$$

where y_{ci} is a primary outcome index for individual i in country c ; α_c is a country fixed effect; d_{ci} is a vector of treatment group indicators; x_{ci} is a vector of individual characteristics from Panel B of Table 9 in our main specification—namely age, gender, ethnicity, urban/rural indicators, marital status, education level, employment status, and support for current government—but we will also estimate versions with no controls; z_{ci} is a vector of standardized variables measuring respondents’ “first-stage” beliefs and perceptions. z_{ci} comprises:

- As a measure of **respondent self-interest**, *Family currently/expects to benefit* (Panel C), standardized to facilitate comparisons across different determinants of support. (We will show robustness to including the *Self-interest index* instead.)
- As measures of **general views on poverty**, the *Perception of poverty*, *Concern for the poor*, *Internal causes of poverty* (poverty attributed to individual factors), *External causes of poverty* (poverty attributed to structural factors), and *Role of government* indices (all Panel E).

- As measures of **Program effectiveness**, the *Waste and corruption*, *Targeting accuracy*, *Benefits-for-beneficiaries* and *distortions and spillovers* indices, separately for each program (all Panel G).¹⁵ Where the outcome is a Relative support index, we will also include respondents' view of beneficiaries' implicit valuation (relative to an unconditional cash transfer) for the three other core program types (Panel J) as additional regressors.

The treatment indicators absorb any post-treatment shifts in outcomes, so the analysis can be conducted on the full sample. The parameters of interest are the δ s, which capture the conditional correlations between the beliefs and perceptions set out in our conceptual framework and the primary outcome variables. In particular, we expect to show:

- For regressions where the **dependent variable is one of the three different versions of the Overall support index** (see section 3.6.2), we will, at a minimum, show the coefficients and corresponding confidence intervals on the measure of respondent self-interest and on the five different indices of general views on poverty. If the coefficient on the measure of respondent self-interest is positive and significantly different from zero, this will support hypothesis H1a; if the coefficient on the *Perception of poverty index* is positive and significantly different from zero, this will support hypothesis H1b. We expect to show the relevant coefficients and confidence intervals in a coefficient plot with different symbols for the three different Overall support indices.
- For regressions where the **dependent variable is a Relative support index for a specific program type**, we will, at a minimum, show the coefficients and corresponding confidence intervals on the Program effectiveness indices (*Waste and corruption*, *Targeting accuracy*, *Benefits-for-beneficiaries* and *distortions and spillovers*) for the same program type (values for all program types are included in all regressions). If the coefficients on these indices are positive and significantly different from zero, this will support hypothesis H1c. We expect to show the relevant coefficients and confidence intervals in a coefficient plot with different symbols for the six different Relative support indices (see Panel A of Figure A2 in Appendix B for an example based on data from the small-scale pilots).

In an appendix, we expect to show the full coefficient vector with standard errors for the regressions described above, as well as from a regression of the same form where the dependent variable is the Tax exemption index. For predictors where the effect is heterogeneous across countries, we additionally show country-specific coefficients as in Dechezleprêtre et al. (2025).

4.2.3 Variance decomposition

We will decompose the variance in each primary outcome index that is attributable to the different “first-stage” beliefs and perceptions using the LMG method (Lindeman et al., 1980; Grömping, 2007). LMG allocates R^2 across predictors in a way that is invariant to predictor ordering, making it well-suited to comparing the relative importance of conceptually distinct predictor groups. The analysis will be conducted on the full sample and include country fixed effects α_c in addition to the vector of “first-stage” beliefs and perceptions z_{ci} .

¹⁵We will also report results for a version with only own-program effectiveness beliefs as a supplementary analysis to address potential concerns about multicollinearity.

Results will be presented as bar charts of the LMG-attributed R^2 shares by outcome; we will group together predictors and/or restrict the chart to the most important ones (see Panel B of Figure A2 in Appendix B for an example based on data from the small-scale pilots). We will report the shares attributable to all predictors individually in an appendix.

4.2.4 Additional descriptive analysis

All results in this section will be reported both for the full sample and separately for the control group only.

Spending support and allocation alignment. To address H3a, we will compare respondents' expressed spending support across the seven hypothetical budget allocations in Panel I of Table 9: an even split across programs, each program in isolation, the respondent's own preferred allocation, and the current government allocation. H3a predicts that spending support is highest when the allocation corresponds to the respondent's own preferred program mix. We test this directly by comparing, within respondents, the spending levels elicited under the preferred allocation versus the even-split and government allocations. The Δ *Spending vs preferred* variable (Panel I) summarizes the within-respondent difference.

Country-level program alignment. To address H3b, we will examine whether support for specific program types is higher in countries where those programs are more widely used. Using country-level data on program prevalence from the World Bank's ASPIRE database, we identify the most prominent program type in each country. We then compare support for that program type across countries, controlling for individual-level characteristics. As a supplementary analysis, we will also compare respondents' preferred level of total spending with actual total spending in a given country. This cross-country comparison is purely descriptive and does not have a causal interpretation.

Differences in support and views by background characteristics. We will investigate how the raw means reported in Section 4.2.1 vary by background characteristics (Panel B of Table 9). We will report key conditional means and use k-means clustering to detect patterns of heterogeneity in baseline support and views.

Perceived beneficiary trade-offs. We will report the distribution of the pairwise comparison outcomes from Panel J of Table 9 separately for each program pair (CCT vs. UCT, in-kind vs. UCT, and public works vs. UCT) and by country. These measures address H3c (do respondents believe beneficiaries require greater compensation to accept programs with conditions or ordeals?). We will also examine how these quantitative measures of perceived benefits for beneficiaries relate to the qualitative measures we collect in the program effectiveness module of the survey; as only the quantitative measures are explicitly from the beneficiaries' point of view, this could be informative about respondents' level of paternalism.

4.3 Main experimental analysis

Our identification strategy relies on the random assignment of individuals to each of the experimental groups. We exploit this random assignment to estimate the causal effects of being assigned to the different treatment groups using linear regression. We will estimate average treatment effects by comparing outcome variables among those who were randomly assigned to each treatment group to those that were randomly assigned to the control group.

4.3.1 Balancing Checks

We will assess baseline balance across treatment arms using two complementary approaches. First, we will report group means (or proportions) for all baseline variables by treatment arm. Second, we will formally test for differences using two-sample t -tests for continuous variables and chi-square tests of independence for categorical variables. All variables showing statistically significant differences in these tests will be noted explicitly.

To evaluate potential attrition bias, we will construct an indicator for completing the primary outcome section and compare baseline characteristics between completers and non-completers using the same methods (t -tests and chi-square tests). We will also test whether completion rates differ by treatment arm. Balance diagnostics (both across treatment arms and between attritors and non-attritors) will be reported alongside summary tables of sample composition.

4.3.2 Empirical model

We use the following empirical specification:

$$y_{ci} = \alpha_c + \mathbf{d}'_{ci}\boldsymbol{\beta} + \mathbf{x}'_{ci}\boldsymbol{\gamma} + \epsilon_{ci} \quad (4)$$

Here, y_{ci} are the primary outcome indices and their components, as well as the *Supports petition* real stakes outcome (Panel L of Table 9) and the *Manipulation check* responses (Panel M of Table 9) for individual i in country c . All covariates are defined as in Section 4.2.2. Including the pre-specified individual characteristics $\mathbf{x}_{ci}$ improves precision without affecting identification.

The coefficients of interest are those included in $\boldsymbol{\beta}$, which capture the average treatment effect of assignment to each treatment. This will be how we evaluate whether our data supports hypotheses H2a, H2b, H2c and H2d. For H2a–H2c the outcomes are the overall support and relative support indices; for H2d the relevant outcome is the tax exemptions index (see Section 3.6.2).

Reporting. The regressions shown in Equation 4 are estimated by OLS with HC1 heteroskedasticity-robust standard errors. Treatment effects on each index and its components are reported in the corresponding tables (for examples of the main tables based on data from the small-scale pilots, see Tables A1 to A4 in Appendix B). Each table separately reports regressions of the outcome on a single pooled treatment indicator (combining the *Conditions faced by the Poor* and *Effectiveness* arms) and regressions of the outcome separately on the two treatment indicators, with the control group as the omitted category. Among the relative support outcomes, we show results for UCT and for UCT and CCT combined (the “cash transfer support index” and its components) in the main tables

(see Table A3 in Appendix B); results for other programs are shown in an appendix.¹⁶ The main results are pooled across countries; we will show results by country-income group and individual countries in an appendix. Control-group means are reported in raw units for the component outcomes and on the standardized scale for the indices.

