

Deconstructing the (un)affordability of clean cooking fuels through a randomized trial in rural Tanzania

Received: 2 August 2024

Accepted: 16 April 2025

Published online: 03 June 2025

 Check for updates

Annelise Gill-Wiehl ^{1,2}✉, Isha Ray ¹, Robert Katikiro ³,
Daniel M. Kammen ^{1,4} & Alan Hubbard⁵

Low-income users struggle to save for clean cooking fuel costs. We test whether a lockbox intervention paired with micro-saving nudges could alleviate the unaffordability of clean fuels. In a year-long stepped-wedge randomized control trial in Tanzania ($n = 511$), we find that compared to savings nudges only, a lockbox and savings nudges increased annual refills of liquefied petroleum gas (LPG) by 1.4 (0.054 per 2 weeks ([95% confidence interval: 0.043, 0.066], $P < 0.0001$)) and minimally decreased firewood use but had no effect on lags between LPG refills or the frequency of LPG and charcoal use. We find that easing liquidity constraints is insufficient for exclusive LPG use when LPG is the financial responsibility of only women, who ration LPG purchases to meet other household needs and social expectations. The financial and gendered realities of low-income consumers demand clean energy policies beyond easing liquidity constraints or targeted subsidies.

About 2.4 billion people rely on smoky solid fuels, resulting in ambient and household air pollution (HAP), to which over eight million annual untimely deaths are attributed¹. Near-exclusive use of ‘clean’ stoves and fuels, as categorized by the World Health Organization, is needed to avert negative health outcomes². Affordability remains one of the largest barriers to the consistent use of clean cookstoves and fuels^{3,4}. The challenge for energy policy is thus to provide affordable clean stoves and fuels that minimize users reverting to biomass (the ‘refill gap’)⁵ or using multiple stoves in parallel (stove ‘stacking’)⁶.

Clean cooking has two cost components: the upfront stove cost and the ongoing fuel and maintenance costs^{4,6–10}. The fuel cost consists of everyday sums or occasional lump sums. Traditional fuels (wood, agricultural waste) are often purchased or collected day to day, whereas cleaner fuels such as liquefied petroleum gas (LPG) are purchased on a monthly (or longer) basis. Many individuals without clean fuels have neither regular cash flows nor financial tools^{11,12}, making it difficult to

save even modest sums. Further, individuals under the stress of chronic poverty often ‘tunnel’ or focus exclusively on immediate priorities, struggle to make forward-looking choices and must exert constant vigilance to not spend the small amounts that they could (and often must) save^{13,14}.

The clean cooking literature has focused more on the upfront cost of clean stoves^{15–17} than on financing the recurring fuel costs or reducing the refill gap. Micro-saving programmes could potentially accommodate low-income users’ spending and saving patterns for either upfront or recurring costs. Numerous randomized control trials (RCTs) have investigated banked savings programmes^{18–20}, organized savings groups^{21–23} and portable devices such as lockboxes^{23–25}, with varying levels of commitment. Lockboxes have been proven to help increase savings overall²⁶ as well specifically for healthcare²³ and education^{24,25}. We hypothesized that tools to encourage small savings could help to increase (on time) clean fuel refills.

¹Energy and Resources Group, University of California, Berkeley, CA, USA. ²Environmental Health Sciences, Columbia University, New York, NY, USA.

³Agricultural and Natural Resources Economics and Business, University of Dar es Salaam, Dar es Salaam, Tanzania. ⁴Goldman School of Public Policy, University of California, Berkeley, CA, USA. ⁵Division of Biostatistics, School of Public Health, University of California, Berkeley, CA, USA.

✉ e-mail: ag5050@cumc.columbia.edu

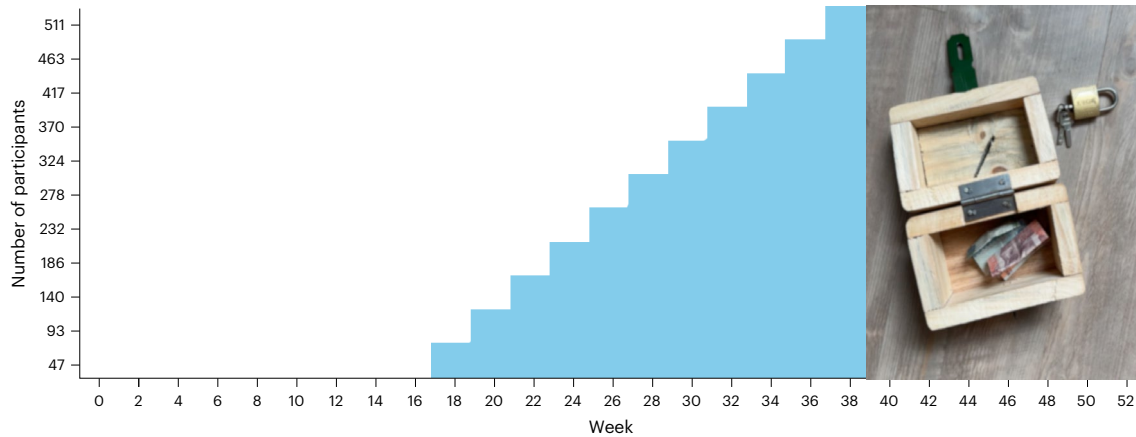


Fig. 1 | Schematic of study design. This illustration depicts how individuals transitioned into the treatment of a locked deposit box. We enrolled 511 participants at baseline ($t = 0$) and then randomly assigned each to transition into treatment after weeks 16 through 32. The first -46 participants received the intervention (blue) at week 18, then -46 more participants at week 20 and so forth

until all 511 respondents received the treatment. Once in treatment, participants remained thus until the end of the study. The randomized rollout was intended to create balanced comparison groups between the control (white) and the treatment (blue). Photo of lockbox and key taken by A.G.-W. Lockboxes were hand crafted by a local artist.

The behavioural economics literature has developed a range of interventions to encourage socially desirable habits, including nudges, boosts, gamification, priming, framing and so on^{27,28}. Nudge theory²⁹ is based on structuring choices to guide individuals towards better decisions without restricting their freedom or requiring notable changes to existing practices. Savings nudges in the literature make strategic use of such small interventions^{30,31} to protect against procrastination or valuing current consumption over future needs. Another suggested solution to the savings challenge is setting up dedicated saving accounts (earmarked for specific purchases) for low-income individuals. These accounts do not alleviate scarcity itself but could remind participants of their savings goals³², lower the mental burden of remembering to save and give individuals a sense of control even under constrained circumstances.

Combining these literatures, we designed an intervention that incorporates a dedicated account (drawing on behavioural economics) as a financial tool for small savings (drawing on clean cooking research) and acknowledges that nudging or encouragement (also from behavioural economics) may be necessary to facilitate a pro-saving behaviour change. We provided individuals a locked deposit box ('lockbox'), earmarked for LPG cylinders, which served as an informal financial device and a physical reminder to save, but without formal commitment or enforcement. Our theory of change was that if individuals saved in the lockbox, this would increase their on-hand liquidity (having eased their liquidity constraints) to spend on LPG refills. We present here the results on changes on LPG purchases on account of any change in savings. We also present new ethnographic evidence to explain our experimental results, grounded in understanding intra-household gender dynamics and the portfolio of household needs as key determinants of purchasing and using clean fuels.

Experimental design

We designed a year-long stepped-wedge RCT in which participants ($n = 511$) across Shirati, a rural town in Mara Region, Tanzania, started in control and then crossed over in waves into the intervention/treatment group. At the baseline, no one used LPG. Each participant received two free 6-kg LPG cylinder/burner sets, along with information from their community technology worker (CTW). All future refills of the LPG cylinder were the participant's responsibility (-US\$10). Participants were randomly assigned to roll over into the treatment (that is, receiving the lockbox) (Fig. 1). The CTWs followed up every 2 weeks to encourage clean fuel use and nudge participants to save, for both treatment and control. That is, the control group also received encouragement nudges. At these visits, the CTWs also conducted surveys. From those surveys, LPG

Table 1 | Socio-demographic characteristics of our sampled respondents, $n = 511$

Female respondent (main cook), n (%)	480 (94%)
Female-headed households, n (%)	145 (29%)
Age, mean [range]	41 [18,88]
Luo tribe, n (%)	446 (88%)
Number of individuals living together, mean [range]	6 [1,20]
Children <12, mean [range]	2 [0,15]
Number of individuals eating together, mean [range]	6 [1,16]
Occupation of main cook	
Farmer, n (%)	411 (81%)
Business, n (%)	51 (10%)
Nurse, n (%)	5 (1%)
Cares for the home, n (%)	24 (5%)
Other occupation, n (%)	15 (3%)
Education of main cook	
No formal education, n (%)	18 (4%)
Primary school, n (%)	450 (88%)
Secondary school, n (%)	39 (8%)
University, n (%)	1 (0%)
Expenditure and saving	
Respondent's weekly expenditure, mean (s.d.)	US\$17.2 (10.4) ^a
Respondent's weekly expenditure PPP, mean (s.d.)	US\$56.7 (34.4) ^b
Ability to save at all, n (%)	286 (56%)
Female saves alone, n (%), endline ^c	433 (89%)
Attrition	1%
Maximum missingness within a single visit	6%

^aAn average exchange rate of US\$1=2,325.96 Tanzanian shillings was used in January 2022.

