UNITED NATIONS UNIVERSITY
UNU-WIDER

WIDER Working Paper 2018/59

NGOs and the effectiveness of interventions

Faraz Usmani,[1] Marc Jeuland,[2] and Subhrendu K. Pattanayak[2]

May 2018

United Nations University World Institute for Development Economics Research
wider.unu.edu

Abstract: Interventions in remote, rural settings face high transaction costs. We develop a model of household decision-making to evaluate how non-governmental organizations (NGOs) address these implementation-related challenges and influence intervention effectiveness. To test our model's predictions, we create a sample of observationally similar Indian villages that differ in their prior engagement with a local development NGO. In partnership with this NGO, we then stratify a randomized technology promotion intervention on this institutional variable. We uncover a large, positive, and statistically significant 'NGO effect': prior engagement with the NGO increases the effectiveness of our intervention by at least 30 per cent. Our results have implications for the generalizability of experimental research conducted jointly with NGOs. In particular, attempts to scale-up findings from such work may prove less successful than anticipated if the role of NGOs is insufficiently understood. Alternatively, policy makers looking to scale-up could achieve greater success by enlisting trusted local partners.

Keywords: external validity, field experiments, NGOs, non-governmental organizations
JEL classification: L31, C93, O12

Acknowledgements: We thank Anna Birkenbach, Laura Grant, Esther Heesemann, David Kaczan, Brian Murray, Jennifer Orgill-Meyer, Seth Sanders, and seminar participants at AERE 2016, SEA 2016, CECFEE 2016, EAERE 2017, the University of Gothenburg, the 2017 Occasional Workshop in Environmental and Resource Economics, EfD 2017, and CSAE 2018 for valuable comments and suggestions. This study was made possible by the support of the American people through the United States Agency for International Development (USAID), and was partially funded by USAID under Translating Research into Action, Cooperative Agreement No. GHS-A-00-09-00015-00. The contents of this publication are the sole responsibility of the authors and do not necessarily reflect the views of USAID or the United States Government. Usmani gratefully acknowledges support from the Duke Global Health Institute, Duke University Energy Initiative, Sustainable Energy Transitions Initiative, and the United Nations University World Institute for Development Economics Research (UNU-WIDER). Any remaining errors are our own.

[1] Nicholas School of the Environment and Sanford School of Public Policy, Duke University, Durham, NC, USA, corresponding author: faraz.usmani@duke.edu; [2] Sanford School of Public Policy and Duke Global Health Institute, Duke University, Durham, NC, USA.

This study has been prepared within the UNU-WIDER Visiting PhD Fellowship programme.

Information and requests: publications@wider.unu.edu

ISSN 1798-7237 ISBN 978-92-9256-501-5 https://doi.org/10.35188/UNU-WIDER/2018/501-5

Typescript prepared by Gary Smith.

The United Nations University World Institute for Development Economics Research provides economic analysis and policy advice with the aim of promoting sustainable and equitable development. The Institute began operations in 1985 in Helsinki, Finland, as the first research and training centre of the United Nations University. Today it is a unique blend of think tank, research institute, and UN agency—providing a range of services from policy advice to governments as well as freely available original research.

The Institute is funded through income from an endowment fund with additional contributions to its work programme from Finland, Sweden, and the United Kingdom as well as earmarked contributions for specific projects from a variety of donors.

Katajanokanlaituri 6 B, 00160 Helsinki, Finland

1 Introduction

Implementation challenges stymie policies and programmes in virtually every sector. Transaction costs borne to identify beneficiaries and service providers, gauge their trustworthiness, bargain to reach consensus, and monitor agreements to ensure they are fulfilled can undermine the effectiveness of interventions—particularly in low-income settings in which such frictions are especially important (Holloway et al. 2000; Ito 2009; Jack and Suri 2014; Schaner 2016). Where public-service delivery has proven difficult, a variety of non-state, non-market institutions—most notably nongovernmental organizations (NGOs)—have emerged (Werker and Ahmed 2008). Indeed, NGOs increasingly play lead roles in implementation on the ground; by some estimates, India alone is home to over three million of them—a figure that outnumbers its public hospitals, schools, and police force (Anand 2015).[1] That they are (at least in theory) nimble and efficient has made them attractive partners for international donors. In 2012, for instance, Organisation for Economic Co-operation and Development (OECD) countries channelled over $17 billion of overseas development assistance to—and through—NGOs (Aldashev and Navarra 2018). Private charitable giving to international development NGOs may be considerably higher (Micklewright and Wright 2003). Yet little is known about the ways in which NGOs directly impact the effectiveness of the interventions they implement.

In this study, we develop a model of household decision-making motivated by transaction costs to evaluate how NGOs influence intervention outcomes. We then use a novel matched-experimental study design to test our model's main theoretical claim: that an NGO lowers transaction costs where it has been more active, which in turn increases intervention effectiveness. We test this claim in the context of an intervention designed to promote improved cookstoves (ICS) in rural India. Nearly three billion people globally rely on traditional stoves and solid fuels for their primary energy needs. ICS, which are designed to increase the efficiency of solid-fuel combustion, have long been seen as a potential solution to the environmental-health-development burden these energy-use patterns impose (Jeuland and Pattanayak 2012).[2] Yet widespread uptake has proven challenging (e.g. Mobarak et al. 2012) and—as in other arenas—the role played by local institutions in driving adoption remains poorly understood despite considerable scholarship on what encourages uptake (Lewis and Pattanayak 2012).