4.4 Heterogeneous Effects

4.4.1 Prespecified dimensions of heterogeneity

We prespecify four dimensions of heterogeneity motivated by hypotheses H4a–H4d. The first three (H4a–H4c) are each tied to a specific block of the conceptual framework in Figure 2 and apply to overall support for poverty reduction programs.¹⁷ The fourth (H4d) applies specifically to the consumption tax exemption outcomes (Panel K of Table 9) and is motivated by the self-interest channel (Box I of the conceptual framework in Figure 2). For each dimension, we define the relevant baseline moderator using items measured prior to treatment, split the sample at the within-country median of the moderator, and estimate an augmented version of Equation (4):

$$y_{ci} = \alpha_c + \beta_1 \text{Cond}_{ci} + \beta_2 \text{Eff}_{ci} + \beta_3 (\text{Cond}_{ci} \cdot M_{ci}) + \beta_4 (\text{Eff}_{ci} \cdot M_{ci}) + \beta_5 M_{ci} + x'_{ci} \gamma + \varepsilon_{ci}, \quad (5)$$

where Cond_{ci} and Eff_{ci} are the treatment indicators for the *Conditions faced by the Poor* and *Effectiveness* arms, M_{ci} is the prespecified moderator (defined separately for each dimension below), and x_{ci} is the vector of pre-treatment controls used in Equations (3) and (4). Standard errors are heteroskedasticity-robust (HC1). The interaction coefficients β_3 and β_4 are the primary objects of interest. A hypothesis is supported if the relevant coefficient has the predicted sign and differs significantly from zero at the five percent level; we use two-sided p -values, as for the primary hypotheses (Section 4.8.7). For H4b, which concerns both treatment arms, we additionally test $\beta_3 = \beta_4$. These tests are not part of the set of primary treatment–outcome tests defined in Section 4.8.7, and we do not adjust them for multiple testing. They are prespecified and directional, and are reported as secondary throughout. The sample is restricted to respondents who answered the questions about general views on poverty and program-specific effectiveness beliefs *before* the treatments, i.e., for whom genuinely pre-treatment beliefs are observed (see Section 3.3).

Baseline perception of poverty (H4a). H4a predicts that the *Conditions faced by the Poor* video will increase support most strongly among respondents whose pre-treatment perception of the extent and severity of poverty is low. The moderator M_{ci} is an indicator equal to one for respondents in the lower half of the within-country distribution of the *Perception of poverty index* defined in Panel E of Table 9. The coefficient of primary interest for H4a is β_3 in Equation (5). H4a is supported if β_3 is positive and significantly different from zero, indicating that the *Conditions faced by the Poor* treatment raises support more where baseline perception of the extent and severity of poverty was lower. We have no directional prior for β_4 .

¹⁶Showing results separately for UCT and for UCT and CCT combined is partly motivated by the fact that in the small-scale pilots, respondents perceived the two potential downsides of cash transfers that the *Effectiveness* treatment is intended to address—wasteful spending and reduced labour supply by beneficiaries—to be an issue for UCTs but not CCTs (see Figure A1 in Appendix B).

¹⁷That said, in line with hypothesis H2c, we would expect effects of the *effectiveness* treatment to be strongest on support for cash transfers.

Preferred role of government (H4b). H4b predicts that both video treatments will increase support most strongly among respondents who believe that the government has the responsibility to reduce poverty. The moderator M_{ci} is an indicator equal to one for respondents in the upper half of the within-country distribution of the *Role of government index* (Panel E of Table 9). H4b is supported if β_3 and β_4 in Equation (5) are both positive and significantly different from zero. We additionally test $\beta_3 = \beta_4$; rejecting equality would indicate that openness to government action is more of a precondition for persuasion through one informational channel than the other (concern in the case of *Conditions faced by the Poor*; effectiveness in the case of *Effectiveness*).

Baseline beliefs about program effectiveness (H4c). H4c predicts that the *Effectiveness* video will increase support most strongly among respondents who, prior to treatment, were most skeptical about the effectiveness of poverty reduction programs. For each respondent in this subsample we construct a composite *Baseline effectiveness index* by averaging the four program-specific effectiveness indices defined in Panel G of Table 9 (reversing the *Waste and corruption index*, the only one of the four oriented so that higher values indicate less favorable beliefs) across items and across the four core programs (UCT, CCT, in-kind, public works), so that higher values correspond to more favorable baseline beliefs. The moderator M_{ci} is an indicator for being in the lower half of the within-country distribution of this composite index. The coefficient of primary interest for H4c is β_4 . H4c is supported if β_4 is positive and significantly different from zero, indicating that the *Effectiveness* treatment raises support more where baseline effectiveness beliefs were less favorable. We have no directional prior for β_3 . As a complement, we re-estimate Equation (5) using, in turn, each of the two Panel F items that the *Effectiveness* video addresses directly: *Recipients waste help* and *Reduces labor supply*, in both cases for unconditional and conditional cash transfers, the programs to which both claims apply (Table 6). Each item is coded one if the respondent attributes the characteristic to the program, so higher values indicate less favorable beliefs. The moderator M_{ci} is therefore the indicator itself, and a positive β_4 again indicates larger effects among respondents who were initially more skeptical.

Perceived self-interest in consumption tax exemptions (H4d). H4d applies to the consumption tax exemption outcomes (Panel K of Table 9) rather than the primary support outcomes. The moderator M_{ci} is an indicator equal to one for respondents in the upper half of the within-country distribution of the *Exposure to consumption tax (gain)* or *Exposure to consumption tax (loss)* variable (Panel B of Table 9), which captures self-reported financial gain/loss if consumption taxes were cut/raised, relative to the average family, depending on which version was asked (respondents are asked either version of the question with probability $p = 0.5$). We estimate Equation (5) with the consumption tax exemption index as the outcome.¹⁸ H4d is supported if β_4 is positive and significantly different from zero. Because M_{ci} here identifies respondents who expect to be *more* affected, a positive β_4 means a smaller reduction in support for tax exemptions among them, which is equivalent to the larger reduction among the less affected that H4d predicts. We have no directional prior for β_3 .

¹⁸We will also estimate the effect for the subsamples with different phrasing of the question individually in order to check the sensitivity to the phrasing.

Additional prespecified moderators. Beyond the four dimensions tied to H4a–H4d, we will report heterogeneity estimates along seven further prespecified moderators that are motivated by the descriptive hypotheses and by the discussion of the online sample in Section 3.1. These are: (i) self-interest, captured by an indicator equal to one for respondents who currently benefit, or expect to benefit, from any of the listed social assistance programs (Panel C of Table 9); (ii) subjective income position, captured by an indicator equal to one for respondents who place their household in the bottom two subjective income groups; (iii) education, captured by an indicator equal to one for respondents with at least secondary education; (iv) gender; (v) membership of the largest ethnic or religious group in the respondent’s country (*Largest group*, Panel B of Table 9); (vi) membership of the largest ethnic or religious group *among the poorest 20%* of people in the respondent’s country (*Largest group among poor*, Panel B of Table 9); and (vii) support for the current government (*Supports current government*, Panel B). Each moderator is entered separately in Equation (5). These analyses are prespecified but are reported as secondary; we will not adjust them for multiple testing across moderators, and we treat them as informative only in conjunction with the H4a–H4d analyses.

Power and interpretation. As noted in Section 3.7, the median-split heterogeneity tests are well powered for pooled cross-country analysis but only modestly powered within individual countries. We will therefore report pooled estimates of Equation (5) as our primary heterogeneity results, and report country-specific estimates only in appendix. Where median splits are mechanically sensitive to ties or to near-median respondents, we report as a robustness exercise the continuous-moderator analog of Equation (5), in which M_{ci} is replaced by the standardized continuous index.

4.4.2 Ex-post identification of the main dimensions of heterogeneity

We will use Generalized Random Forests (GRF; Wager and Athey, 2018; Athey et al., 2019) to estimate heterogeneous treatment effects after data collection is complete. GRF non-parametrically estimates the conditional average treatment effect (CATE) as a function of pre-treatment covariates, accommodating high-dimensional interactions without requiring a priori specification of which moderators matter most. We will apply GRF separately to each primary outcome index and each treatment arm using the `grf` package in R. The pre-treatment moderators entered into the forest are the same variables included in Equation (4) together with the full set of “first-stage” beliefs and perceptions as defined in Section 4.2.2. The sample is restricted to respondents for whom all moderators are measured prior to treatment.

We will assess whether heterogeneity is statistically significant using the calibration test of Athey et al. (2019), which tests whether the estimated CATE is a better linear predictor of realized outcomes than the average treatment effect alone. We will also report the variable importance scores from each forest to identify which baseline characteristics are the strongest moderators of treatment effects. These results will be reported in an appendix as supplementary analysis, with any particularly salient patterns of heterogeneity discussed in the main text.

4.5 Additional experimental analysis

There are six additional types of experimental analysis that will be conducted. Except for the first and the second, the results will be presented in an appendix.

First, we will measure the effect of the treatments on respondents' "first-stage" beliefs and perceptions. The analysis is identical in form to Equation (4) above, but is restricted to respondents who answered the relevant modules later in the survey (after the treatment for the treatment groups), and the outcomes of interest are the general views on poverty and perceived-effectiveness indices rather than the main support outcomes. As a supplementary analysis, we will also show results where we instead use respondents who answered the relevant modules earlier in the survey as the comparison group (this is a less preferred comparison because the effect of the treatment is potentially confounded by a question-order effect).

Second, we will test H5a by examining the effect of the three different ways of closing the budget constraint (higher taxes, lower government spending, higher foreign aid) on the level of spending on poverty reduction programs respondents support. For this purpose, we estimate a version of Equation (4) with indicators added to the treatment vector for which hypothetical way of closing the budget constraint was presented to respondents. The dependent variable is the quantitative overall support index, as it includes precisely those questions where the hypothetical budget constraint was presented. H5a is supported if preferred spending is significantly higher under the foreign-aid assumption than under each of the other two. We will additionally report whether the difference between the higher-taxes and lower-other-spending assumptions varies with perceived exposure to consumption taxes or with subjective income group (Panel C of Table 9), again tested through an interaction. However, this comparison will likely be underpowered.