^bWe use a PPP of 3.30 for 2022. ^cThis value was from the final visit where 489 respondents were contacted that week, and 18 did not answer this question as they were struggling to save.

purchase receipts and triangulation with photos, we constructed our pre-specified primary outcomes: purchased refills, the days between running out of LPG and purchasing the next cylinder ('the refill gap'), categories of LPG use (exclusive, habitual, occasional and none) and

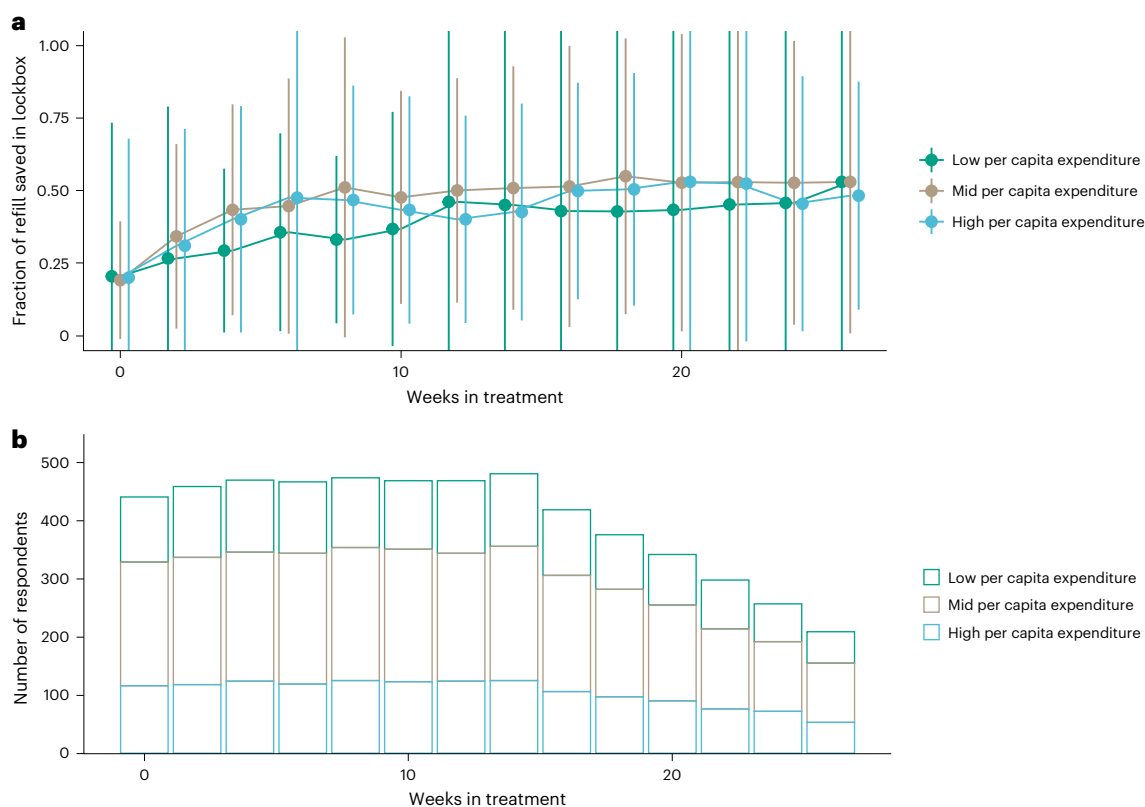


Fig. 2 | Amount saved in the lockbox as a fraction of the cylinder refill cost.

a, We tracked the amount respondents had saved in their lockboxes as a fraction of the total cost to refill a single liquefied petroleum gas cylinder (- US\$10; -US\$26 purchasing power parity) across expenditure categories ($n = 6,044$ observations from 498 participants); data are presented as

mean values \pm standard deviation. **b**, How many participants were enrolled in treatment at these times (as participants transitioned to treatment in waves). We stop the x axis at 26 weeks in treatment or control to keep observations around 50 individuals for each time point.

biomass (firewood/charcoal) usage³³. Savings outcomes are reported in another paper (A.G.-W., manuscript in preparation). Following the full trial, we conducted a 6-month period of qualitative work that included observations, focus groups, budgetary games and 90 semi-structured interviews (Methods).

Sample characteristics

We spoke with the mostly female main cooks (94%) with an average weekly expenditure of US\$56.7 (standard deviation (s.d.) 34.4) after adjusting for purchasing power parity (PPP) (Table 1). Women earned small amounts of income from selling products from their farms, selling small fish, braiding hair and so on. LPG use in the region was very low (~ 2% in 2018)³⁴. Confirming previous anthropological literature from sub-Saharan Africa^{35–37}, spouses kept finances completely separate: “There are very few couples who save their money together” (Respondent 4, Focus Group 1). Women were responsible for daily household expenses (for example, food, soap, charcoal and children’s needs), while men were responsible for larger purchases (building the house, furniture, television and so on). Couples could come together for certain purchases, but this was not the norm. Spouses rarely disclosed their financial status to each other. Thus, our main respondents reported only their own available weekly expenditures. At the baseline, 56% of respondents reported being able to save anything at all.

Treatment effects on clean fuel and biomass usage

We assess treatment impact on refills by constructing a user’s rate of refilling while in control versus treatment (total refills in control or treatment divided by cumulative time in control or treatment). For all other outcomes, we analyse immediate and cumulative impacts of

offering participants a lockbox. The immediate impact of treatment is the outcome for the following 2-week time period, whereas the cumulative impact of treatment is the average impact of each additional time period in treatment. For all outcomes, we report the average treatment effect (ATE), that is, the difference in the mean outcomes between treatment and control groups (Methods) and conduct subgroup analysis for first (low), second and third (mid) and fourth quartile (high) expenditure strata (Supplementary Fig. 4).

Regarding savings, we find that uptake of the lockbox was very high (~ 96% by the final visit) and that treatment immediately caused an additional 13% (95% confidence interval (CI): (10.7%, 15.4%, $P < 0.001$) of participants to save within the next 2 weeks, compared to control (A.G.-W., manuscript in preparation). On average, respondents had half the refill amount saved in the lockbox (Fig. 2).

On average, participants refilled 3.5 (min: 0, max: 18) times throughout the year-long study. We estimate the ATE on total refills to be 0.054 ([95% CI: 0.043, 0.066], $P < 0.001$) per 2-week period (Table 2). Our results when extrapolated indicate that if everyone had received treatment over the year-long study period, average refills would have increased by -1.4 (0.054×26).

We observe very few instances of individuals reporting a lag between LPG refills. They were stacking with other fuels and seemed to time running out of LPG for when they had money to refill (Supplementary Fig. 1). Respondents’ rationing behaviour avoided any measured LPG ‘refill gap’. During follow-up interviews, some women acknowledged rationing LPG use to avoid absolutely running out: “I decided to reduce my gas use so that it should take at least one month to run out... if I decide to use it only, it will not even take one month [to run out] ...so I decide to use it to make only tea and reheat the food”

Table 2 | Intent-to-treat ATE estimates, immediate and cumulative

Analysis of immediate impact	Observations	ATE (95% CI)
LPG use categories		
Exclusive use	12,314	-0.00057 (-0.010, 0.0092)
Habitual use	12,312	0.0078 (-0.043, 0.059)
Occasional use	12,312	0.0063 (-0.036, 0.049)
No LPG use	12,312	-0.018 (-0.065, 0.028)
Biomass use		
Total firewood (kg d ⁻¹)	8,012	-0.083 (-0.37, 0.20)
Total charcoal (kg d ⁻¹)	3,324	0.15 (-0.10, 0.40)
Analysis of cumulative impact		ATE (95% CI)
LPG refill		
Total refills	11,132	0.054*** (0.043, 0.066)
LPG use categories		
Exclusive use	12,314	0.000044 (-0.00074, 0.0013)
Habitual use	12,312	0.0039 (-0.0018, 0.0093)
Occasional use	12,312	0.0032 (-0.0017, 0.0083)
No LPG use	12,314	-0.0079 (-0.010, 0.0051)
Biomass use		
Total firewood (kg d ⁻¹)	8,012	-0.038* (-0.078, 0.002)
Total charcoal (kg d ⁻¹)	3,324	0.025 (-0.012, 0.063)

Note the immediate impact of treatment is the outcome for the following 2-week time period, whereas the cumulative impact of treatment refers to the average impact of each additional time period in treatment. We modelled immediate and cumulative impact of treatment on binary and continuous outcomes using logistic and linear regressions, respectively. We model the rate of refilling over the time periods in control vs treatment using Poisson regression. For all, we report robust inference and use two-sided tests. No adjustment was made for multiple comparisons (Methods). *** $P < 0.01$, * $P < 0.1$.

(Respondent 2, Focus Group 1). Many individuals reported stacking, reducing their LPG use to occasional (such as reheating food) to control when they would run out; others simply never refilled (Fig. 3 and Supplementary Fig. 1).