We first use ex-ante propensity score matching to create a sample of observationally similar villages that are differentiated by prior exposure to a local development NGO. In partnership with this NGO, we then randomly assign nearly 100 geographically distinct hamlets within these villages (covering a sample of almost 1,000 households) to treatment and control groups as part of an experimental ICS-promotion intervention. Our results suggest that the intervention increases adoption: nearly half of all households targeted by the promotion campaign purchased an ICS. However, our study design also allows us to identify the direct impact of the NGO on ICS adoption and energy-use patterns. We uncover a large, positive, and statistically significant 'NGO effect'—purchase rates are nearly 13 percentage points (28 per cent) higher in treated communities with prior interactions with the NGO. This finding is robust to multiple ways of defining the scope of the NGO's prior relationships with communities. Placebo tests inspired by randomization-based inferential procedures also reveal that it is highly unlikely to have been the result of the chance selection of the set of villages in our sample. Using a triple-differences specification—which further relaxes our identifying assumptions—we find that treated households in NGO communities are also 16 percentage points more likely to use intervention stoves than treated

---

[1] In contrast, the total number of international NGOs is more conservatively estimated to be closer to 30,000 (Aldashev and Navarra 2018).

[2] Reliance on traditional fuels and stoves imposes a tremendous burden on household health (via exposure to household air pollution); local forests (via unsustainable extraction of fuelwood); and the global climate (due to widespread biomass burning). Bailis et al. (2015), Jeuland and Pattanayak (2012), Jeuland et al. (2015), and Venkataraman (2005) provide an overview of the global nature and magnitude of these environmental, health, and welfare challenges.

households in communities without a prior relationship with the NGO, representing a 50 per cent increase in the size of the treatment effect. Consistent with these patterns of adoption and use, treated households in NGO communities exhibit significant reductions in the use of solid fuels and in fuel-collection times. In contrast, we find no evidence of similar improvements in energy-use patterns for treated households in non-NGO communities. Our stratified study design, therefore, reveals that we would have considerably overestimated the effectiveness of our intervention had it been a typical randomized evaluation conducted in partnership with the NGO.

These findings, thus, highlight some of the implications of the increasing roles NGOs play as partners in research, especially in recent years in response to growing calls for more 'evidence-based' policies (Banerjee et al. 2007, 2017). Charity evaluators (such as GiveWell and Giving What We Can) routinely release lists of top NGOs whose effectiveness is rigorously evaluated for would-be donors in search of causes. In the fiercely competitive world of fundraising, such signals do not go unnoticed, and NGOs are often keen to work with researchers and have the impacts of their programmes assessed. For applied researchers, partnerships with NGOs provide ready access to target populations, local expertise, and human resources and operational infrastructure. These lower the costs of doing research and create opportunities for researchers to test new theories directly. Randomized controlled trials (RCTs) in developing countries have particularly benefited from these trends. Indeed, proponents argue that 'many of the best RCTs have come from long-term partnerships between researchers and NGOs or other local partners' (Glennerster 2015). NGO–researcher collaborations undoubtedly yield valuable insights about the effectiveness of environmental, health, and development interventions (e.g. Banerjee et al. 2015; Brooks et al. 2016; Jayachandran et al. 2017; Miguel and Kremer 2004). That said, we show that this seemingly symbiotic relationship may also mask the unique roles NGOs often play in remote, rural settings—with serious implications for the generalizability and scalability of solutions deemed effective in applied research (Berge et al. 2012; Peters et al. 2018; Vivalt 2017).

More broadly, our study is motivated by calls for research that sheds light on how and why certain interventions are effective, not simply whether or not they are (Deaton 2010; Pattanayak et al. 2017; Rosenthal et al. 2017). In particular, we demonstrate how solutions from applied research that is tied to the places, populations, and programmes associated with specific NGOs may have very different impacts when implemented in alternative settings (Ravallion 2009).[3] Our study is also related to a growing literature on the role of implementer identity, which finds that NGO-led interventions are generally more effective than comparable efforts by other actors (Bold et al. 2016; Cameron and Shah 2017; Grossman et al. 2016; Henderson and Lee 2015). Yet such comparisons implicitly assume that there is something inherently different about NGOs that leads to implementation effectiveness. As we show, this is not necessarily the case, and even heretofore 'effective' NGOs struggle to overcome implementation-related barriers when forced to operate in new settings in which their stock of social capital is low. This is because effective local NGOs often spend years fostering trust in the communities in which they operate. These typically unobserved context- and NGO-specific characteristics interact with aspects of the intervention, influence transaction costs associated with implementation, and ultimately help determine final outcomes.[4]

[3] This is often referred to as 'site selection bias' (Allcott 2015) and can arise in at least two different ways. First, as Lin et al. (2012) show in the context of forestry interventions, the spatial distribution of social-welfare programmes is non-random. The presence of such programmes for the purposes of evaluation is correlated with past levels of NGO activity, which is itself strategically determined by NGOs (Brass 2012; Fruttero and Gauri 2005; MacLean et al. 2015). Second, NGOs capable of managing complex RCTs (designed to provide evidence of programme effectiveness) are also likely to implement programmes more effectively than the average implementer.