Third, we will examine the within-survey randomization of the comparator program used in the final consumption tax exemption module. In this section of the survey, respondents are presented with a comparison between a reduction in consumption taxes on essential goods and a randomly assigned comparator program—either an unconditional cash transfer or a public works program. This randomization enables us to test whether respondents' relative preference for consumption tax exemptions over targeted social assistance depends on the form of targeted assistance to which they are compared. The analysis follows Equation (4) above, restricted to the consumption tax outcomes, with the comparator program (UCT or public works) used as additional treatment indicators. If respondents express greater support for tax exemptions when the comparator is an unconditional cash transfer than when it is a public works program, this would suggest that opposition to less restrictive forms of targeted assistance is an important driver of support for tax exemptions. Conversely, similar levels of support across both comparators would indicate that support for tax exemptions reflects underlying preferences for universal or less administratively complex forms of support, rather than a specific objection to particular targeted programs. We will also examine whether the main treatment effects on tax-exemption support are moderated by the assigned comparator, by interacting the comparator indicator with the main treatment indicators.

Fourth, we investigate whether treatment effects differ across countries and whether there are systematic differences between groups of countries. The first step will be to reproduce the analysis

in Equation (4) above for each of the eight countries (i.e., without pooling data across countries and without country fixed effects), so that country-specific point estimates and confidence intervals can be compared directly. The second step will be to reproduce the analysis in Equation (4) pooling subsets of countries defined by country-level characteristics (e.g., separating countries into those with high versus low social assistance spending as a share of GDP, those that are classified as lower-middle versus upper-middle income, or those that predominantly fund cash transfers versus other program types). The full country-by-country results and the country-subset analyses will be reported in an appendix. The country-specific estimates from the first step will also be the inputs to the hierarchical analysis described in Section 4.6.

Fifth, we will conduct a “reverse IV” consistency check to assess whether the estimated treatment effects on support are consistent with the proposed belief-updating mechanism (Dechezleprêtre et al., 2025). We compute “belief-pathway predicted effects” by multiplying the first-stage belief shifts from the first additional analysis above by the corresponding belief–support slopes from Section 4.2.2 and summing over all belief dimensions. These predicted effects are compared with the directly estimated reduced-form treatment effects on support. Close agreement between the two quantities supports the interpretation that the treatments operate through the belief-updating channel; a large divergence would suggest either that the channel is misspecified or that the treatments affect support through pathways not captured by the measured beliefs. This analysis is purely descriptive and does not constitute an additional hypothesis test; it will be presented in an appendix and discussed only briefly in the paper.

Sixth, we will test directly whether eliciting beliefs before treatment changes the treatment effects we estimate. As described in Section 3.8, the placement of the *Views on poverty* and *Beliefs about program effectiveness* modules relative to the treatment stage is randomized with probability one half, independently of treatment assignment. We therefore estimate a specification identical in form to Equation (5), with an indicator for pre-treatment placement in place of the moderator, taking the primary support indices as outcomes. Consistency bias, whereby respondents hold to positions consistent with views they have already stated, would predict a smaller treatment effect for those who see the beliefs and perceptions modules before the treatment. In contrast, experimenter demand effects might lead to higher estimated treatment effects for this group if prior elicitation makes the purpose of the videos more transparent. We note that this test is powered to detect differences in treatment effects of the magnitudes we regard as substantively meaningful, but not to rule out small ones.

4.6 Cross-country generalizability

The pooled specification in Equation (4) estimates a precision-weighted average of the eight country-specific treatment effects. This is the appropriate estimand for the primary hypotheses, which ask whether the treatments shift support on average across the countries we study, and it remains valid if the effects differ across countries. It is silent, however, on *how much* they differ, and so provides no basis for statements about countries we do not survey. We therefore supplement it with a Bayesian hierarchical analysis following Rubin (1981). Richer versions of this approach have been used to aggregate evidence across multi-site experiments in development economics by Meager (2019, 2022).

Let $\hat{\beta}_c$ denote the estimated treatment effect in country c , obtained by estimating Equation (4) separately by country as described in Section 4.5, and let σ_c denote its standard error. We model

$$\hat{\beta}_c \sim \mathcal{N}(\beta_c, \sigma_c^2), \quad \beta_c \sim \mathcal{N}(\mu, \tau^2), \quad (6)$$

where μ and τ are the mean and standard deviation, respectively, of the distribution of true country-specific treatment effects.

This structure lets the true treatment effects vary across countries while still allowing estimates from one country to contain information about treatment effects in another ('partial pooling'). The key parameter of interest is τ : the smaller τ , the stronger would be the external validity of our estimated treatment effects. Two polar cases are that true treatment effects are the same across countries ($\tau = 0$), so all cross-country variation in estimated treatment effects arises purely from sampling noise; and that treatment effects in one country are completely uninformative about treatment effects in another ($\tau \rightarrow \infty$).

The central aim of the analysis is to construct a 95% prediction interval for the true treatment effect in a hypothetical additional country, which would account for the underlying standard deviation of the treatment effect τ as well as uncertainty about both μ and τ . We regard this interval as the most direct summary of what the study implies for generalization and will report it alongside the pooled treatment effects. It should be read as applying to a country *exchangeable* with the eight we study, which need not be true for middle-income countries in general. We will also report 'partially-pooled' country-specific estimates implied by the same model, which are shrunk toward the posterior mean of μ .

Estimation and the prior on τ . We estimate Equation (6) in a Bayesian framework. With eight sites, point estimates of τ will often take the value zero exactly even if positive values also have substantial likelihood. We place independent uniform priors on μ and τ , as in Rubin (1981). This is a conservative choice: the uniform prior admits values of τ we would regard as implausibly large, and therefore yields a wider prediction interval than a substantively motivated prior would. We will therefore also report sensitivity to a half-normal prior on τ (Gelman, 2006), with scale set so that the prior places 1% of its mass above one standard deviation of the outcome variable, a between-country standard deviation of true treatment effects we regard as implausibly large. The posterior is computed by numerical integration over τ , with μ and the β_c marginalized analytically. We treat σ_c as known, which is reasonable with approximately 2,000 respondents per country.

Relation to the pooled estimates. We will not replace the pooled estimates for the treatment effects with the posterior means of μ , which will in general be in between the pooled estimate and an equally weighted average. The pooled specification delivers a precision-weighted average of the country-specific treatment effects and remains valid even if the true effect differs across countries. Our primary hypothesis tests rely on the pooled estimates, while the hierarchical model is added to support statements about generalization and to discipline the interpretation of the country-specific estimates.

4.7 Analysis of Open-Ended Responses

In addition to the quantitative outcomes and mechanisms, the survey includes several open-ended questions designed to capture respondents' reasoning and reflections in their own words. These items include brief prompts about why respondents selected the program they did for a top choice and why they selected the portfolio composition that they did. We also include a final question asking respondents for their feedback on the survey.

We plan to analyze these qualitative responses descriptively to complement the quantitative findings and to contextualize treatment effects. Responses will be processed into keywords using LLM-based text analysis with manual validation. Keyness tests will be conducted to measure the relative frequency of keywords between subgroups of respondents (Geiecke and Jaravel, 2024; Haaland et al., 2024b). For example, we will examine differences in the reasons respondents provided between the treatment and control groups. We will also seek to identify salient patterns in reasoning and narrative framing that help explain heterogeneity in quantitative responses.

4.8 Robustness

Listed below are a series of robustness checks that will be conducted to examine the strength of the results.

4.8.1 Alternative specifications

As discussed above our main analysis will be based on OLS regressions, but we will also examine whether using alternative statistical models would have a material impact on our findings. Specifically, we intend to rerun our main analysis with a logit or ordered logit regression. Based on previous studies we do not anticipate either of these alternative statistical models will substantively change our results. All of this additional analysis will be included in an appendix of the Stage 2 Report.

4.8.2 Data quality checks

It is possible that some respondents may not devote their full attention to the survey, and to detect whether this impacts our results we will conduct four main robustness checks. Firstly, we will rerun our analysis after excluding respondents that took too little or too much time to complete the survey. Specifically, we will exclude respondents in the bottom/top 1, 2.5 and 5 percent of the completion time distribution. Secondly, we follow Alesina et al. (2018) and include an explicit data quality check in the form of a knowledge quiz after the neutral explainer video (see Section 3.3). We exclude all respondents that do not have a passing mark in the second attempt.¹⁹ Thirdly, we will conduct an implicit careless-response check based on straight-lining in the qualitative support module. A respondent is flagged as a straight-liner if they select the same response option for all five program-support items in the qualitative support module (i.e., the inter-item standard deviation across the five Likert responses equals zero). These items span meaningfully different programs (unconditional cash transfer, conditional cash transfer, in-kind transfer, public works,

¹⁹As this exclusion occurs for all respondents and before treatment assignment, it does not impact the internal validity of our experiment.

and either an asset transfer or a personalized tax refund), so a plausible and attentive respondent who genuinely holds similar views of every program type would be unlikely to be excluded by this criterion. We will rerun the main analysis excluding straight-liners and report the proportion excluded by country and treatment arm.