Most respondents claimed to need only one cylinder, holding the second as reserve for when the first cylinder ran out: “While using one cylinder I save for another one” (Respondent 387). This was also a form of rationing: “...it’s hard to get money to refill them both at once” (Respondent 178). Using only one cylinder was also a way to maintain savings for other items: “I was just saving and not only for gas...I have not refilled the second one because for now we have a lot of funerals” (Respondent 54). Money for refills was especially tight as LPG purchases were solely in the women’s expenditure domain; their ‘share’ of the household expenses was being saved towards LPG (and food-related expenses overall). Men rarely assisted with LPG refills: “In all cooking matters women are the one responsible... as a woman you will struggle to make sure you refill the cylinder and cook, so when he comes home, he only eats” (Respondent 4).

Regarding self-reported LPG use categories, we find, while controlling for weeks into the experiment, neither an immediate nor cumulative impact. Descriptively, the percentage of participants reporting no LPG use rose with number of weeks in control across all income strata. The percentage reporting no LPG use was, on average, higher in treatment than in control but did not increase measurably with weeks in treatment (Table 2, Fig. 3 and Supplementary Tables 1–3). The stability of LPG use categories reinforces our result that users were rationing their LPG.

Each additional 2-week period in treatment decreased average firewood use by -0.038 kg (95% CI: $(-0.078, 0.002)$) (extrapolates to -1 kg a day by the end of the year), particularly for participants with mid-level expenditure (Supplementary Table 3). Stacking was stable throughout control and treatment, consistent with our treatment effects, which found minimal effect on biomass usage (Fig. 3 and Supplementary Fig. 1). While stacking fuels, respondents across expenditure strata were spending well over the 5% threshold (using only the women’s reported expenditures as denominators) set by the Energy Sector Management Assistance Program (ESMAP) as the affordability bar for cooking fuel³⁸. Low expenditure participants spent the largest percentage (Supplementary Fig. 3). The post-treatment fuel expenditure ratio is largely consistent with that reported at baseline; polluting fuels thus account for substantial costs, often similar to clean fuels per unit of energy generated (Supplementary Fig. 2). Our interviews and focus groups explained these results through two recurring themes: LPG was but one priority among many, and the nature of cooking fuel purchases is highly gendered.

Easing liquidity is insufficient given a portfolio of needs

The treatment was an effective nudge for respondents to save, thus easing their liquidity constraints (A.G.-W., manuscript in preparation). They dipped into those savings for refills, essential items as well as emergencies: “I am the one buying each and everything in this house [so] the money that I was supposed to buy gas with I use to buy food” (Respondent 67). Or “the child was sick, so we used the [lockbox] money” (Respondent 400). Most women we interviewed claimed to be saving for LPG but ultimately saved whatever they could, whenever they could (Fig. 4).

The lockbox increased participants’ recognition of the need to save for LPG: “it is good to save because it’s hard to get all 24,000 shillings at once” (Respondent 202). And they did in fact save. This awareness could have contributed to the stabilizing LPG use and continued rationing: “I can even see how much I am left with to reach the total amount” (Respondent 273). The lockbox visualized for participants the fraction of the hefty LPG refill cost that they had on hand: “if I use gas daily it will run out quickly before I get enough money to refill” (Respondent 13). Participants used the lockbox to time when the cylinder would run out to when they had enough saved (Fig. 4).

When the gas ran out, they would look in the lockbox to see if there was enough to refill. If there was enough, or almost enough, they tended to refill. If there was any extra, they would buy other things, they would not leave the residual towards a future refill. If there was not enough, they waited to refill, turning to the second cylinder and reducing their LPG use: “in the lockbox we were just saving so that when the cylinder runs out, we go and refill and if there is some amount of money left after refilling, we just use the money to buy other items in the house” (Respondent 500). Having enough money on hand in their lockboxes (that is, easing liquidity constraints) was a necessary but by no means sufficient condition to increase LPG use given the portfolio of needs.

The literature suggests many reasons for continued stove stacking^{8,39}, but the slight reduction in firewood and no change in charcoal that we find is largely underpinned by low and fluctuating incomes. The instability of their financial situations made saving particularly hard: “I am now saving if I get money, it’s not every week... because I depend on the business of oranges and as you know it’s seasonal” (Respondent 272). Distance to refill, often cited as a separate barrier, contributed to the affordability challenge in that transport typically added another dollar to the refill cost). The widespread prevalence of rationing (using LPG when in a hurry, reducing LPG to never run out, refusing to use LPG for beans because beans have a long cooking time and so on) points to unaffordability being a dominant reason for persistent stacking. For instance: “I don’t want to run out” [when cooking beans] (Respondent 101). To our probe: “So is it something

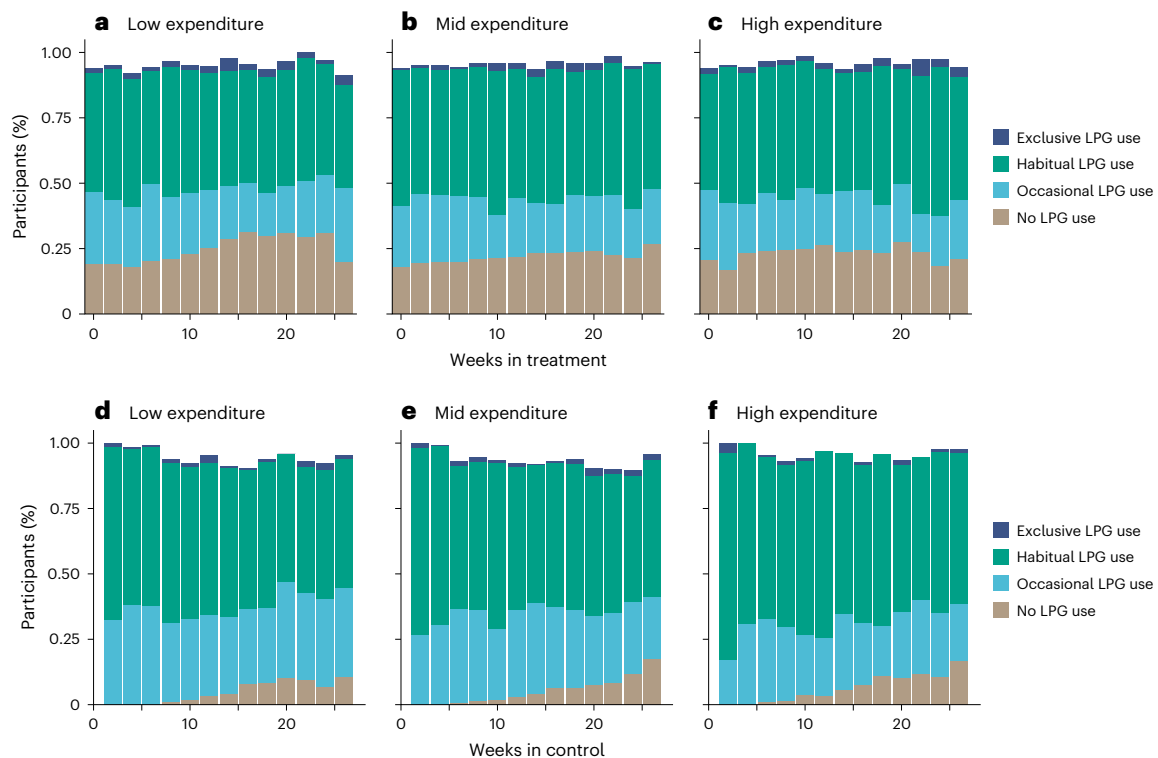


Fig. 3 | Categories of LPG use during study. a–f, We tracked categories of LPG use across low (a,d), mid (b,e) and high (c,f) per capita expenditure strata during treatment (a–c) and control (d–f) periods. Most respondents were using a

combination of LPG and firewood and/or charcoal (that is, stacking stoves). We stop the x-axis at 26 weeks in treatment or control to keep observations around 50 individuals for each time point.

about the beans?,” the response was: “No, it’s about the money”. Participants expressed a strong preference for LPG: “if gas was free, I’d refill everyday...even if there was still available firewood and charcoal” (Respondent 13).

The literature already documents the unaffordability of clean cooking fuels^{40–42}. Our results support that literature but show that alleviating liquidity constraints alone is insufficient to mitigate unaffordability, in contrast to previously posed hypotheses⁴. We show that even when there are savings increases, LPG is not always prioritized amid myriad household needs. With saved money, many items become more affordable, not LPG alone. There have not been any randomized trials of pay-as-you-go LPG⁴³, but our results suggest similar competing needs would arise. We also find that fluctuating business opportunities make for fluctuating savings: “[my savings] depends on how much I sell, because sometimes I sell 500 or 1000 or 3000 TSH” (Respondent 179). Small irregular cash flows make it easier to purchase fuel in small irregular quantities. In reality, even if LPG is not substantially more expensive than charcoal per unit of energy, it is more ‘unaffordable’ because it is a lumpy purchase. We conclude that LPG is unaffordable not as a stand-alone commodity but in the context of other unmet needs and uncertain income streams.

The gendered nature of clean fuel purchases

We initially targeted main cooks as representing their households. However, we found that not just the physical but the financial burden of cooking falls on the woman. A survey enumerator explained: “Women have always been the ones to collect firewood or purchase charcoal, so even though it is now purchasing gas, it is still her responsibility”. Using their own earnings, a stipend from their husbands or both, women alone were responsible for cooking fuel purchases. We highlight here the under-appreciated separation of financial responsibility for fuels rather than the intra-household power asymmetry highlighted in other work^{44,45}. In this part of Africa, LPG is not a ‘household’ expense, so the

total income in ‘a household’ is not necessarily a predictor of LPG’s affordability.