[4] For instance, if beneficiaries are unsure about the benefits and costs associated with unfamiliar welfare-improving technologies (such as ICS), trusted NGOs can leverage these relationships to reduce households' perceptions of the risks associated with technology adoption.

Nevertheless, we recognize that ours is not the only study grappling with these questions, and that there are tangible opportunities for us to incorporate lessons from this broader literature on the importance of implementer identity. To evaluate the relative strength of the evidence we uncover, we use estimates from Bold et al. (2016) and Cameron and Shah (2017)—who compare the effectiveness of an NGO-led intervention to a government-led one—to formally specify a distribution for our prior understanding of the effects NGOs might have on final outcomes. How does our study contribute to this prior knowledge? To answer this question, we use this prior distribution as a critical component of a multilevel Bayesian analysis, with which we revisit our main results. We show that a synthesis of existing insights and our results in this way yields posterior distributions of the magnitude of the 'NGO effect' that are considerably more precise. Importantly, we demonstrate that our results overwhelmingly point to the direction of this effect being positive.

This paper proceeds as follows: in Section 2, we develop our theoretical model of household decision-making in the presence of NGOs and transaction costs; Section 3 provides an overview of our data and sample-selection methods; Section 4 outlines our empirical framework and identification strategy; Section 5 presents results; Section 6 demonstrates how insights from the related literature inform our results within a Bayesian framework; and Section 7 concludes.

2 NGOs, transaction costs, and household decision-making

Transaction costs are inherent in the adoption of new technologies—especially in remote, rural settings characterized by low information (Bernard et al. 2017; Foster and Rosenzweig 2010; Suri 2011). For example, while the material costs of new technologies are usually borne immediately, benefits are often uncertain and realized far in the future; resources are thus required to learn about the valuable attributes of the technology, and gauge the full extent of its costs and benefits. In addition, the search and exchange process that technology adoption entails can impose significant costs. These are often related to the size of the market. Relatively large markets—such as major urban areas—feature a multitude of retailers and suppliers, competing along technology price, quality, and differentiability criteria. In contrast, relatively small markets—such as remote, rural settings—are characterized by weak supply chains, a paucity of suppliers, and limited options. These contextual characteristics influence households' decisions about technology adoption, particularly in smaller markets in which kinship ties, reciprocal exchange, and repeated dealings (in other words, social capital) are more salient (Kranton 1996).

These insights motivate our model of household decision-making. Drawing on Jeuland et al. (2015)—who in turn build on more fundamental work in environmental and health economics (Grossman 1972; Pattanayak and Pfaff 2009)—we develop a model in which households decide whether to invest in technologies that avert environmental health risks (such as ICS). These decisions necessarily involve a trade-off with consumption of other goods and leisure: households must maximize utility by allocating limited resources to environmental or health investments, consumption, and leisure. Accordingly, household utility u is a function of consumption c, leisure l, technology adoption a, time spent sick s, and household environmental quality e. Sickness is determined by household environmental quality, which is itself determined by the household's adoption decisions. The household's utility function is assumed to be twice differentiable, continuous, and concave.

The household faces a 'full income' constraint, whereby its exogenous income y must be allocated to consumption c as well as the materials m, time t, and knowledge k—with prices p, w, and r, respectively—required for technology adoption. Similarly, total available time T must be allocated to leisure, time spent sick, and time allocated to the technology-adoption decision.

The Lagrangian associated with the household's optimization problem is as follows:

$$\max_{a,l,c,t,m,k} \mathscr{L} = u\left[c,l,a,s(e),e(a)\right] + \lambda\left[y - c - p(\xi)\cdot m - r(\xi)\cdot k + w\left(T - s(e) - l - t(\xi)\right)\right] \tag{1}$$
$$+ \mu\left[T - l - s(e) - t(\xi)\right].$$

Note that the material and time costs associated with the technology adoption are presented in Equation (1) as a function of the activities of an NGO, denoted by ξ.[5] In addition, while NGO activity is relatively less likely to influence the price of time (the local market wage rate), it can influence the amount of time a household needs to allocate to the technology-adoption decision; accordingly, t is also a function of ξ.