Fourthly, we will test directly whether responses deteriorate with survey position, using the survey's built-in position randomizations. Half of respondents answer the *Views on poverty* and *Beliefs about program effectiveness* modules near the start of the survey and half near the end (Section 3.3); within the control group, who see no treatment video under either placement, this is a pure randomization of the position of identical items. Between the two placements in the control group we will compare item means, item nonresponse and 'don't know' rates, straight-lining rates, inter-item standard deviations, and per-module response times. The randomized order of the seven budget-allocation questions, and of items within questions without a natural order, provides further position variation. The primary support outcomes are elicited at the same position for all respondents, so the primary hypothesis tests are unaffected; these comparisons inform the interpretation of the belief-level and heterogeneity analyses and will be reported in the Stage 2 appendix.

4.8.3 Weighting

All main results are presented for the unweighted online sample and are therefore sample average treatment effects for the achieved sample (Section 3.1). As a robustness exercise, survey-weighted results are reported in an appendix, adjusting the sample toward the adult population as a whole. Weights are constructed by inverse probability weighting: within each country, respondents are grouped into cells defined by sex interacted with age group (18–24, 25–34, 35–44, 45–54, 55 and over), and each respondent's weight equals the ratio of the cell's share in the adult population (from the most recent census or comparable national statistics) to the cell's share in our achieved sample—that is, inverse probability weighting with fully saturated cells. If there are representativeness concerns regarding other demographic characteristics (for example, urban/rural status or education), for which joint population distributions are typically unavailable, weights will be computed by raking (iterative proportional fitting) to the marginal distributions. Survey weights will be trimmed to the interval $[0.25, 4.0]$, following Dechezleprêtre et al. (2025), so that no single respondent receives a weight more than four times the minimum; this prevents extreme weights from unduly distorting the reweighted estimates. We will also report the results in an appendix using untrimmed weights.

4.8.4 Experimenter demand effects

A potential concern in any survey experiment, particularly one providing information from an authoritative source, is that respondents may infer the researcher's expected response and adjust their answers accordingly (De Quidt et al., 2018; Mummolo and Peterson, 2019). We therefore conduct several complementary tests to assess the extent to which experimenter demand effects could plausibly account for our findings.

Direct test using bias perceptions. We will examine whether the treatments increase the probability that respondents report perceiving the survey as biased or as pushing them toward a partic-

ular answer. A treatment-induced increase in perceived bias would be consistent with respondents detecting an experimenter-preferred response. We will also implement the De Quidt et al. (2018) logic by examining whether estimated treatment effects are systematically larger among respondents who perceive the survey as biased relative to those who do not. If demand effects are an important driver of our results, treatment effects should be concentrated among the former group; if they are not, treatment effects should be similar across the two groups.

Comparison across outcomes with differential susceptibility to demand. Our primary outcomes vary in the extent to which they are susceptible to experimenter demand. Stated Likert-scale support items are the most susceptible, since they involve no cost or commitment from the respondent and are the most transparent indicators of the researcher’s intended outcome. The budget-allocation tasks impose internal trade-offs (more spending on one program means less on another) and are therefore intermediate in susceptibility. The real-stakes outcome—willingness to sign a petition highlighting the benefits of the respondent’s preferred poverty reduction program—is the least susceptible, because it involves a more concrete commitment than a Likert item and cannot be answered to please the experimenter without consequence. If treatment effects are concentrated in the Likert-scale outcomes and disappear in the real-stakes outcome, this would suggest that demand effects may play a role. If treatment effects are broadly consistent across the three types of outcomes, this would provide evidence that the estimates capture genuine preference change rather than experimenter-induced compliance.

Qualitative analysis of open-ended feedback. The final survey question invites respondents to provide feedback on the survey. We will conduct a structured analysis of these responses, searching for terms and themes related to perceived bias, pressure, advocacy, or distrust of the survey’s framing (e.g., “biased”, “one-sided”, “trying to convince”, “leading questions”). We will compare the prevalence of such themes across treatment and control groups, and across treatment arms. Treatment-induced increases in references to perceived bias could suggest that the treatments are detected as advocating for a particular position.

Heterogeneity that is inconsistent with demand effects. Finally, we note that several patterns of heterogeneous effects, if found, would be difficult to reconcile with a pure demand-effect interpretation. In particular, hypotheses H4a–H4c predict that treatment effects will be concentrated among respondents with specific prior beliefs (e.g., low baseline perception of the severity of poverty for the *Conditions faced by the Poor* video, or low baseline beliefs about program effectiveness for the *Effectiveness* video). If treatment effects vary systematically with these priors in the predicted directions, this would be more consistent with genuine belief and preference updating than with uniform compliance with experimenter expectations.

Collectively, these complementary tests will help us assess the contribution of experimenter demand effects to our estimated treatment effects. If multiple lines of evidence point in the same direction, we will interpret our findings as more likely to reflect genuine preference change.

4.8.5 Attrition

In this study, “attrition” refers to post-randomization non-completion of the outcome section (“Support for spending on poverty reduction programs”, see Section 3.3). Screen-outs during eligibility checks and exits before randomization are not counted as attrition. Item non-response within completed surveys (i.e., skipping one or more outcome items) is tracked separately. We expect this to be minimal or zero, as we have set up the survey to require a response to each question.

In the second and third small-scale pilots, post-randomization attrition was 12%, with no sign of differential attrition across treatment arms. Item non-response was minimal and only possible due to minor technical issues. These results suggest that attrition and item non-response are unlikely to bias main-wave estimates. Statistical power is unaffected by attrition because we will recruit additional respondents until the target for survey completions is reached (see Section 3.10).

Design features to minimize attrition include (i) short, mobile-friendly videos, (ii) limited open-ended responses, and (iii) transparent compensation aligned with completion. We will monitor completion rates by country and treatment arm in real time and report overall and differential attrition in the analysis.

In the event that differential attrition occurs during the survey experiment, we will produce Lee bounds estimates of our main results (Lee, 2009). The tightness of these bounds will be informative about whether differential attrition is an issue of concern. This is the same approach that has been taken by other studies in this literature (e.g. Kuziemko et al., 2015; Hoy and Mager, 2021).

4.8.6 Handling outliers

The susceptibility of our outcomes to extreme values varies considerably across the four primary outcome families. Most outcomes are bounded by construction and therefore not affected by outliers: the qualitative support items are five-point Likert scales bounded between 1 and 5; the relative support measures are bounded between 0 and 100 (budget shares) or 0 and 1 (top-choice dummies); and most of the program indifference and tax exemption outcomes are binary indicators. Outliers are only a substantive concern for the continuous outcomes, namely the quantitative support measures (preferred social assistance spending in local currency under each hypothetical budget allocation) and the monetary premium amounts elicited in the program indifference module (Panel J of Table 9).

For these continuous outcomes, we will winsorize all monetary outcomes at the 99th percentile within each country and main treatment group, replacing values above this threshold with the 99th-percentile value. We winsorize rather than trim because the upper tail of stated spending preferences is informative about the strength of preferences and we do not want to discard this information entirely. Country and treatment-group-specific winsorization is used to avoid biasing comparisons across countries and treatment groups. As a robustness check, we will rerun our main analyses with continuous outcomes winsorized at the 95th percentile and with the unwinsorized data, and we will report any cases where conclusions are sensitive to the winsorization threshold. These additional results will be reported in the appendix.

For the index outcomes, the standardization and averaging procedure described in Section 3.6.2 provides a degree of robustness to outliers in any individual component. Because each index is the unweighted average of multiple component z-scores, a single extreme value in one component has bounded influence on the index. Nonetheless, we will apply the winsorization rules described above to the underlying continuous components before standardization, so that the indices and their components are constructed consistently.

4.8.7 Multiple outcome and multiple hypothesis testing

Testing treatment effects across multiple outcome families simultaneously raises the risk of spurious discoveries. Although we have a very small number of pre-specified primary outcome families, we nonetheless address this concern explicitly. Multiple-testing adjustments will be treated as a robustness exercise rather than the primary inference criterion, given that each family is motivated by a distinct theoretical channel and the tests are not exchangeable. We will apply the Benjamini–Hochberg false discovery rate (FDR) correction (Benjamini and Hochberg, 1995), which controls the expected proportion of false discoveries among rejected hypotheses. This approach is less conservative than familywise error rate (FWER) procedures and is well suited to a setting with multiple pre-specified outcome families where some true effects are anticipated. We follow Dechezleprêtre et al. (2025) in adopting this approach.

We interpret the pre-specified primary tests using unadjusted two-sided p-values, and report Benjamini–Hochberg FDR-adjusted q-values across the set of primary treatment–outcome tests. Results are considered confirmatory when the direction, magnitude, and unadjusted inference align with the pre-specified hypothesis; FDR-adjusted results are used to assess robustness to multiple testing.

4.8.8 Non-parametric bootstrap standard errors

As a complement to the HC1 heteroskedasticity-robust standard errors used throughout, we will report non-parametric bootstrap standard errors for the primary hypothesis tests (H2a–H2d) and for the treatment effects on the belief indices in Panels E and G of Table 9. For each of 2,000 bootstrap replications, we resample respondents with replacement within each country, holding fixed the country sample sizes, and reconstruct the relevant outcome indices using exactly the same standardization and aggregation procedure as in the main analysis (Section 3.6.2). We then re-estimate the corresponding treatment effects using the same regression specifications. The bootstrap standard error for each treatment-effect estimate is calculated as the standard deviation of the estimated treatment effects across the 2,000 bootstrap samples. The resulting bootstrap standard errors will be reported in an appendix table alongside the HC1 standard errors for the same estimates, and the main text will note any case in which they differ materially. Because the empirical distribution of the observed outcomes is resampled directly, this procedure naturally preserves their discreteness and concentration and does not require the indices to have continuous support. Reconstructing the indices within each bootstrap replication also incorporates sampling variation arising from their standardization and aggregation.