Our results suggest that successful efforts to increase clean fuel use demand a gendered understanding of household responsibilities. We need to understand what women are asked to give up to purchase clean fuels consistently. Where household finances are entirely in women’s hands, they are the ones who must choose between LPG and clothes for their children. Some researchers have implied that non-habitual use of clean fuels or continued stacking means that the users do not truly value the benefits of clean stoves^{9,46,47}, but we find that women do value and purchase LPG with higher savings. While they do not refill LPG at the frequencies needed to see health benefits (exclusive use)², we argue that this reflects long-standing practices of women paying for household needs with their own health when they cannot pay with cash. Feminist scholarship, echoing the work of Sen⁴⁸, has argued that an individual’s overt preferences do not reflect individual welfare where the preferences themselves have had to adapt to constrained circumstances⁴⁹. We, therefore, cannot conflate LPG non-use with low valuations for clean energy; rather, we must re-conceptualize this choice as reflecting low valuations of women’s health. This re-conceptualization strongly supports the need for the energy and health sectors to account for the full costs of their interventions and for the individual who is expected to pay these costs.

Discussion and conclusions

Our study demonstrates that an effective ‘nudge’ to (modestly) increase LPG consumption and keep individuals from running out completely of LPG is possible at the level of individual cooks. Although we only find an increase in -1.4 refills (8.4 kgs) over the year and no change in LPG use categories, prominent LPG policies have been even less successful. India’s policies for individuals below-the-poverty line (for example, Pradhan Mantri Ujjwala Yojana), which covered the upfront cost of the stove, led to an annual refill increase of 0.68 kg of LPG⁵⁰.

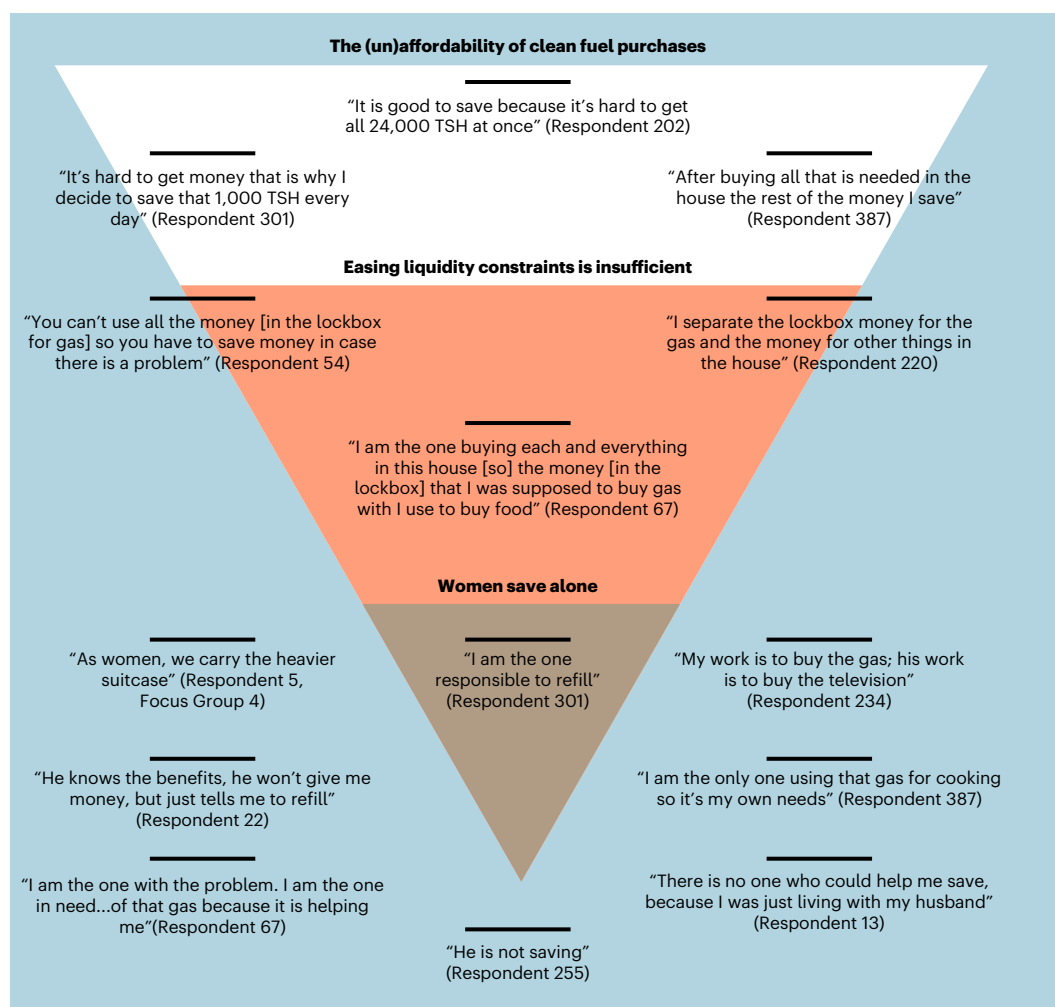


Fig. 4 | Qualitative results. Our qualitative results reveal that the unaffordability of LPG (white) is underpinned by (1) a portfolio of (competing and urgent) needs that prevent easing liquidity constraints from fully addressing the financial challenge (orange) and (2) the gendered expectations of household budgets that place LPG in the woman’s domain (brown). TSH, Tanzanian shillings.

Experimental evidence indicates that LPG demand is very sensitive to price⁵¹, even with pay-as-you-go models⁵². Our intervention increased LPG consumption, but not exclusive use, the level needed for potential health benefits². Even with eased liquidity constraints, therefore, and intention to save for clean cooking fuels, financial and gendered realities forced our participants to continue to stack LPG and solid-fuel stoves. Our findings have research and policy implications well beyond rural Tanzania, as intra-household budgeting separation and multiple household needs are common in sub-Saharan Africa.

Our results are a call to reconsider a key metric used to measure clean fuel affordability: the conventional affordability ratio⁴⁰. The conventional affordability ratio is typically constructed as the ratio of fuel costs to the household monthly expenditure⁴⁰. However, we have shown that fuel affordability may be primarily a function of the woman’s, not the household’s budget, and relevant on weekly or even daily timescales, given our respondents’ earning patterns. ‘Household’ monthly affordability measures may not accurately estimate a low-income user’s ability to purchase clean fuel. Further, our work reveals that this measure on its own, without the context of other needs, will not be a useful predictor of LPG use. Residual income approaches to affordability that consider essential items (rent, food and so on) and gender-specific affordability metrics⁴ would provide a more realistic picture of clean fuel affordability.

Our findings call for energy policy to consider that increasing incomes/saving, subsidizing LPG and/or providing purchasing options

with relaxed liquidity constraints (for example, smaller quantities⁴, pay-as-you-go technologies⁵²) are necessary but probably not sufficient to ensure habitual or exclusive LPG use and abandonment of polluting fuels. Simply alleviating liquidity constraints or providing cheaper LPG will not be enough if there are too many other needs and only a portion of household income available for them. To foster exclusive clean fuel use, we may need to pursue direct fuel provisions, such as vouchers for the entire cylinder (unlike the partial cylinder voucher explored in Peru⁵³) or direct, targeted delivery (such as nutritional programming⁵⁴ or drinking water through pipes or tanker trucks⁵⁵), that can ensure a public health-consistent level of use. Energy research and policy must understand how fuel purchases are embedded in gendered, that is, separate but not necessarily asymmetric, intra-household dynamics and within a portfolio of urgent unmet needs to increase clean fuel use.

Methods

Study setting

We conducted this work in Shirati, Rorya District, Mara Region, Tanzania. Shirati is a rural town of ~50,000 on the edge of Lake Victoria near the Kenyan border. Shirati has a tropical climate with light rains from October to December and heavy rains from March to June.

According to the 2022 census, when our study was completed, the average number of ‘persons who spent the census night’⁵⁶ (pg. 77) living together in the Mara region was five individuals. In Mara, 42.3% of households were female headed, but over half of adults were married

(56%)⁵⁷. According to Tanzania's 2018 Household Budget survey, most individuals in Mara were 47 years old, had completed primary school (62%); many used a rechargeable lamp (35.3%) or solar lamps (26.6%) as their main source of lighting³⁴. Average monthly household consumption was -US\$207 in 2017/2018. Female heads of households were self-employed (including farming, petty trade and so on), 14% were unemployed, never worked or did not know and 17.5% were unpaid household workers. Only 11.6% of households in the region had at least one member (typically not a woman) with a bank account. We note that these characteristics are for the region and of all individuals not only main cooks (our sample). In Mara, most used firewood (72.9%) or charcoal (23.5%) for cooking, although LPG was available.