The first-order conditions associated with the maximization problem outlined in Equation (1) are as follows:

$$\mathscr{L}_a = u_a + u_s s_e e_a + u_e e_a - \lambda\left[p(\xi)\cdot a_m + r(\xi)\cdot a_k + w\left(a_{t(\xi)} + s_e e_a\right)\right] - \mu\left[a_{t(\xi)} + s_e e_a\right] = 0 \tag{2}$$
$$\mathscr{L}_l = u_l - \lambda\cdot w - \mu = 0 \tag{3}$$
$$\mathscr{L}_c = u_c - \lambda = 0 \tag{4}$$
$$\mathscr{L}_t = u_a a_{t(\xi)} + u_s s_e e_a a_{t(\xi)} + u_e e_a a_{t(\xi)} - \lambda w\left(1 + s_e e_a a_{t(\xi)}\right) - \mu = 0 \tag{5}$$
$$\mathscr{L}_m = u_a a_m + u_s s_e e_a a_m + u_e e_a a_m - \lambda\left(p(\xi) + w s_e e_a a_m\right) - \mu = 0 \tag{6}$$
$$\mathscr{L}_k = u_a a_k + u_s s_e e_a a_k + u_e e_a a_k - \lambda\left(r(\xi) + w s_e e_a a_k\right) - \mu = 0 \tag{7}$$
$$\mathscr{L}_\lambda = y - c - pm - rk + w\left(T - s(e) - l - t(\xi)\right) = 0 \tag{8}$$
$$\mathscr{L}_\mu = T - l - s(e) - t \geq 0; \mu\left(T - l - s(e) - t(\xi)\right) = 0. \tag{9}$$

Assuming that all individuals work some non-zero hours, it follows from Equation (9) that $\mu = 0$. Then, from Equation (2), the optimal level of technology adoption (a^*) must fulfil

$$\underbrace{\frac{u_a + u_s s_e e_a + u_e e_a}{\lambda} - w s_e e_a}_{\text{Marginal benefit, } MB(a)} = \underbrace{p(\xi)\cdot a_m + r(\xi)\cdot a_k + w a_{t(\xi)}}_{\text{Marginal cost, } MC(a,\xi)}. \tag{10}$$

The first term on the left-hand side of Equation (10) represents the monetary value of the change in utility arising from marginal investments in welfare-improving technologies; the second term is the opportunity cost of sickness valued at the price of time (the wage rate, w). The right-hand side represents the marginal cost of investing in an additional unit of a, disaggregated by the costs associated with materials, knowledge, and time commitment necessary for technology adoption.

Let $\pi(a,\xi)$ represent household welfare (or net benefit). Then:

$$\left.\frac{\partial \pi(a,\xi)}{\partial a}\right|_{a=a^*} = MB(a) - MC(a,\xi) = 0 \tag{11}$$
$$\frac{\partial^2 \pi(a,\xi)}{\partial a^2} = \frac{\partial MB(a)}{\partial a} - \frac{\partial MC(a,\xi)}{\partial a} < 0, \tag{12}$$

where Equation (12) follows from the concavity of the household's utility function.

[5] Most fundamentally, differences in NGO activity arise because NGOs are present in certain locations and not in others. However, such differences may also arise due to the amount of time an NGO has been active in a particular community, or due to the nature and intensity of the programmes it chooses to implement there.

We further assume that:

$$p'(\xi) < 0, p''(\xi) < 0 \tag{13}$$

$$r'(\xi) < 0, r''(\xi) < 0 \tag{14}$$

$$t'(\xi) < 0, t''(\xi) < 0. \tag{15}$$

For the intuition behind the identities outlined in Equations (13)–(15), consider the roles local NGOs might play in facilitating technology adoption in remote, rural settings. Effective NGOs are often intimately familiar with local infrastructure limitations (such as access to electricity) and work to identify the technologies most suited to local contexts on behalf of their beneficiaries. This reduces households' time costs. Their status as trusted sources of relevant information may similarly reduce costs associated with the acquisition of knowledge, reducing the need for households to independently attempt to verify claims: believing that an NGO is reliable and trustworthy—the NGO's existing stock of social capital within its community—lowers the risks of adopting technologies recommended by it. Finally, NGOs may also provide beneficiaries with subsidies or discounts, lowering costs associated with the acquisition of materials directly. These subsidies or discounts need not be externally funded. For instance, an NGO's relatively larger size may allow it to exploit economies of scale associated with acquiring technologies in bulk, resulting in cost savings that it could pass onto its beneficiaries.[6]

Given these conditions, we can evaluate how NGO activity influences technology adoption directly using the implicit-function theorem as follows:

$$\frac{da}{d\xi} = -\frac{\partial^2 \pi / \partial a \partial \xi}{\partial^2 \pi / \partial a^2} \tag{16}$$

$$= -\frac{-\partial MC(a,\xi)/\partial \xi}{\partial^2 \pi / \partial a^2} \tag{17}$$

$$> 0. \tag{18}$$

The numerator in Equation (17) is positive. This follows from Equations (13)–(15), which imply that the marginal cost of technology adoption is decreasing in ξ. The negative sign of the denominator is established in Equation (12), implying that the sign of the expression in Equation (17) is positive. In other words, the welfare-maximizing level of technology adoption increases with an increase in NGO activity, all else being constant.[7]

---

[6] It is worth noting settings in which these assumptions might not hold. For instance, while an NGO's positive reputation may increase the willingness of beneficiaries to adopt solutions suggested by it, its assessment of its own 'reputation risks' (Herman et al. 2003) could lead it to restrict the set of interventions it chooses to implement. Indeed, our local implementation partner expressed reservations about implementing a potentially welfare-improving ICS-promotion intervention in the study setting, highlighting in particular the reputation risks the NGO faced should ICS-promotion activities related to our study (described in Section 4.1) be received poorly by its beneficiaries. In addition, while our implementing partner lobbied during the design phase for reductions in the baseline price of ICS that its beneficiaries would ultimately face, it is not difficult to imagine a rent-seeking NGO behaving very differently. Needless to say, there are a variety of organizations active in remote, rural settings, and we do not intend to insinuate that our model provides general principles for how these actors will behave under all conditions.