4.9 Expert prediction survey

Alongside the main empirical analysis, we will collect forecasts of our primary results from a sample of researchers and policy experts using the Social Science Prediction Platform (SSPP; DellaVigna et al., 2019). Collecting expert priors before starting to disseminate the results serves three purposes that complement the pre-registered hypotheses and estimation strategy described above. First, the experts provide an external benchmark against which to assess which of our results are genuinely *informative*, in the sense of departing from what an informed prior would have expected (DellaVigna and Pope, 2018; DellaVigna et al., 2019). Second, they discipline our own interpretation of the results as broadly matching expert priors or diverging sharply from priors. Third, the expert predictions are of substantive interest for the political-economy questions motivating the study, as they reveal whether researchers and practitioners working on social protection in middle-income countries hold accurate beliefs about public attitudes toward poverty reduction programs.

Expert sample. We will target three groups of experts. The first is academic researchers working on development economics, public economics, or the political economy of redistribution. The second is policy practitioners at multilateral organizations (e.g. the World Bank, IMF, UN agencies), bilateral donors, and relevant government ministries in the study countries. The third is researchers and analysts at think tanks and NGOs working on social protection. We will recruit experts through the SSPP standing pool of registered forecasters, supplemented by targeted outreach via professional networks, mailing lists, and direct invitations to authors of recent publications on social protection in middle-income countries. We aim for at least 100 completed expert responses, in line with sample sizes in recent applications (e.g. DellaVigna et al., 2019; Vivalt and Coville, 2023). We will report the number of experts, their professional category, their country of work, and their self-reported familiarity with the relevant literature.

Materials shown to experts. Experts will see a short briefing document. This will include: (i) the research questions and the conceptual framework summarized in Figure 2; (ii) the sampling strategy, including the countries surveyed and the online nature of the sample; (iii) the structure of the survey instrument, including the explainer video, the two video treatments, and the main outcome modules; and (iv) the text of the two video treatments and the wording of the primary outcome questions.

Outcomes to be forecast. We will elicit forecasts of the two treatment-specific effects on the following primary outcomes: (i) overall support for poverty reduction programs; (ii) support for each program type on a Likert scale; (iii) the ranking of program types from most to least preferred; and (iv) beliefs about whether tax exemptions are a good policy, help the poor, and are fairer than alternatives, as well as preferences for tax exemptions relative to UCTs and public works programs. We will also collect forecasts for our descriptive analysis on the extent to which support is shaped by self-interest, general views on poverty and the government, program-specific effectiveness, and respondents' perceptions of beneficiaries' relative willingness to pay for programs. To keep the elicitation burden manageable, we do not request forecasts of country-specific effects or of heterogeneous effects, though we will invite open-ended comments on dimensions along which

forecasters expect particularly large heterogeneity.

Use of forecasts in interpretation. In the Stage 2 report, we will present a side-by-side comparison with the realized treatment-specific effects for each of the outcomes listed above. For each outcome, we will report the mean and median forecast, and the share who correctly predicted the sign of the effect. The full set of forecasts, together with the comparison to realised effects, will be reported in an appendix.

5 Interpreting Results

The results of this study could fall into several broad categories. In this section, we describe how we will interpret each type of finding. This is intended to clarify the evidentiary value of the study regardless of whether the main experimental hypotheses are confirmed.

5.1 Strong baseline support for all or some poverty reduction programs

One possible finding is that respondents express high baseline support for poverty reduction programs, even in the absence of information. This would be most directly relevant to hypotheses H1a–H1c on the descriptive determinants of support. High baseline support would suggest that public opposition is less of a constraint on poverty reduction spending than policymakers often assume. If baseline support is high across countries and program types, the main contribution of the study would be to show that the political economy constraint is unlikely to be general opposition to helping poor households.

If support is high for cash transfers in particular, this would challenge the common perception that the public strongly prefers more restrictive forms of assistance, such as in-kind transfers or public works programs. Such a finding would imply that governments may have more political space than expected to expand less restrictive forms of social assistance. Conversely, if respondents strongly prefer in-kind transfers, public works programs, or other more restrictive programs, this would provide direct evidence that public preferences may help explain limited political support for cash transfers and point to a trade-off between the restrictiveness and the amount of support provided. Patterns consistent with H1a—in which support for less restrictive programs is concentrated among current and prospective beneficiaries—would be consistent with political constraints on cash transfer adoption resting partly on the preferences of non-beneficiary voters.

5.2 Low baseline support for all poverty reduction programs

A second possible finding is that respondents express low baseline support for all poverty reduction programs. This would suggest that public attitudes may constitute an important political constraint on expanding social assistance in middle-income countries. The interpretation of this result will depend on which components of the conceptual framework are most strongly associated with support, and will draw directly on the descriptive evidence from H1a–H1c.

If low support is concentrated among respondents who do not expect to benefit personally (consistent with H1a), self-interest is the leading candidate channel. If it is instead concentrated among

respondents who perceive poverty as not very severe (relevant to H1b), express limited concern for poor households, or believe poverty is primarily caused by lack of effort, distributive values and beliefs about deservingness would be the leading candidates. If low support is most strongly associated with concerns about waste, corruption, weak targeting, or negative behavioral responses (relevant to H1c), beliefs about program effectiveness would be. In each case the descriptive pattern identifies a candidate rather than establishing a channel: the treatments provide the causal variation, and the consistency check in Section 4.5 asks whether the effects we measure are of the size implied by the belief shifts the treatments produce.

5.3 Information increases overall support for poverty reduction programs

A third possible finding is that one or more of the video treatments increases overall support for poverty reduction programs, consistent with hypotheses H2a and H2b. If the *Conditions faced by the Poor* video increases support (H2a), this would suggest that the perception of poverty is an important margin for persuasion. In this case, respondents may not initially appreciate the severity of poverty, the limited opportunities available to poor households, or the extent to which poor households face little room for error. If the *Effectiveness* video increases support (H2b), this would suggest that beliefs about program effectiveness are an important barrier to support. In particular, it would indicate that concerns about misuse of transfers, labor supply reductions, or the distributional incidence of tax exemptions are at least partly responsive to accurate information.

However, a positive effect would not by itself establish that baseline opposition rests on incomplete or inaccurate information, since a video can shift stated support by making considerations a respondent already accepted more salient as well as by correcting a belief that was mistaken. Four patterns bear on which account is correct. None is decisive on its own, but they point in different directions under the two interpretations.

The first is whether the beliefs the videos assert move, and whether the movement is confined to them. The informational account requires that the targeted beliefs move: support that rises while they are unchanged would be difficult to attribute to the content of the videos. It further predicts that the movement is specific, leaving beliefs the videos do not address unaffected, whereas a general increase in attention to poverty would move beliefs across the battery, including ones neither video mentions. The manipulation checks alone cannot distinguish these cases, since all six of their items are claims the videos assert. The comparison instead rests on the first additional analysis in Section 4.5, which estimates treatment effects on the full set of beliefs in Panels D–G of Table 9.

The second is whether the size of the effect on support is accounted for by the belief movement. The reverse-IV consistency check in Section 4.5 compares the reduced-form treatment effect with the effect implied by the first-stage belief shifts and the belief–support slopes. Under the informational account, the two should be of similar magnitude. An effect on support substantially larger than the belief pathway implies would indicate that something beyond the measured belief change was doing the work, of which increased salience is one candidate.

The third is how the effect varies with prior beliefs. Correcting a belief requires that the belief was

mistaken, so the informational account predicts larger effects among respondents whose priors were least accurate, which is what H4c tests for the *Effectiveness* treatment. Saliency predicts no such gradient and may predict the reverse, since a consideration can be made more prominent only for a respondent who already accepts it. A profile that is flat, or that runs the other way, would weigh against the informational reading.

The fourth is whether the two treatments differ from one another. The videos assert different things and are predicted to move different beliefs and different outcomes (Tables 5 and 6): *Conditions faced by the Poor* to raise support without reallocating it across programs, and *Effectiveness* to shift its composition. A general increase in the saliency of poverty would move the outcomes in common, so differentiated effects across the two arms would indicate that the specific content of each is responsible.

5.4 Information changes the types of programs respondents support

A fourth possible finding is that information changes the allocation of support across program types. This corresponds to hypothesis H2c and would be particularly important for policy. Governments not only decide how much to spend on poverty reduction; they also choose how to allocate support across cash transfers, in-kind transfers, public works programs, and potentially asset transfers or personalized tax refunds.

If the *Effectiveness* video increases relative support for unconditional or conditional cash transfers (consistent with H2c), we will interpret this as evidence that opposition to less restrictive programs is partly driven by incorrect beliefs about their effects. For example, if respondents become more supportive of cash after learning that cash transfers do not generally increase harmful spending or reduce labor supply, this would suggest that information about program effectiveness can shift support toward less restrictive forms of social assistance.