In Shirati, firewood and charcoal could be collected and purchased in every village, and charcoal making was a common small business. For LPG users, 6- or 12-kg cylinders could be exchanged at retail points in the village with the main market, Obwere. To refill, participants had to either walk with the heavy cylinder or pay the extra cost of a motorcycle or three-wheeled delivery cart (- US\$1). We intentionally targeted a rural setting as rural communities are disproportionately understudied and are often targeted for improved but not (truly) clean stoves.

Sampling and randomization

We randomized both how participants were selected to participate and the order in which selected participants were assigned to the intervention. We relied on a combination of ward-level shapefiles, Google Maps, interviews with village leaders, local knowledge and Python to randomly place -511 points across Shirati and then invited the main cook from the closest homestead to participate. After baseline data collection, the first author determined treatment distribution order using a random number generator in R.

Participants

Main cooks were approached by our local team to participate. We intentionally targeted main cooks as the most pertinent individual involved in cooking for the household. Further, we needed to train main cooks on LPG use and safety. After a few days to consider our request, we returned to obtain written informed consent. The only eligibility criterion for our selected participants (Supplementary Fig. 1) was that they did not already use LPG for cooking. This was not a limiting criterion as LPG use in the Mara region was 1.4% as of 2018 and similarly low in the study area⁵⁸. We decided to target non-LPG users to isolate the phenomenon of continued consumption (as opposed to conflating continued use with the upfront cost of the stove that existing LPG users would have paid for already). At baseline in 2021, our respondents reported an average monthly expenditure of US\$182 (or US\$462.28 at PPP). Most individuals engaged in subsistence agriculture, raising goats or cattle or pursuing small business.

Procedures

After establishing eligibility, explaining the study and obtaining consent, we conducted a baseline survey. This baseline collected demographic, socio-economic and energy information. Within 2 weeks, we delivered two 6-kg LPG cylinder/burner sets and trained the main cook on how to safely use the sets. We interviewed LPG suppliers in Shirati when we piloted the study, and they reported that LPG supply was not a concern. However, this was before the trial. During the study, no participant ever reported failing to purchase a cylinder due to a supply shortage.

The every-other-week follow-up surveys obtained information on the user's cylinder purchases, the lag time between cylinder refills and their cooking fuel use. During these visits, we also confirmed that both LPG cylinders were still in the home even if empty. A few households asked to give one away to their relative or friend because they were rationing LPG and only needed one cylinder/burner set. We asked that they keep both at least until the end of the trial. We had survey

enumerators take photos of all cylinders (and all stoves for that matter) at each visit and report the number of cylinders within the home and those being used. Establishing rapport was of the utmost importance to our research and field team. The local women we hired to work in their own communities assured participants that the cylinders were theirs to keep and just asked that they keep both until the end of the year-long study. Finally, we conducted an endline survey that mirrored the baseline. Each visit lasted about 30 min; our work was made possible by the labour of the women working for and participating in the study.

Supervisors of the local team delivered a locked deposit box with a key that respondents could choose to keep or give to their local outreach worker, whom we titled Community Technology Workers (CTWs). The CTWs encouraged the use of the lockbox to save for LPG refills. We did not, however, seek to commit participants to saving only for LPG. Most respondents chose to keep the key with the CTW (Supplementary Figs. 5 and 6). The first author worked on the ground with the local team throughout the entire study and held weekly meetings where logistics or any safety/adverse events could be addressed (although none arose).

Outcome construction

Pre-specified primary outcomes were purchased refills, the days between running out of LPG and purchasing the next cylinder ('the refill gap'), categories of LPG use (exclusive, habitual, occasional and none) and biomass (firewood/charcoal) usage³³. We counted purchased refills for each participant over their times in control and treatment, respectively. Respondents self reported the day that one or both cylinders ran out and the date on which they refilled one or both cylinders. This allowed us to calculate the refill gap. We categorized each participant at each visit into one of four LPG categories based on their responses to numerous questions, indicating whether they were exclusively, habitually or occasionally using LPG or had failed to refill for an extended period. We prescribed how these categories would be defined in Appendix 6 of our pre-analysis plan (PAP). For example, if a user was categorized as an Exclusive LPG User if the number of meals prepared and the number of meals prepared with gas were equal, the user reported only using LPG, the user reported no firewood or charcoal use in the past 7 days and yesterday and neither the firewood nor charcoal stove is warm. Respondents reported their firewood or charcoal consumption in typically purchased quantities; we translated these into kilograms after weighing numerous bundles/bucket sizes.

Sample calculation

We overpowered the sample size to provide higher than 80% power to detect a 10% difference in intervention versus control for the rarest outcome we tracked within the entire study, exclusive LPG use, as detailed in our PAP³³. We pursued individual randomization at the level of the main cook (our independent unit of analysis) after finding similar coverage and standard errors across our simulations of individual and cluster/community randomization. These simulations were conducted including and ignoring spillover effects and with different levels of contamination and intra-cluster correlation coefficients (PAP Appendix 4). We timed the 'steps' in the step-wedge design to start at week 18 until 30 to balance the number of control and treatment observations. Every participant contributed on average 25 observations (of 26 possible). We had 6,605 control and 6,221 treatment observations. We define attrition as an individual whom we were not able to contact again or who asked to not be involved in the study anymore. We define missingness as a participant whom we were unable to contact during a 2-week period, whether that be because they were travelling or had conflicting schedules, but we were able to survey again; they continued in the study.

Statistical analysis

We describe the mean and percentage, range or standard deviation of baseline demographic and socio-economic characteristics

of participants. We conducted a balance test on key covariates with weights for how many time periods each household contributed to control and treatment respectively to ensure balance across control and treatment observations (Supplementary Table 1).

We assess treatment impact on refills by constructing a user's rate of refilling while in control versus treatment (total refills in control or treatment divided by cumulative time in control or treatment). For all other outcomes, we analyse immediate and cumulative impacts of offering participants a saving lockbox. For all, we investigated the immediate treatment impact, which refers to the outcome for the first 2-week time period after entering into treatment. For all, we estimate the average treatment effect (ATE) using logistic regression, which indicates the difference in the mean outcomes between treatment and control groups.

In an intent-to-treat analysis, we modelled immediate impact of treatment on binary and continuous outcomes using logistic and linear regressions, respectively. We model the rate of refilling over the time periods in control vs treatment using Poisson regression. For binary outcomes, we report robust inference through the bootstrap method, clustered at the level of the participant. For continuous outcomes, we report robust inference, clustered at the level of the participant.

We also estimated the cumulative treatment impact by finding average impact of each additional 2-week time period in treatment. The initial impact may change over time, as users may adapt to better saving habits or drop off.

To do so, we model the outcomes against the cumulative time in treatment (9 to 18 2-week periods) with a logistic or linear regression (if binary/continuous) and predict the average. For binary outcomes, we then model this average with a linear regression against all possible time in treatment. That is, we find the additional treatment effect for one more period of being treated. We report robust inference through the bootstrap method, clustered at the level of the participant.

We control for time since the study began to isolate the effect of treatment from the effect of time as cooking fuel needs can change seasonally and participants over time may dig themselves into a financial LPG hole before even reaching treatment status. This was not pre-specified, so we also include the analysis without controlling for time since the study began in Supplementary Information.

Although not pre-specified, as a secondary investigation, we repeated all analyses for sub-groups of per capita expenditure: the first quartile (low), the second and third quartiles (mid) and the fourth quartile (high).

As additional sensitivity analysis for the analysis not controlling for time, we repeat the intent-to-treat immediate impact analyses that does not control for time using double-robust, targeted maximum likelihood estimation (Supplementary Tables 9 and 10). This estimation technique is based on both the calculation of the propensity score and the parametric model specified; if either is correct, the model is robust to bias. We partly modelled this analysis off of cluster RCT evaluated HIV treatment in Uganda (although we pursue individual randomization)⁵⁹. Although not necessary in an RCT as treatment was randomized, we pursued this approach as sensitivity analysis as it does not rely on any parametric assumptions about the distribution of the underlying data and allow for the adjustment of covariates to increase efficiency without requiring any additional modelling assumptions^{60,61}.

We conducted a complete case analysis as overall missingness was <5% of observations (Supplementary Figs. 7 and 8). We had -1% ($n = 7$) attrition (for example, participants left the study) and missing observations (that is, field team could not contact the participant that specific week) never exceeded 6% for a single time period (Supplementary Fig. 7). No analysis was conducted until after the end of data collection. All analyses were conducted in STATA 16.1 and R studio 4.2.1. Data were collected with SurveyCTO 2.71.4. The study protocol was approved by the University of California, Berkeley's Institutional Review Board (protocol 2020-02-13013) and by Tanzania's COSTECH (permit number 2021-465-NA-2021-112

(renewed as 2022-886-NA-2021-112). The study PAP is available at <https://www.socialscisearch.org/trials/8465>. Full replication files for this study are available online, as is the deidentified data.

Qualitative methods

Following the full trial, we conducted a 6-month period of qualitative work that included observations, focus groups, budgetary games and 90 semi-structured interviews. We collected quantitative and qualitative data sequentially⁶² to understand the reasons behind the respondents' reported saving and cooking fuel consumption practices. We conducted all qualitative work after the experiment had ended so as to not interfere with the experiment. Further, to align with best practice⁶², no analysis was conducted until after the end of data collection. Therefore, to structure the interviews to understand the results, we conducted all qualitative work after the end of the quantitative data collection. We used our qualitative results to help interpret the quantitative survey results. Although the trial had ended before the interviews started, participants were allowed to keep both cylinders and their lockboxes. Therefore, the present tense used by interviewees is indicative of the continued use of both the lockbox and LPG.