[7] The implicit-function theorem cannot be used to evaluate the impact of discrete changes in NGO activity. One may be interested, for instance, in how an NGO's decision to begin operating in a community in which it had not previously done so affects households' avoidance behaviour. To see the effects of such a change, note that an increase in ξ from ξ' to ξ^* causes a discrete reduction in marginal cost of avoidance behaviour, analogous to the continuous case. Under this condition, household welfare exhibits increasing marginal returns to ξ, as demonstrated below:

$$\begin{aligned}\frac{\partial \pi(a,\xi^*)}{\partial a} - \frac{\partial \pi(a,\xi')}{\partial a} &= (MB(a) - MC(a,\xi^*)) - \left(MB(a) - MC\left(a,\xi'\right)\right) \\ &= -MC(a,\xi^*) + MC\left(a,\xi'\right) \\ &> 0.\end{aligned}$$

Let $\arg\max \pi(a,\xi') = a'$ and $\arg\max \pi(a,\xi^*) = a^*$. Then $a^* > a'$ by the strict-monotonicity theorem (Edlin and Shannon 1998).

Our model thus yields a tractable definition of the transaction costs associated with technology adoption: the opportunity costs of allocating time, materials, and knowledge to the adoption process. In structuring risks and variability, fostering collective action, and influencing existing practices, NGOs influence the set of costs associated with investments in welfare-improving technologies—and, indeed, in participating in environmental, health, and development interventions more broadly. Recent work has looked at the roles NGOs play in enhancing monitoring and enforcement (Aldashev et al. 2015; Grant and Grooms 2017); improving public-service delivery (Devarajan et al. 2013); and working with heterogeneous beneficiaries (Bengtsson 2013). Outside of a nascent stream of research on the strategic nature of NGOs' location decisions (Brass 2012; Fruttero and Gauri 2005; MacLean et al. 2015), however, little is known about the roles that they might play in directly determining the outcomes of applied interventions. Indeed, while some have noted the presence of differential impacts across communities with and without NGO activity (e.g. Niehaus and Sukhtankar 2013; Sharma et al. 2015), to the best of our knowledge no study has rigorously examined how the presence of an effective local NGO and the specific institutional context that represents might directly influence household decision-making. In contrast, our model provides a clear, empirically testable hypothesis: where NGO activity has been higher, intervention effectiveness—in our case, in the form of uptake of ICS technologies—will also be higher. The remainder of this paper turns to an empirical evaluation of this claim.

3 Data and descriptive statistics

Our sampling frame consists of over 1,000 households across 97 geographically distinct hamlets (*toks*) located in 38 Census-delineated villages (*gram panchayats*) in the northern Indian state of Uttarakhand. In this section we (1) describe the creation of observationally equivalent groups of these 38 NGO and non-NGO villages using ex-ante propensity score matching; (2) outline the random selection of households in hamlets located within these villages; and (3) present descriptive statistics for our full sample of households.

3.1 Creation of the sample using propensity score matching

Given our interest in the influence of community-level institutional factors—as well as the fact that interventions are often initiated at the community level—we developed a sampling strategy designed to minimize the risk associated with community-level confounders influencing our analysis. Specifically, our strategy for the pre-baseline selection of communities (and distribution of households across these communities) relied on a matched-experimental design that sought to ensure sufficient variation to create the contextual strata of interest for the study, namely previous engagement with our known institutional partner (an Uttarakhand-based environmental NGO).[8]

We first conducted an exhaustive enumeration of all villages in two districts of Uttarakhand—Bageshwar and Nainital—that had previously been targeted in a programme implemented by a local development NGO.[9] In total, we identified 148 distinct villages lying in our NGO stratum, which we refer to as

[8] Pattanayak et al. (2009) present a step-by-step process of creating such a sample with an application to a similar environmental health programme in rural India. King et al. (2007) describe a large-scale application of a similar approach in Mexico to evaluate the impacts of universal health insurance.

[9] The NGO leads activities related to agriculture and forestry (promotion of sustainable agricultural practices, sustainable fodder cultivation, and promotion of culinary herbs), health (local hospitals/clinics), education (local schools), village-level groups (self-help groups, youth groups, and vocational cooperatives), and water management (watershed renewal, and spring-water recharge).