If information increases overall support but does not change relative support across program types, this would suggest that respondents become more favorable toward poverty reduction in general, but that their views about program design are more stable. Conversely, if information shifts relative support without changing total desired spending, this would suggest that information affects preferences over the composition of social assistance spending more than overall preferences for redistribution.

5.5 No meaningful impacts from the treatments

It is also possible that the treatments will have no meaningful effect on either overall support for poverty reduction programs or relative support across program types. This would be informative. A null result would suggest that short, accurate information interventions are unlikely to substantially shift attitudes in this context.

There are several possible interpretations for such a result. First, respondents may already hold accurate beliefs, leaving little scope for information to change their views. Second, respondents may update their beliefs but not their policy preferences, implying that attitudes toward poverty reduction are shaped more by ideology, self-interest, distrust, or deeper values than by factual be-

liefs. Third, the treatments may be too light-touch to affect preferences, especially if respondents have strong prior views or if policy preferences require more sustained deliberation to change. We will distinguish between these interpretations by examining whether the treatments affect intermediate beliefs and whether baseline beliefs are already close to the information provided in the treatments. A pattern in which first-stage effects on beliefs are large but downstream effects on policy preferences are not would itself be an informative finding, suggesting that the gap between belief updating and preference change is a binding constraint.

5.6 Evidence of backlash or reduced support

Another possible result is that one or more treatments reduces support for poverty reduction programs. This could occur if information makes the costs of these programs more salient, reinforces concerns about targeting or corruption, or leads respondents to view poverty reduction as more difficult than they had previously believed. Such a finding would be important, because it would show that information provision does not mechanically increase support for social assistance.

If the *Conditions faced by the Poor* video reduces support, we will examine whether respondents react by becoming more pessimistic about the effectiveness of government action. If the *Effectiveness* video reduces support, we will examine whether correcting some misconceptions makes other concerns more salient. For example, respondents may accept that cash transfers do not reduce labor supply but still oppose them because of concerns about financing, targeting, or government capacity. Any such result would be interpreted cautiously and in conjunction with the mechanism outcomes.

5.7 Treatment effects exist but are inconsistent with the hypotheses

A further possibility is that treatment effects exist but do not match the patterns specified in our pre-registered hypotheses. We anticipate two broad scenarios in which this could occur, each with distinct interpretations.

Heterogeneous responses across subgroups. The hypotheses are framed in terms of average treatment effects, but a treatment may resonate only with specific subgroups, with offsetting effects elsewhere in the sample. For instance, the *Effectiveness* video could increase support for cash transfers among respondents who initially held inaccurate beliefs about labor supply and recipient behavior, while reducing support among respondents who interpret the corrective information as politically motivated or who become more skeptical of the government's capacity to administer transfers. Similarly, the *Conditions faced by the Poor* video might increase the perceived severity of poverty and thus support for poverty reduction programs among non-beneficiaries while having little or even negative effects on respondents who already perceive themselves as economically vulnerable. The pre-specified heterogeneity analyses (H4a–H4d) are designed to detect such patterns, and where average effects are null or inconsistent with the hypotheses, we will pay close attention to subgroup-level effects identified through the generalized random forest approach described in Section 4.4.

Differential effects across outcomes. The treatments may affect some outcome families but not others. Effects on Likert-scale support measures may not translate into changes in budget allocations, or vice versa. Effects on stated preferences may not appear in the petition outcome from the real-stakes module. Such patterns would themselves be informative about the depth of preference change and the extent to which information shifts policy preferences as opposed to merely shifting how respondents express those preferences in low-stakes settings.

5.8 Cross-country heterogeneity

The study may also find substantial heterogeneity across countries. Such heterogeneity would not necessarily weaken the study. If treatment effects or baseline preferences differ across countries, we will interpret these differences in light of the existing policy environment, the composition of social assistance spending, and baseline beliefs about government capacity and program effectiveness. Where appropriate, we will pool data only across subsets of countries that share relevant characteristics (e.g., upper-middle versus lower-middle income, or countries with high versus low social assistance spending). However, country-specific estimates will generally be less precisely estimated than pooled estimates. We will therefore treat pooled estimates as the primary evidence on average effects across the study countries, while using country-specific estimates to understand external validity and possible contextual mechanisms.

We will use the hierarchical model set out in Section 4.6 to understand what our results imply about the underlying heterogeneity in true treatment effects across countries. The indicator we focus on is the 95% prediction interval for the true treatment effect in a hypothetical additional country. A narrow predictive interval would indicate that comparable countries are likely to have similar treatment effects. A wide interval would indicate that the average effect may not be a good guide to any particular country and would direct attention to the country-level characteristics that might account for the variation. However, given the small sample of eight countries, we expect prediction intervals to be relatively wide even where the underlying variation is small, because of limited statistical power.

5.9 Interpreting perceived beneficiary trade-offs

The program-indifference outcomes will be interpreted as respondents' beliefs about how poor households value different program types relative to unconditional cash transfers. These measures do not directly estimate beneficiaries' own willingness to pay. Rather, they capture public perceptions of beneficiary preferences. This distinction is important. If respondents believe poor households would require a large monetary premium to accept in-kind transfers, public works programs, or conditional transfers instead of cash, this would suggest that respondents recognize the non-monetary costs of restrictions, ordeals, or conditionality. If respondents believe beneficiaries would prefer restrictive programs even at equal monetary value, this would suggest that the public perceives non-cash features as valuable to beneficiaries, not merely as politically appealing to non-beneficiaries.

5.10 Interpreting support for consumption tax exemptions

Results on consumption tax exemptions will be interpreted separately from results on targeted poverty reduction programs. Consumption tax exemptions are often justified as pro-poor policies, but they differ from social assistance programs because they are broad-based, (much) less targeted, and often less visible as government transfers. If respondents strongly support tax exemptions, this may indicate a preference for universal or less administratively complex forms of support, even if these policies are less progressive. If information reduces support for tax exemptions relative to cash transfers or public works programs, this would suggest that respondents are responsive to evidence about incidence and targeting.

5.11 Overall interpretation

Overall, the study will be informative under a wide range of possible results. If information changes support, the results will identify which beliefs are most malleable and which communication strategies are most effective. If information does not change support, the results will indicate that attitudes toward poverty reduction programs are relatively stable in response to short-run information provision. If baseline support differs sharply across program types or countries, the results will help explain why governments may prioritize some poverty reduction instruments over others. In all cases, the findings will be interpreted as evidence on stated preferences among internet-using adults in middle-income countries, rather than as direct evidence on voting behavior, long-run persuasion, or the preferences of the full adult population.

6 Limitations and potential challenges

This study is designed to provide credible experimental evidence on how information shapes public support for poverty reduction programs. Nevertheless, several limitations and implementation challenges should be considered when interpreting the results.

First, the study will be conducted online, and respondents are recruited from an online panel rather than randomly sampled. We therefore do not describe the sample as representative of a national population, though we expect it to be broadly representative of the internet-using adult population (at least in terms of age and gender). As discussed in Section 3.1, internet-using adults are a substantive population of interest in their own right. Nevertheless we will report detailed comparisons between our sample and national benchmarks, and report survey-weighted results in an appendix. The interpretation of the results will emphasize that the study identifies attitudes among the online population, while being cautious about claims regarding the full adult population.

Second, cross-country comparability is an important challenge. Poverty reduction programs differ substantially across the eight countries in their design, visibility, scale, and political salience. Respondents may therefore interpret terms such as cash transfers, public works programs, in-kind transfers, asset transfers, and consumption tax exemptions through the lens of existing programs in their own country. This is both a feature and a limitation of the design. It allows us to study how attitudes vary across policy environments, but it also means that identical survey wording may not be interpreted identically in all contexts. To reduce this concern, all respondents will re-

ceive a neutral explainer video describing the core features of each program type before the main outcomes are elicited. In the analysis, we will report pooled estimates with country fixed effects, but also show country-specific estimates to assess whether the main patterns are driven by particular settings.

Third, some elements of the survey require respondents to evaluate relatively complex policy trade-offs. These include budget-allocation exercises, comparisons across program types, and questions about how beneficiaries would value non-monetary features such as conditionality, in-kind provision, and public works requirements. If respondents find these tasks difficult, measured preferences may partly reflect confusion or heuristics rather than stable policy views. The survey design addresses this concern in several ways. Respondents receive a common explanation of each program type, complete knowledge checks, and are prevented from progressing without answering required questions. The small-scale pilot surveys were also used to identify confusing wording and tasks that generated inconsistent responses. In the Stage 2 report, we will present robustness checks excluding respondents who complete the survey unusually quickly or provide internally inconsistent responses.

Fourth, the experimental treatments should be interpreted as information interventions rather than as complete policy communication strategies. The video treatments provide short, accurate, and standardized information. These interventions allow us to test whether relatively light-touch information can shift attitudes, but they do not capture the full range of ways governments, political parties, civil society organizations, or media outlets communicate about poverty reduction programs. In addition, any measured treatment effects will be short-run effects, since outcomes are elicited within the same survey. The study therefore cannot determine whether information effects persist over time, whether they would survive exposure to countervailing political messages, or whether they would translate into voting behavior or sustained policy advocacy.