After data collection from the survey instrument documenting LPG use and saving patterns ended in March 2023, we randomly selected interview respondents across the RCT participants, stratified by expenditure groups (low, medium and high) and LPG categories (exclusive, habitual, occasional and no LPG use) (Supplementary Table 1). We slightly overweighted smaller categories and conducted interviews in June–July of 2023 and then again in January 2024. The first author and our main research assistant conducted interviews in Swahili; however, if the participant only spoke Luo, the research assistant led the interview. We conducted a total of 90 interviews by which time we had reached qualitative saturation, wherein new interviews were not revealing new insights.

We conducted four focus groups in May–June 2023 and January 2024 to better understand how respondents spent and saved within their household units across key expenditure categories. Each group consisted of five or six women from a representative sample across villages and expenditure groups. We conducted two more focus groups to present and ground truth our findings (July 2023) with a random selection of participants and with our survey enumerators, who helped us to better understand household responses (January 2024).

We asked all participants for verbal consent to audio record and also hand-wrote notes during focus groups and interviews. The interviews were transcribed and translated in the weeks following their completion. We performed initial and intermediate coding from interview notes and advanced coding for all transcripts in Dedoose 9.0.107, a qualitative data analysis software. Within this platform, we coded and kept track of frequency of codes to ensure that key themes were representative and comprehensive of all -90 interviews and focus groups. We discussed emergent themes, category identification and final codes/themes throughout the fieldwork and while analysing the results.

Limitations

Our work may be vulnerable to self-selection bias. We believe this to be minimal as participants did not self select into our programme but rather were chosen randomly from a spatial random sampling approach. Further, only -5% ($n = 24$) approached individuals did not agree to participate.

As with all participant-reported data (survey or interview), our results are vulnerable to recall and social desirability bias. We protected against recall and social desirability bias by asking about the previous day's cooking activities, recent practices saving and purchasing patterns, building rapport and trust with our participants, and triangulating among receipts, photos of all stoves in use and any biomass stoves, survey answers and interview responses.

Our work was also vulnerable to the risk of contagion where control participants could hear of the treatment and obtain their own

lockbox. We modelled this contagion in our PAP to inform our decision to not cluster our intervention; there was minimal impact on our ability to detect an effect (Methods, Sample calculation and our PAP).

To make our interviews and focus groups as representative as possible, we randomly sampled across LPG categories and expenditure groups from our trial. In the focus groups, we ensured that individuals were representative of the different LPG categories and expenditure categories. We do not, however, present these as experimental results but rather as context and explanation to understand our experimental results. We waited until the experiment was over to interview respondents on their savings habits so as to not change their behaviour mid-experiment and affect our internal validity.

Inclusion and ethics

The research has included local researchers throughout the research process. The first author built a relationship with the author based at the University of Dar es Salaam in 2016 before she started any research in Tanzania. The local implementation team was included in designing and implementing the study from before the pre-pilot work to ensure it was locally relevant. Study participants from an earlier feasibility pilot for the concept of community technology workers even became enumerators for the larger study, wanting to shape the new study based on their experience. The study protocol was approved by the University of California, Berkeley's Institutional Review Board (protocol 2020-02-13013) and by Tanzania's COSTECH (permit number 2021-465-NA-2021-112 (renewed as 2022-886-NA-2021-112)). The study posed minimal risk to participants. We have taken local and regional research into account in our citations.

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

The deidentified data for this study are available via Github at <https://github.com/agillwiehl/Deconstructing-the-un-affordability-of-clean-cooking-fuels-evidence-from-a-randomized-trial-in-rur> and figshare at <https://doi.org/10.6084/m9.figshare.28327115> (ref. 63).

Code availability

Full replication files for this study are available via Github at <https://github.com/agillwiehl/Deconstructing-the-un-affordability-of-clean-cooking-fuels-evidence-from-a-randomized-trial-in-rur>.

References

1. Abbafati, C. et al. Global burden of 87 risk factors in 204 countries and territories, 1990–2019: a systematic analysis for the global burden of disease study 2019. *Lancet* **396**, 1223–1249 (2020).
2. Pope, D. et al. Are cleaner cooking solutions clean enough? A systematic review and meta-analysis of particulate and carbon monoxide concentrations and exposures. *Environ. Res. Lett.* **16**, 083002 (2021).
3. *The Energy Progress Report SDG7* (IEA, IRENA, UNSD, WB & WHO, 2023).
4. Gill-Wiehl, A., Ray, I. & Kammen, D. Is clean cooking affordable? A review. *Renew. Sustain. Energy Rev.* **151**, 111537 (2021).
5. Cabiyo, B., Ray, I. & Levine, D. I. The refill gap: clean cooking fuel adoption in rural India. *Environ. Res. Lett.* **16**, 014035 (2021).
6. Masera, O. R., Saatkamp, B. D. & Kammen, D. M. From linear fuel switching to multiple cooking strategies: a critique and alternative to the energy ladder model. *World Dev.* **28**, 2083–2103 (2000).
7. Hollada, J. et al. Perceptions of improved biomass and liquefied petroleum gas stoves in Puno, Peru: implications for promoting sustained and exclusive adoption of clean cooking technologies. *Int. J. Environ. Res. Public Health* **14**, 182 (2017).
8. Shankar, A. V. et al. Everybody stacks: lessons from household energy case studies to inform design principles for clean energy transitions. *Energy Policy* **141**, 111468 (2020).
9. Mobarak, A. M., Dwivedi, P., Bailis, R., Hildemann, L. & Miller, G. Low demand for nontraditional cookstove technologies. *Proc. Natl Acad. Sci. USA* **109**, 10815–10820 (2012).
10. Kar, A., Pachauri, S., Bailis, R. & Zerriffi, H. Using sales data to assess cooking gas adoption and the impact of India's Ujjwala programme in rural Karnataka. *Nat. Energy* **4**, 806–814 (2019).
11. Collins, D., Morduch, J., Rutherford, S. & Ruthven, O. *Portfolios of the Poor* (Princeton Univ. Press, 2010).
12. Mani, S., Jain, A., Tripathi, S. & Gould, C. F. The drivers of sustained use of liquefied petroleum gas in India. *Nat. Energy* **5**, 450–457 (2020).
13. Mullainathan, S. & Shafir, E. *Scarcity: Why Having Too Little Means So Much* (Times Books, 2013).
14. DellaValle, N. People's decisions matter: understanding and addressing energy poverty with behavioral economics. *Energy Build.* **204**, 109515 (2019).
15. Hsu, E. et al. Microfinance for clean cooking: what lessons can be learned for scaling up LPG adoption in Kenya through managed loans? *Energy Policy* **154**, 112263 (2021).
16. Alem, Y., Ruhinduka, R., Berck, P. & Bluffstone, R. *Final Report Credit, LPG Stove Adoption and Charcoal Consumption: Evidence from a Randomised Controlled Trial Credit, LPG Stove Adoption and Charcoal Consumption: Evidence from a Randomised Controlled Trial* (International Growth Centre, 2015).
17. Pattanayak, S. K. et al. Experimental evidence on promotion of electric and improved biomass cookstoves. *Proc. Natl Acad. Sci. USA* **116**, 13282–13287 (2019).
18. Ashraf, N., Karlan, D. & Yin, W. Tying Odysseus to the mast: evidence for a commitment savings product in the Philippines. *Q. J. Econ.* **121**, 635–672 (2006).
19. McConnell, M. *Between Intention and Action: An Experiment on Individual Savings* Working Paper (The Abdul Latif Jameel Poverty Action Lab, 2012).
20. Dupas, P., Karlan, D., Robinson, J. & Ubfal, D. Banking the unbanked? Evidence from three countries. *Am. Econ. J.: Appl. Econ.* **10**, 257–297 (2018).
21. Beaman, L., Karlan, D. & Thuysbaert, B. Saving for a (not so) rainy day: a randomized evaluation of savings groups in Mali. *Natl Bur. Econ. Res.* **53**, 1689–1699 (2013).
22. Ksoll, C., Lilleør, H. B., Lønborg, J. H. & Rasmussen, O. D. Impact of village savings and loan associations: evidence from a cluster randomized trial. *J. Dev. Econ.* **120**, 70–85 (2016).
23. Dupas, P. & Robinson, J. Why don't the poor save more? Evidence from health savings experiments. *Am. Econ. Rev.* **103**, 1138–1171 (2013).
24. Berry, J., Karlan, D. & Pradhan, M. The impact of financial education for youth in Ghana. *World Dev.* **102**, 71–89 (2018).
25. Karlan, D. et al. *Loose Knots: Strong Versus Weak Commitments to Save for Education in Uganda* Working Paper 19863 (National Bureau of Economic Research, 2014); <http://www.nber.org/papers/w19863>
26. Steinert, J. I., Vasumati Satish, R., Stips, F. & Vollmer, S. Commitment or concealment? Impacts and use of a portable saving device: evidence from a field experiment in urban India. *J. Econ. Behav. Organ.* **193**, 367–398 (2022).
27. Loewenstein, G. & Chater, N. Putting nudges in perspective. *Behav. Public Policy* **1**, 26–53 (2017).
28. DellaValle, N. & Sareen, S. Nudging and boosting for equity? Towards a behavioural economics of energy justice. *Energy Res. Soc. Sci.* **68**, 101589 (2020).