'NGO villages'. Using data on community-level characteristics for these villages from the 2001 round of the Indian Census, we next applied propensity score matching to identify observationally similar 'non-NGO villages' in which the NGO had not previously intervened (Rosenbaum and Rubin 1984).[10] The matching approach allowed us to purposely select for variation within our institutional stratum, while maintaining similarity across the strata in terms of a large number of observable Census characteristics. In the first stage of this approach, we excluded any village with fewer than 40 households from our sampling frame, and then estimated logistic regression models for the outcome of selection into a previous NGO intervention as a function of community-level Census variables.[11]

In the first stage we find that, in general, NGO villages are slightly smaller and have proportionally more Scheduled Caste members than the average village (Table B1).[12] NGO villages are also more likely to have infrastructure or village-level institutions (schools, credit societies, and bus facilities), but are more remote (further away from large towns, and with less access to paved roads and telephones) than non-NGO villages.

In the second stage, following the logit estimation, we estimated the probability of receiving a previous NGO intervention for all villages in our eligible districts. These predicted probabilities constitute the propensity score for each village. We matched NGO villages to non-NGO villages with the most similar propensity scores (allowing for replacement, and limiting matches to the support region with the greatest overlap in density of NGO and non-NGO villages). In this way, we ensured that the communities selected for our sampling frame were as similar as possible. For each model, we developed a matching routine that restricted our potential sample villages on the basis of size (greater than 40 households) and distance from the base of operations for our survey (working in sub-districts that could be reached within one day). We next eliminated the worst 10 per cent of matches on the basis of propensity score distance—a process known as 'trimming'—to ensure that pairs that are poor matches are not selected simply due to the inclusion of villages that do not happen to have good matches (Crump et al. 2009).

Finally, to draw our precise sample of matched pairs, we studied each of the individual pairs remaining after trimming in detail. Here, we paid particular attention to match quality and overall balance between the NGO and non-NGO strata with regards to key contextual factors (such as population, distance from nearby towns, and the presence of village-level groups and societies). Our final sample consisted of 38 villages (19 matched pairs of NGO and non-NGO villages). At the conclusion of our matching exercise,

---

[10]This does not mean that no NGO had ever worked in these villages, only that our local implementation partner had not. Our analyses, thus, focus on how relationships with particular implementing NGOs influence intervention effectiveness.

[11]In total, we estimated three distinct specifications of these logit models. In the first specification of this model, we included sub-district (block) fixed effects, which helped restrict the set of controls to very local villages, and omitted variables that were frequently missing in the Census data (e.g. access to bus services, tap water, and/or electricity availability characteristics). The second specification dropped these sub-district fixed effects, whereas the third specification was similar to the second except that it included the additional controls (with missing values in the Census assumed to be zero, or omitted in the case of tap-water availability). We eliminated the second specification because it was clearly less robust to the distance restrictions for our sample—that is, fewer matches were preserved when dropping the sub-districts far away from our base of operations. We also found that the quality of matches on a number of variables was greatly reduced with inclusion of the sub-district-level controls used in the first specification. Thus, our final analysis is based on the third specification (see Figure A1 and Table B1 for additional details).

[12]The Scheduled Castes and Scheduled Tribes are various historically disadvantaged or indigenous groups in India that have received official recognition as such from the Indian government.

Table 1: Post-match balance on village-level characteristics across NGO and non-NGO villages

Village-level characteristic	(1) NGO villages Mean	(2) NGO villages Std dev.	(3) Non-NGO villages Mean	(4) Non-NGO villages Std dev.	(5) Normalized difference
Area (km^2)	146.7	94.5	175.0	268.6	−0.10
Total population	386.8	136.6	376.7	130.0	0.05
Scheduled Caste† population (proportion)	0.26	0.29	0.28	0.30	−0.03
Scheduled Tribe† population (proportion)	0.0065	0.028	0.00036	0.0016	0.21
Number of primary schools	1.11	0.46	1.05	0.23	0.10
Number of middle schools	0.37	0.50	0.32	0.48	0.08
Number of health centres	0	0	0	0	–
Number of primary health centres	0	0	0	0	–
Number of telephone connections	0.26	0.56	0.42	0.51	−0.20
$\mathbb{1}$ (Bus services)	0.11	0.32	0.053	0.23	0.13
$\mathbb{1}$ (Credit societies)	0	0	0	0	–
$\mathbb{1}$ (Approach to village: paved road)	0.21	0.42	0.16	0.37	0.09
Distance from nearest town (km)	25.1	16.1	19.9	11.9	0.25
Forest area (hectares)	34.1	63.9	26.0	51.8	0.10
$\mathbb{1}$ (Tap water)	0.89	0.32	1	0	−0.32
$\mathbb{1}$ (Electricity for all purposes)	0.053	0.23	0	0	0.22
Observations	19		19		38

Notes: columns (1) and (3) show means for NGO and non-NGO villages, respectively, for each of the village-level variables from the 2001 round of the Indian Census used for our final propensity score matching exercise. Columns (2) and (4) show the respective standard errors for these means. Column (5) shows the normalized difference between the two means. †The Scheduled Castes and Scheduled Tribes are various historically disadvantaged or indigenous groups in India that have received official recognition as such from the Indian government. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Source: authors' calculations based on data from the 2001 round of the Census of India.

our matched pairs of NGO and non-NGO villages were balanced on all community-level characteristics used during our propensity-score estimation (Table 1).[13]

3.2 Sub-cluster survey samples and household surveys

Uttarakhand is one of India's least densely populated states, with terrain that gives rise to 'sub-clusters' (geographically distinct hamlets known as *toks*) within villages.[14] These hamlets typically vary in terms of cultural and socioeconomic characteristics. To maximize sample variation along these socioeconomic lines we randomly selected between two and four hamlets within each matched village for our final sample. Specifically, we determined that our survey teams would work within at least two hamlets in small villages, at least three in medium villages, and at least four in large villages (owing to population differences across these three groups of villages). Our final sample consisted of 97 hamlets.