That said, evidence from experiments that do recontact respondents suggests that effects of this kind tend to decrease rather than disappear: of the fifteen information-provision treatments surveyed by Stantcheva (2023), fourteen still show detectable effects at horizons between one week and four months, almost always at reduced magnitude. The two closest to this study in subject matter both persist: Kuziemko et al. (2015) on information about inequality and taxation, and Alesina et al. (2018) on information about intergenerational mobility, each measuring support for redistribution. The priming treatment in Kuziemko et al. (2015) had no detectable effect after one month while their information treatment persisted, which is the pattern one would expect if such treatments work by changing beliefs rather than by making considerations respondents already hold temporarily salient. In any case, the question this study is designed to answer is whether attitudes toward poverty reduction programs are malleable at all, rather than whether a single short video changes them permanently: were the effects to fade within weeks, that would not establish that a sustained, multi-channel information campaign would leave attitudes unchanged.

Fifth, each treatment video bundles multiple distinct empirical claims. The Conditions faced by the Poor video combines three claims about living conditions, economic vulnerability, and limited opportunities faced by poor households. The Effectiveness video addresses three separate myths: that recipients of cash transfers waste the money on harmful goods, that cash transfers reduce

labor supply, and that consumption tax exemptions primarily benefit poor households. If either treatment changes stated support for poverty reduction programs, it will not be possible to determine definitively which specific bundled claim drove the effect. However, we check directly which messages from the videos landed with respondents in the Manipulation check section (see Section 3.3). Each component of the videos is also expected to move different underlying beliefs, as pre-specified in Tables 5 and 6 based on the theoretical mechanisms in Section 2.1.²⁰ In the Stage 2 report, examining which beliefs actually shifted will therefore allow us to infer which parts of each video were likely effective at updating respondents' beliefs, even though the overall treatment effect on support cannot be reliably disentangled.

Sixth, attrition and inattentive responding may threaten the quality of the data. This is a common concern in online surveys, particularly when surveys are relatively long or include video and interactive components. The design minimizes this risk by using short videos, mobile-friendly survey pages, required responses, attention checks, and knowledge checks. We discuss in detail how we will deal with attrition in Section 4.8.5.

Seventh, the descriptive relationships between beliefs and policy preferences should not be interpreted causally. For example, respondents who believe poverty reduction programs are effective may be more supportive of these programs, but this association could reflect prior ideology, self-interest, trust in government, or motivated reasoning. The descriptive hypotheses are therefore intended to document patterns consistent with the conceptual framework, not to identify causal effects of baseline beliefs. The experimental treatments provide the main source of causal variation, while the descriptive analysis helps interpret the mechanisms and heterogeneity underlying the treatment effects.

Finally, there are several external circumstances that could affect implementation. Recruitment may be more difficult in countries with smaller or less representative online panels. Political events, elections, economic shocks, or major policy announcements during fieldwork could temporarily change the salience of poverty reduction programs. Technical issues could also arise if videos do not load properly. We will document any such issues in the Stage 2 report, including the timing of fieldwork in each country, deviations from the planned sample, and any changes to the survey instrument made before launch. If major contemporaneous policy events occur during data collection, we will report the timing of responses and, where feasible, conduct robustness checks excluding affected periods or controlling for fieldwork date.

Overall, these limitations imply that the results should be interpreted as evidence on the short-run effects of information on stated preferences among internet-using adults in middle-income countries. The study is not designed to estimate the long-run political feasibility of particular reforms, nor to measure preferences among the full national population with the same precision as a household survey. Its contribution is instead to provide large-scale, experimentally identified evidence on whether and how accurate information changes public support for different poverty reduction programs, and to document the beliefs and trade-offs that shape these preferences, across a diverse set of countries.

²⁰The exception is the "Living conditions" and "Limited margin for error" parts of the "Conditions faced by the Poor" video, which are expected to shift the same belief.

7 Administrative information

Funding. We gratefully acknowledge financial support from the Global Tax Program of the World Bank, the Centre for Tax Analysis in Developing Countries (TaxDev) through the UK Government Public Finance Centre of Expertise funded by UK International Development, and the ESRC Centre for the Microeconomic Analysis of Public Policy at the Institute for Fiscal Studies (ES/Z504634/1).

Institutional Review Board (ethics approval). Ethics approval for this work was processed through the University College London Economics Local Research Ethics Committee (Project ID 1761).

Declaration of interest. None.

Declaration of generative AI use. The authors used OpenAI’s ChatGPT and Anthropic’s Claude in the preparation of this manuscript for drafting assistance and copyediting. After using these tools, the authors reviewed and edited the manuscript as needed and take full responsibility for its content.

Acknowledgments. We are grateful to Julia Loh, Bansil Malde, David Phillips, Imran Rasul, Yani Tsykerud and seminar participants at the Institute for Fiscal Studies for useful comments and suggestions.

References

- Alesina, Alberto and George-Marios Angeletos, “Fairness and Redistribution,” *American Economic Review*, August 2005, 95 (4), 960–980.
- and Paola Giuliano, “Preferences for Redistribution,” in “Handbook of Social Economics,” Vol. 1, Elsevier, 2011, pp. 93–131.
- , Stefanie Stantcheva, and Edoardo Teso, “Intergenerational Mobility and Preferences for Redistribution,” *American Economic Review*, February 2018, 108 (2), 521–554.
- Almås, Ingvild, Alexander W. Cappelen, and Bertil Tungodden, “Cutthroat Capitalism versus Cuddly Socialism: Are Americans More Meritocratic and Efficiency-Seeking than Scandinavians?,” *Journal of Political Economy*, May 2020, 128 (5), 1753–1788.
- Andersen, Asbjørn G., Simon Franklin, Tigabu Getahun, Andreas Kotsadam, Vincent Somville, and Espen Villanger, “Does wealth reduce support for redistribution? Evidence from an Ethiopian housing lottery,” *Journal of Public Economics*, August 2023, 224, 104939.
- Anderson, Michael L., “Multiple Inference and Gender Differences in the Effects of Early Intervention: A Reevaluation of the Abecedarian, Perry Preschool, and Early Training Projects,” *Journal of the American Statistical Association*, December 2008, 103 (484), 1481–1495. Publisher: Taylor & Francis _eprint: <https://doi.org/10.1198/016214508000000841>.
- Athey, Susan, Julie Tibshirani, and Stefan Wager, “Generalized Random Forests,” *The Annals of Statistics*, 2019, 47 (2), 1148–1178. Publisher: Institute of Mathematical Statistics.

- Bachas, Pierre, Anders Jensen, and Lucie Gadenne**, "Tax Equity in Low- and Middle-Income Countries," *Journal of Economic Perspectives*, February 2024, 38 (1), 55–80.
- , **Lucie Gadenne, and Anders Jensen**, "Informality, Consumption Taxes, and Redistribution," *The Review of Economic Studies*, October 2024, 91 (5), 2604–2634.
- Banerjee, Abhijit, Esther Duflo, and Garima Sharma**, "Long-Term Effects of the Targeting the Ultra Poor Program," *American Economic Review: Insights*, December 2021, 3 (4), 471–486.
- , —, **Nathanael Goldberg, Dean Karlan, Robert Osei, William Parienté, Jeremy Shapiro, Bram Thuysbaert, and Christopher Udry**, "A multifaceted program causes lasting progress for the very poor: Evidence from six countries," *Science*, May 2015, 348 (6236), 1260799. Publisher: American Association for the Advancement of Science.
- , **Rema Hanna, Benjamin A. Olken, and Diana Sverdlin Lisker**, "Social Protection in the Developing World," *Journal of Economic Literature*, December 2024, 62 (4), 1349–1421.
- Banerjee, Abhijit V., Rema Hanna, Gabriel E. Kreindler, and Benjamin A. Olken**, "Debunking the Stereotype of the Lazy Welfare Recipient: Evidence from Cash Transfer Programs," *The World Bank Research Observer*, August 2017, 32 (2), 155–184.
- Benabou, R. and E. A. Ok**, "Social Mobility and the Demand for Redistribution: The Poup Hypothesis," *The Quarterly Journal of Economics*, May 2001, 116 (2), 447–487.
- **and J. Tirole**, "Belief in a Just World and Redistributive Politics," *The Quarterly Journal of Economics*, May 2006, 121 (2), 699–746.
- Benjamini, Yoav and Yosef Hochberg**, "Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing," *Journal of the Royal Statistical Society: Series B (Methodological)*, 1995, 57 (1), 289–300. [_eprint: https://rss.onlinelibrary.wiley.com/doi/pdf/10.1111/j.2517-6161.1995.tb02031.x](https://rss.onlinelibrary.wiley.com/doi/pdf/10.1111/j.2517-6161.1995.tb02031.x).
- Besley, Timothy and Stephen Coate**, "Workfare versus Welfare: Incentive Arguments for Work Requirements in Poverty-Alleviation Programs," *The American Economic Review*, 1992, 82 (1), 249–261. Publisher: American Economic Association.
- Bursztyn, Leonardo**, "Poverty and the Political Economy of Public Education Spending: Evidence from Brazil," *Journal of the European Economic Association*, October 2016, 14 (5), 1101–1128.
- Cerkez, Nicolas, Adnan Q Khan, Imran Rasul, and Anam Shoaib**, "Big Push Pro-poor Policies and Economic Circumstances: Reality, Perceptions and Attitudes," 2025.
- Cruces, Guillermo, Ricardo Perez-Truglia, and Martin Tetaz**, "Biased perceptions of income distribution and preferences for redistribution: Evidence from a survey experiment," *Journal of Public Economics*, February 2013, 98, 100–112.
- Currie, Janet and Firouz Gahvari**, "Transfers in Cash and In-Kind: Theory Meets the Data," *Journal of Economic Literature*, May 2008, 46 (2), 333–383.
- de la Feria, Rita and Artur Swistak**, "Designing a Progressive VAT," *IMF Working Papers*, April 2024, 2024 (078), 1.