29. Thaler R. H., Sunstein C. R. *Nudge: Improving Decisions About Health, Wealth, and Happiness* (Penguin Books, 2008).
30. Duflo, E., Kremer, M. & Robinson, J. Nudging farmers to use fertilizer: theory and experimental evidence from Kenya. *Am. Econ. Rev.* **101**, 2350–2390 (2011).
31. Thaler, R. H. & Benartzi, S. Save more tomorrow™: using behavioral economics to increase employee saving. *J. Political Econ.* **112**, S164–S187 (2004).
32. Bertrand, M., Mullainathan, S. & Shafir, E. A behavioral-economics view of poverty. *Am. Econ. Rev.* **94**, 419–423 (2004).
33. Gill-Wiehl, A., Ray, I., Hubbard, A. E., Levine, D. I. & Kammen, D. M. Nudging towards micro-savings: a step-wedge experiment on LPG adoption in rural Tanzania. *AEA RCT Registry* <https://doi.org/10.1257/rct.8465-1.0> (2022).
34. *The 2017–2018 Household Budget Survey-Dataset* (Tanzania National Bureau of Statistics, 2018).
35. Guyer, J. in *A Home Divided: Women and Income in the Third World* (eds Dwyer, D. & Bruce, J.) 155–172 (Stanford Univ. Press, 1988).
36. Guyer, J. *Family and Farm in Southern Cameroon* (Boston Univ., 1984); <https://open.bu.edu/handle/2144/23138>
37. Peters P. E. *Dividing the Commons: Politics, Policy, and Culture in Botswana* (Univ. Press of Virginia, 1994); <https://books.google.com/books?id=sIUrIgdgBLUC>
38. Angelou, N. & Bhatia, M. *Beyond Connections: Energy Access Redefined* (Energy Sector Management Assistance Program, 2015).
39. Jewitt, S., Atagher, P. & Clifford, M. “We cannot stop cooking”: stove stacking, seasonality and the risky practices of household cookstove transitions in Nigeria. *Energy Res. Soc. Sci.* **61**, 101340 (2020).
40. Gill-Wiehl, A. & Ray, I. Affording a clean stack: evidence from cookstoves in urban Kenya. *Energy Res. Soc. Sci.* **105**, 103275 (2023).
41. Khavari, B., Ramirez, C., Jeuland, M. & Fuso Nerini, F. A geospatial approach to understanding clean cooking challenges in sub-Saharan Africa. *Nat. Sustain.* **6**, 447–457 (2023).
42. Poblete-Cazenave, M., Pachauri, S., Byers, E., Mastrucci, A. & van Ruijven, B. Global scenarios of household access to modern energy services under climate mitigation policy. *Nat. Energy* **6**, 824–833 (2021).
43. Shupler, M. et al. Pay-as-you-go liquefied petroleum gas supports sustainable clean cooking in Kenyan informal urban settlement during COVID-19 lockdown. *Appl. Energy* **292**, 116769 (2021).
44. Ashraf, N. Spousal control and intra-household decision making: an experimental study in the Philippines. *Am. Econ. Rev.* **99**, 1245–1277 (2009).
45. Anderson, S. & Baland, J. M. The economics of roscas and intrahousehold resource allocation. *Q. J. Econ.* **117**, 963–995 (2002).
46. Lewis, J. J. & Pattanayak, S. K. *Who adopts improved fuels and cookstoves? A systematic review.* *Environ. Health Perspect.* **120**, 637–645 (2012).
47. Beltramo, T., Blalock, G., Harrell, S., Levine, D. & Simons, A. M. The effects of fuel-efficient cookstoves on fuel use, particulate matter, and cooking practices: results from a randomized trial in rural Uganda. *UC Berkeley: Center for Effective Global Action* <https://escholarship.org/uc/item/365778pn> (2019).
48. Sen, A. in *Collective Choice and Social Welfare* vol 11 (ed. Sen, A.) 173–186 (North-Holland, 1979); <https://doi.org/10.1016/B978-0-444-85127-7.50026-5>
49. Peter, F. Gender and the foundations of social choice: the role of situated agency. *Feminist Econ.* **9**, 13–32 (2003).
50. Gill-Wiehl, A., Brown, T. & Smith, K. The need to prioritize consumption: a difference-in-differences approach to analyze the total effect of India’s below-the-poverty-line policies on LPG use. *Energy Policy* **164**, 112915 (2022).
51. Jeuland, M. et al. A randomized trial of price subsidies for liquefied petroleum cooking gas among low-income households in rural India. *World Dev. Perspect.* **30**, 100490 (2023).
52. Witinok-Huber, R. et al. Impact of randomly assigned “pay-as-you-go” liquefied petroleum gas prices on energy use for cooking: experimental pilot evidence from rural Rwanda. *Energy Sustain. Dev.* **80**, 101455 (2024).
53. Troncoso, K. & Soares da Silva, A. LPG fuel subsidies in Latin America and the use of solid fuels to cook. *Energy Policy* **107**, 188–196 (2017).
54. Kreider, B., Pepper, J. V. & Roy, M. Identifying the effects of WIC on food insecurity among infants and children. *South. Econ. J.* **82**, 1106–1122 (2016).
55. Post, A. & Ray, I. *Hybrid Modes of Urban Water Delivery in Low- and Middle-Income Countries* (2020); <https://doi.org/10.1093/acrefore/9780199389414.013.679>
56. *2002 Population and Housing Census: Volume I-Methodology Report* (Central Census Office, National Bureau of Statistics, 2003).
57. *Age and Sex Distribution Report: Tanzania* (United Republic of Tanzania, 2022); <https://sensa.nbs.go.tz/publication/report7.pdf>
58. Gill-Wiehl, A., Sievers, S. & Kammen, D. M. The value of community technology workers for LPG use: a pilot in Shirati, Tanzania. *Energy Sustain. Soc.* **12**, 5 (2022).
59. Havlir, D. V. et al. HIV testing and treatment with the use of a community health approach in rural Africa. *N. Engl. J. Med.* **381**, 219–229 (2019).
60. Moore, K. L. & van der Laan, M. J. Covariate adjustment in randomized trials with binary outcomes: target maximum likelihood estimation. *Stat. Med.* **28**, 39–64 (2009).
61. Moore, K. L., Neugebauer, R., Valappil, T. & Laan, M. J. Robust extraction of covariate information to improve estimation efficiency in randomized trials. *Stat. Med.* **30**, 2389–2408 (2011).
62. Shaffer P. *Q-Squared: Combining Qualitative and Quantitative Approaches in Poverty Analysis* (Oxford Univ. Press, 2013); <https://doi.org/10.1093/acprof:oso/9780199676903.001.0001>
63. Gill-Wiehl, A. Deconstructing the (un)affordability of clean cooking fuels: evidence from a randomized trial in rural Tanzania. *figshare* <https://doi.org/10.6084/m9.figshare.28327115> (2025).

Acknowledgements

This work was funded by the National Science Foundation’s Graduate Research Fellowship Program (A.G.-W.), the Global Distribution Fund (A.G.-W.), University of California, Berkeley’s Center for African Studies’ Rocca Dissertation Grants (A.G.-W.), the James and Katherine Lau Climate Equity Fellowship (A.G.-W. and D.M.K.), the Foreign Language and Areas Studies Fellowship (A.G.-W.). We thank the women of Shirati for their incredible generosity in participating in this study.

Author contributions

A.G.-W., I.R. and A.H. conceived and designed the experiments. A.G.-W. performed the experiments. A.G.-W. and A.H. analysed the data. R.K., D.M.K. and A.H. contributed materials and analysis tools. A.G.-W. and I.R. wrote the paper. R.K., D.M.K. and A.H. contributed materials and analysis tools.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41560-025-01778-w>.

Correspondence and requests for materials should be addressed to Annelise Gill-Wiehl.

Peer review information *Nature Energy* thanks the anonymous reviewers for their contribution to the peer review of this work.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.

© The Author(s), under exclusive licence to Springer Nature Limited 2025

Reporting Summary

Nature Portfolio wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Portfolio policies, see our [Editorial Policies](#) and the [Editorial Policy Checklist](#).

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

- | n/a | Confirmed |
|-------------------------------------|--|
| <input type="checkbox"/> | <input checked="" type="checkbox"/> The exact sample size (n) for each experimental group/condition, given as a discrete number and unit of measurement |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly |
| <input type="checkbox"/> | <input checked="" type="checkbox"/> The statistical test(s) used AND whether they are one- or two-sided
<i>Only common tests should be described solely by name; describe more complex techniques in the Methods section.</i> |
| <input type="checkbox"/> | <input checked="" type="checkbox"/> A description of all covariates tested |
| <input type="checkbox"/> | <input checked="" type="checkbox"/> A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons |
| <input type="checkbox"/> | <input checked="" type="checkbox"/> A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals) |
| <input type="checkbox"/> | <input checked="" type="checkbox"/> For null hypothesis testing, the test statistic (e.g. F , t , r) with confidence intervals, effect sizes, degrees of freedom and P value noted
<i>Give P values as exact values whenever suitable.</i> |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings |
| <input type="checkbox"/> | <input checked="" type="checkbox"/> For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes |
| <input checked="" type="checkbox"/> | <input type="checkbox"/> Estimates of effect sizes (e.g. Cohen's d , Pearson's r), indicating how they were calculated |

Our web collection on [statistics for biologists](#) contains articles on many of the points above.