[13]Table B2 presents additional tests of cross-sectional balance using more detailed community-level data from the 2011 round of the Indian Census. We did not use this Census round for our matching exercise as it had only been released provisionally at the time. Nevertheless, we find that NGO and non-NGO villages are balanced across multiple dimensions using these data as well, particularly once we make adjustments to account for multiple hypothesis testing using the free step-down resampling methodology of Westfall and Young (1993).

[14]Topographically, Uttarakhand is characterized by 'hilly terrain, rugged and rocky mountains, deep valleys, high peaks, swift streams and rivulets, rapid soil erosion, frequent landslides and widely scattered habitation' (Maurya 2014).

Households were randomly selected to participate in survey activities; baseline surveys occurred during the summer of 2012.[15] If household members were unavailable during the entire day of survey activities—or if they refused to participate—neighbouring households were randomly selected as replacements. Field supervisors performed household introductions and obtained informed consent, recorded GPS coordinates and elevation data, and oversaw quality-control checks in each village. A random subsample of households was also selected for detailed weighing of daily solid-fuel use.[16]

The ICS-promotion intervention (described in detail in Section 4.1) began in August 2013, with midline and endline surveys taking place around November 2013 and November 2014—the 3- and 15-month marks, respectively.

3.3 Descriptive statistics

Table 2 presents an overview of our main sample of 943 households, disaggregated by NGO and non-NGO villages.[17] The average household in our study consists of just under five members. The average household head is 54 years old and has had approximately six years of formal education. Only about one-quarter of surveyed households are headed by women, and more than half fall below the Indian poverty line.

As in many other parts of rural India, reliance on traditional stoves and fuels is practically universal.[18] In contrast, only about one-third of households own any type of improved stove; this is almost exclusively limited to liquefied petroleum gas (LPG) stoves, which are typically owned by relatively wealthy households. Awareness of modern alternatives to traditional cooking technologies is low: only about one-quarter to one-third of households profess an awareness of the existence of stoves or fuels that produce less smoke. This is not necessarily due to a lack of awareness about the harms associated with exposure to household air pollution. Indeed, half of surveyed households believe that the smoke generated by their primary stove is unsafe. Households also report spending up to two hours per day on average collecting traditional fuels for household use. In addition, around one in five households report that at least one family member suffered from a case of cough or cold in the past two weeks. Together, this suggests that the welfare burden imposed by widespread reliance on traditional energy sources is substantial.

Balance tests in Table 2 reveal that randomly selected households in NGO villages are broadly similar to their counterparts in non-NGO villages in terms of demographic and socioeconomic characteristics as well as in terms of stove ownership and use patterns. We note that NGO-village households are somewhat larger, report spending approximately 30 more minutes per day collecting traditional fuels, and report

[15]Highly variable village structures and geographic constraints created variation in the number of hamlets and households sampled in each village. A minimum of 20 surveys were completed in small villages, 30 in medium ones, and 40 in large ones. If a village was divided into distinct geographical sub-units (e.g. half the village was to the north of the main road, while the other half was to the south), the target number of surveys was split equally among these groups.

[16]This process involves asking households to collect an amount of fuelwood and other solid fuels that is slightly more than what they anticipate using over the next day. This amount is weighed by the field team, which returns approximately 24 hours later to weigh the remaining amount.

[17]We interviewed 1,063 households at baseline. We restrict our main analytical sample to the 943 households that were also located and interviewed during the midline (month 3) and endline (month 15) survey rounds.

[18]In our setting, 'traditional stoves' include traditional braziers (*angithi*), clay stoves (*mitti ka chulha*), coal/fuelwood heaters (*sagarh*), pan-shaped coal stoves, and three-stone fires. 'Improved stoves' include stoves fuelled by biogas, electricity, LPG and kerosene, and commercially available efficient biomass cookstoves. Similarly, 'traditional fuels' include crop residue, dung, fuelwood, leaves, and household waste (trash), while 'clean fuels' include biogas, electricity (for cooking), kerosene, and LPG.