- Dechezleprêtre, Antoine, Adrien Fabre, Tobias Kruse, Bluebery Planterose, Ana Sanchez Chico, and Stefanie Stantcheva**, “Fighting Climate Change: International Attitudes toward Climate Policies,” *American Economic Review*, April 2025, 115 (4), 1258–1300.
- DellaVigna, Stefano and Devin Pope**, “Predicting Experimental Results: Who Knows What?,” *Journal of Political Economy*, December 2018, 126 (6), 2410–2456. Publisher: The University of Chicago Press.
- , —, —, and **Eva Vivalt**, “Predict science to improve science,” *Science*, October 2019, 366 (6464), 428–429. Publisher: American Association for the Advancement of Science.
- Evans, David K. and Anna Popova**, “Cash Transfers and Temptation Goods,” *Economic Development and Cultural Change*, 2017, 65 (2), 189–221. Publisher: The University of Chicago Press.
- Ferrario, Beatrice and Stefanie Stantcheva**, “Eliciting People’s First-Order Concerns: Text Analysis of Open-Ended Survey Questions,” *AEA Papers and Proceedings*, May 2022, 112, 163–169.
- Geiecke, Friedrich and Xavier Jaravel**, “Conversations at Scale: Robust AI-led Interviews with a Simple Open-Source Platform,” 2024.
- Gelman, Andrew**, “Prior distributions for variance parameters in hierarchical models (comment on article by Browne and Draper),” *Bayesian Analysis*, September 2006, 1 (3).
- Grömping, Ulrike**, “Estimators of Relative Importance in Linear Regression Based on Variance Decomposition,” *The American Statistician*, May 2007, 61 (2), 139–147. _eprint: <https://doi.org/10.1198/000313007X188252>.
- Haaland, Ingar, Christopher Roth, and Johannes Wohlfart**, “Designing Information Provision Experiments,” *Journal of Economic Literature*, March 2023, 61 (1), 3–40.
- , —, —, **Stefanie Stantcheva**, and **Johannes Wohlfart**, “Measuring What Is Top of Mind,” 2024.
- , —, —, —, and —, “Understanding Economic Behavior Using Open-ended Survey Data,” Technical Report w32421, National Bureau of Economic Research, Cambridge, MA May 2024.
- Hanna, Rema and Benjamin A. Olken, eds**, *The Handbook of Social Protection: Evidence and New Directions for Low- and Middle-Income Countries*, Cambridge, Massachusetts: The MIT Press, 2026.
- Hendren, Nathaniel and Ben Sprung-Keyser**, “A Unified Welfare Analysis of Government Policies*,” *The Quarterly Journal of Economics*, August 2020, 135 (3), 1209–1318.
- Hoy, Christopher**, “How does progressivity impact tax morale? Experimental evidence across developing countries,” *Journal of Development Economics*, January 2025, 172, 103398.
- and **Franziska Mager**, “Why Are Relatively Poor People Not More Supportive of Redistribution? Evidence from a Randomized Survey Experiment across Ten Countries,” *American Economic Journal: Economic Policy*, November 2021, 13 (4), 299–328.
- , **Yeon Soo Kim, Minh Cong Nguyen, Mariano Sosa, and Sailesh Tiwari**, “Attitudes towards reducing fossil fuel subsidies: Evidence across 12 middle-income countries,” *Journal of Development Economics*, January 2026, 178, 103612.

- Inchauste, Gabriela and Nora Lustig**, *The Distributional Impact of Taxes and Transfers: Evidence From Eight Developing Countries*, Washington, DC: World Bank, August 2017.
- Kling, Jeffrey R, Jeffrey B Liebman, and Lawrence F Katz**, "Experimental Analysis of Neighborhood Effects," *Econometrica*, 2007, 75 (1), 83–119. _eprint: <https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1468-0262.2007.00733.x>.
- Kotlikoff, Lawrence J., Guillermo Lagarda, and Gabriel Marin**, "A personalized VAT with capital transfers: a reform to protect low-income households in Mexico," *International Tax and Public Finance*, October 2025, 32 (5), 1573–1633.
- Kreuter, F., S. Presser, and R. Tourangeau**, "Social Desirability Bias in CATI, IVR, and Web Surveys: The Effects of Mode and Question Sensitivity," *Public Opinion Quarterly*, December 2008, 72 (5), 847–865.
- Kuziemko, Ilyana, Michael I. Norton, Emmanuel Saez, and Stefanie Stantcheva**, "How Elastic Are Preferences for Redistribution? Evidence from Randomized Survey Experiments," *American Economic Review*, April 2015, 105 (4), 1478–1508.
- Lee, David S.**, "Training, Wages, and Sample Selection: Estimating Sharp Bounds on Treatment Effects," *The Review of Economic Studies*, July 2009, 76 (3), 1071–1102.
- Lindeman, Richard Harold, Peter Francis Merenda, and Ruth Z. Gold**, *Introduction to bivariate and multivariate analysis*, Glenview, Ill: Scott, Foresman, 1980.
- Liscow, Zachary and Abigail Pershing**, "Why Is So Much Redistribution In-Kind and Not in Cash? Evidence from a Survey Experiment," *National Tax Journal*, June 2022, 75 (2), 313–354.
- Luttmer, Erzo F. P and Monica Singhal**, "Culture, Context, and the Taste for Redistribution," *American Economic Journal: Economic Policy*, February 2011, 3 (1), 157–179.
- Meager, Rachael**, "Understanding the Average Impact of Microcredit Expansions: A Bayesian Hierarchical Analysis of Seven Randomized Experiments," *American Economic Journal: Applied Economics*, January 2019, 11 (1), 57–91.
- , "Aggregating Distributional Treatment Effects: A Bayesian Hierarchical Analysis of the Microcredit Literature," *American Economic Review*, June 2022, 112 (6), 1818–1847.
- Meltzer, Allan H. and Scott F. Richard**, "A Rational Theory of the Size of Government," *Journal of Political Economy*, 1981, 89 (5), 914–927. Publisher: The University of Chicago Press.
- Mummolo, Jonathan and Erik Peterson**, "Demand Effects in Survey Experiments: An Empirical Assessment," *American Political Science Review*, May 2019, 113 (2), 517–529.
- Nichols, Albert L. and Richard J. Zeckhauser**, "Targeting Transfers through Restrictions on Recipients," *The American Economic Review*, 1982, 72 (2), 372–377. Publisher: American Economic Association.
- O, Ana L. De La**, "Do Conditional Cash Transfers Affect Electoral Behavior? Evidence from a Randomized Experiment in Mexico," *American Journal of Political Science*, 2013, 57 (1), 1–14. _eprint: <https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1540-5907.2012.00617.x>.

- Parker, Susan W. and Petra E. Todd**, “Conditional Cash Transfers: The Case of *Progres/Oportunidades*,” *Journal of Economic Literature*, September 2017, 55 (3), 866–915.
- Pellicer, Miquel, Patrizio Piraino, and Eva Wegner**, “Perceptions of inevitability and demand for redistribution: Evidence from a survey experiment,” *Journal of Economic Behavior & Organization*, March 2019, 159, 274–288.
- Quidt, Jonathan De, Johannes Haushofer, and Christopher Roth**, “Measuring and Bounding Experimenter Demand,” *American Economic Review*, November 2018, 108 (11), 3266–3302.
- Redonda, Agustin, Christian von Haldenwang, and Flurim Aliu**, “Global Tax Expenditures Database (GTED),” October 2025.
- Rubin, Donald B.**, “Estimation in Parallel Randomized Experiments,” *Journal of Educational Statistics*, 1981, 6 (4), 377–401. Publisher: [Sage Publications, Inc., American Educational Research Association, American Statistical Association].
- Stantcheva, Stefanie**, “Understanding economic policies: What do people know and learn?,” 2020.
- , “Understanding Tax Policy: How do People Reason?,” *The Quarterly Journal of Economics*, October 2021, 136 (4), 2309–2369.
- , “Understanding of Trade,” 2022.
- , “How to Run Surveys: A Guide to Creating Your Own Identifying Variation and Revealing the Invisible,” *Annual Review of Economics*, September 2023, 15 (1), 205–234.
- Vivalt, Eva and Aidan Coville**, “How do policymakers update their beliefs?,” *Journal of Development Economics*, October 2023, 165, 103121.
- Wager, Stefan and Susan Athey**, “Estimation and Inference of Heterogeneous Treatment Effects using Random Forests,” *Journal of the American Statistical Association*, July 2018, 113 (523), 1228–1242.
- Warwick, Ross, Tom Harris, David Phillips, Maya Goldman, Jon Jellema, Gabriela Inchauste, and Karolina Goraus-Tańska**, “The redistributive power of cash transfers vs VAT exemptions: A multi-country study,” *World Development*, March 2022, 151, 105742.
- World Bank**, *Poverty, Prosperity, and Planet Report 2024: Pathways Out of the Polycrisis*, Washington, DC: World Bank, October 2024.