Software and code

Policy information about [availability of computer code](#)

Data collection

Data analysis

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Portfolio [guidelines for submitting code & software](#) for further information.

Data

Policy information about [availability of data](#)

All manuscripts must include a [data availability statement](#). This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A description of any restrictions on data availability
- For clinical datasets or third party data, please ensure that the statement adheres to our [policy](#)

Data is available at <https://github.com/agillwiehl/Deconstructing-the-un-affordability-of-clean-cooking-fuels-evidence-from-a-randomized-trial-in-rur> and <https://doi.org/10.6084/m9.figshare.28327115>

Research involving human participants, their data, or biological material

Policy information about studies with [human participants or human data](#). See also policy information about [sex, gender \(identity/presentation\), and sexual orientation](#) and [race, ethnicity and racism](#).

Reporting on sex and gender

The gender of participants was determined based on self-reporting. In adjusted analyses, we control for gender as a covariate.

Reporting on race, ethnicity, or other socially relevant groupings

Participants self-reported their classification of different tribes (e.g., Luo).

Population characteristics

We conducted this work in Shirati, Rorya District, Mara Region, Tanzania. Shirati is a rural town of ~50,000 on the edge of Lake Victoria near the Kenyan border. Shirati has a tropical climate with light rains from October to December and heavy rains from March to June. According to the 2022 census, when our study was completed, the average number of “persons who spent the census night” (pg. 77), living together, in the Mara region was 5 individuals. In Mara, 42.3% of households were female headed, but over half of adults were married (56%). According to the Tanzania’s 2018 Household Budget Survey, most individuals in Mara were 47 years old, had completed primary school (62%); many used a rechargeable lamp (35.3%) or solar lamps (26.6%) as their main source of lighting. Average monthly household consumption was 207 USD in 2017/2018. Female heads of households were self-employed (including farming, petty trade, etc.), 14% were unemployed, never worked, or didn’t know, and 17.5% were unpaid household workers. Only 11.6% of households in the region had at least one member (typically not a woman) with a bank account. We note that these characteristics are for the region and of all individuals, not only main cooks (our sample). In Mara, most used firewood (72.9%) or charcoal (23.5%) for cooking, although LPG was available. In Shirati, firewood and charcoal could be collected and purchased in every village and charcoal-making was a common small business. For LPG users, 6 or 12 kg cylinders could be exchanged at retail points in the village with the main market, Obwere. In order to refill, participants had to either walk with the heavy cylinder or pay the extra cost of a motorcycle or three-wheeled delivery cart (~1 USD).

Recruitment

We invited the main cook from the closest homestead to the spatially randomized point to participate. After a few days to consider our request, we returned to obtain written informed consent. The only eligibility criterion for our selected participants (Supplemental Figure 1) was that they did not already use LPG for cooking. This was not a limiting criterion as LPG use in the Mara region was 1.4% as of 2018 and similarly low in the study area. Our work may be vulnerable to self-selection bias; however, we believe this to be minimal as participants did not self select into our program, but rather were chosen randomly from a spatial random sampling approach and only ~5% (n=24) approached individuals did not agree to participate.

Ethics oversight

The study protocol was approved by the University of California, Berkeley’s Institutional Review Board (Protocol 2020-02-13013) and by Tanzania’s COSTECH (Permit # 2021-465-NA-2021-112 (renewed as 2022-886-NA-2021-112).

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

Life sciences Behavioural & social sciences Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see [nature.com/documents/nr-reporting-summary-flat.pdf](https://www.nature.com/documents/nr-reporting-summary-flat.pdf)

Behavioural & social sciences study design

All studies must disclose on these points even when the disclosure is negative.

Study description

In a year-long stepped-wedge randomized control trial in Tanzania (n=511), we test using quantitative and qualitative methods whether, by how much, and how a lockbox intervention paired with micro-saving nudges could alleviate the unaffordability of clean fuels.

Research sample

Our research sample was main cooks in Shirati, Tanzania. We pursued a spatial randomly sampling procedure, as the Tanzanian census was outdated as of 2021. We relied on a combination of ward level shapefiles, Google Maps, interviews with village leaders, local knowledge, and Python to randomly place ~511 points across Shirati and invite the main cook from the closest household. Our sample is mostly female (94%), 41 years old, from households of roughly 6 people. Most rely on farming or petty trade for income, having gone to primary school (88%). Most respondents had a weekly expenditure of 17.2 (s.d.: 10.4) USD. Slightly over half the respondents were able to save at all (56%). This sample is largely representative of Mara region, according to the 2018 data we have. We intentionally targeted main cooks and a rural settings. Main cooks are the most involved household member for cooking, and we needed to safely train individuals on LPG; however, there were no restrictions or instructions on who could deposit into the lockbox. Further, rural communities are disproportionately understudied, and are often targeted for improved, but not (truly) clean stoves. We decided to target non-LPG users in order to isolate the phenomenon of continued consumption (as opposed to conflating continued use with the upfront cost of the stove, that existing LPG users would have paid for already).

Sampling strategy	We overpowered the sample size to provide higher than 80% power to detect a 10% difference in intervention versus control for the rarest outcome we tracked within the entire study, exclusive LPG use, as detailed in our pre-analysis plan (PAP). We pursued individual randomization at the level of the main cook (our independent unit of analysis) after finding similar coverage and standard errors across our simulations of individual and cluster/community randomization. These simulations were conducted including and ignoring spillover effects, and with different levels of contamination and intra-cluster correlation coefficients (see PAP Appendix 4). We timed the “steps” in the step-wedge design to start at week 18 until 30 to balance the number of control and treatment observations. Every participant contributed on average 25 observations (of 26 possible). We had 6605 control and 6221 treatment observations. For the qualitative analysis, we randomly selected interview respondents across the SW-RCT participants, stratified by expenditure groups (low, medium, and high) and LPG categories (exclusive, habitual, occasional, and no LPG use). We slightly overweighted smaller categories. We conducted a total of 90 interviews, by which time we had reached qualitative saturation, wherein new interviews were not revealing new insights.
Data collection	Local outreach workers, whom we titled Community Technology Workers (CTWs), collected survey data on phones through the SurveyCTO application. The CTWs asked participants to be interviewed alone and was not blind to intervention status as questions about the lockbox would not appear for control participants.
Timing	Participants were enrolled in January 2022. The year-long trial lasted between February 2022 to March 2023. We conducted four focus groups in May-July 2023 and January 2024. All interviews were conducted in June and July 2024.
Data exclusions	No data was excluded from the analysis.
Non-participation	Only ~5% (n=24) approached individuals did not agree to participate (non-participation). Overall missingness was <5% of observations (Figure S7/S8). We had ~1% (n=7) attrition (e.g., participants left the study), and missing observations (i.e., field team could not contact the participant that specific week) never exceeded 6% for a single time period.
Randomization	Treatment order was random. After baseline data collection, the first author determined treatment distribution order using a random number generator in R.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
<input checked="" type="checkbox"/>	<input type="checkbox"/> Antibodies
<input checked="" type="checkbox"/>	<input type="checkbox"/> Eukaryotic cell lines
<input checked="" type="checkbox"/>	<input type="checkbox"/> Palaeontology and archaeology
<input checked="" type="checkbox"/>	<input type="checkbox"/> Animals and other organisms
<input checked="" type="checkbox"/>	<input type="checkbox"/> Clinical data
<input checked="" type="checkbox"/>	<input type="checkbox"/> Dual use research of concern
<input checked="" type="checkbox"/>	<input type="checkbox"/> Plants

Methods

n/a	Involved in the study
<input checked="" type="checkbox"/>	<input type="checkbox"/> ChIP-seq
<input checked="" type="checkbox"/>	<input type="checkbox"/> Flow cytometry
<input checked="" type="checkbox"/>	<input type="checkbox"/> MRI-based neuroimaging

Plants

Seed stocks	<i>Report on the source of all seed stocks or other plant material used. If applicable, state the seed stock centre and catalogue number. If plant specimens were collected from the field, describe the collection location, date and sampling procedures.</i>
Novel plant genotypes	<i>Describe the methods by which all novel plant genotypes were produced. This includes those generated by transgenic approaches, gene editing, chemical/radiation-based mutagenesis and hybridization. For transgenic lines, describe the transformation method, the number of independent lines analyzed and the generation upon which experiments were performed. For gene-edited lines, describe the editor used, the endogenous sequence targeted for editing, the targeting guide RNA sequence (if applicable) and how the editor was applied.</i>
Authentication	<i>Describe any authentication procedures for each seed stock used or novel genotype generated. Describe any experiments used to assess the effect of a mutation and, where applicable, how potential secondary effects (e.g. second site T-DNA insertions, mosaicism, off-target gene editing) were examined.</i>