Table 2: Overview of sample households

Household-level characteristic	(1) NGO villages Mean	(2) NGO villages Std dev.	(3) Non-NGO villages Mean	(4) Non-NGO villages Std dev.	(5) Normalized difference
Demographics					
Household size	5.23	2.12	4.52	1.89	0.24^{***}
Mean number of children aged five and under	0.55	0.86	0.38	0.72	0.16^{***}
Age of household head (years)	53.8	13.7	54.2	14.0	-0.023
$\mathbb{1}$ (Female-headed household)	0.25	0.44	0.27	0.44	-0.026
Socioeconomic characteristics					
Education level of household head (years)	6.23	4.51	6.14	4.62	0.014
Education level of primary cook (years)	4.66	4.32	4.55	4.62	0.018
$\mathbb{1}$ (Below poverty line)	0.58	0.49	0.58	0.49	-0.0074
Stove- and fuel-use characteristics					
$\mathbb{1}$ (Owns traditional stove)	0.99	0.12	0.97	0.16	0.061
$\mathbb{1}$ (Owns improved stove)	0.31	0.46	0.28	0.45	0.038
Traditional-stove use (minutes per day)	300.3	142.3	284.5	137.6	0.080
$\mathbb{1}$ (Used an improved stove in past week)	0.29	0.46	0.28	0.45	0.028
$\mathbb{1}$ (Uses traditional fuels)	0.99	0.11	0.97	0.16	0.073
$\mathbb{1}$ (Uses a clean fuel daily)	0.28	0.45	0.26	0.44	0.034
Traditional-fuel collection (minutes per day)	129.6	101.5	101.8	87.5	0.20^{**}
Beliefs and perceptions					
$\mathbb{1}$ (Heard of stoves that produce less smoke)	0.30	0.46	0.22	0.41	0.13^{**}
$\mathbb{1}$ (Heard of fuels that produce less smoke)	0.37	0.48	0.26	0.44	0.16^{***}
$\mathbb{1}$ (Thinks cookstove emissions are unsafe)	0.50	0.50	0.50	0.50	-0.0060
Health status					
$\mathbb{1}$ (At least one case of cough/cold in past week)	0.25	0.43	0.20	0.40	0.086
Observations	469		474		943

Notes: this table presents baseline (pre-intervention) summary statistics for 943 households that are part of the final study sample. 'Traditional stove' includes traditional braziers (*angithi*), clay stoves (*mitti ka chulha*), coal/fuelwood heaters (*sagarh*), pan-shaped coal stoves, and three-stone fires. 'Improved stove' includes stoves fuelled by biogas, electricity, LPG and kerosene, and commercially available efficient biomass cookstoves. 'Traditional fuel' includes crop residue, dung, fuelwood, leaves, and household waste (trash). 'Clean fuel' includes biogas, electricity (for cooking), kerosene, and LPG. Variables for time spent using traditional stoves per day and for time spent collecting traditional fuels per day are winsorized at the 97.5 percentile level. Missing values in variables for knowledge of stoves or fuels that produce less smoke are replaced with zero. p values associated with differences in means—reported using asterisks in column (5)—are obtained from a regression model of the form: $Y_{ij} = \beta_0 + \beta_1 \cdot \mathbb{1}\,(\text{NGO village}) + \nu_{ij}$, where Y_{ij} represents a household-level characteristic for household i in hamlet j; $\mathbb{1}\,(\text{NGO village})$ represents an indicator that equals one if hamlet j is located in an NGO village, and ν_{ij} represents a normally distributed error term. Standard errors are clustered at the hamlet level, and p values are adjusted using the free step-down resampling methodology of Westfall and Young (1993), as operationalized by Jones et al. (2018). $^{*}\ p < 0.10$, $^{**}\ p < 0.05$, $^{***}\ p < 0.01$.
Source: authors' calculations based on data collected during baseline survey activities (see Figure 1).

higher awareness of the existence of cleaner stoves and fuels. We control for these differences in all our analyses explicitly or via the inclusion of household fixed-effects.

4 Empirical framework and identification strategy

Our empirical framework combines ex-ante community-level matching (described in detail in Section 3) with a randomized intervention design and a quasi-experimental difference-in-difference-in-differences estimation approach for identification. In this section we (1) provide an overview of our experimental intervention; and (2) outline our estimation and identification strategy.

4.1 Randomized ICS-promotion intervention design

Figure 1 presents an overview of our intervention design and implementation timeline. The intervention was randomized at the level of the hamlet; roughly 70 per cent of hamlets—and, by implication, 70 per cent of households—stratified across NGO and non-NGO villages were randomly assigned to the ICS-promotion treatment group prior to the start of the intervention in August 2013. As part of ICS-promotion efforts, treated households were visited by trained enumerators who identified themselves as affiliated with our partner NGO. Treated households received a personalized demonstration of two distinct ICS technologies: the Greenway biomass cookstove, and the G-Coil electric stove. At the end of the demonstration, survey teams presented these households with an offer to purchase one or both of the stoves. This offer consisted of a financial plan (the opportunity to make payments in three instalments) combined with one of three randomized level of rebate (high, medium, or low—representing a reduction in the cost of each stove by about 2, 20, and 30 per cent, respectively).[19] These rebates were randomized at the household-level and delivered as a discount counting against the final instalment payment if a household was found to be still using the stove by that time. Households in control hamlets did not receive the intervention; survey teams visited control households at the same time as treated households to conduct follow-up surveys.

The stratified study design shown in Figure 1 enables us to compare the differential impact of the same randomized intervention delivered by the same field team professing to be affiliated with the same NGO across two institutionally distinct settings, namely communities with which the NGO had a preexisting relationship and those with which it did not. This allows us to isolate how NGOs—and the trust and social capital they foster in their local communities—influences the outcomes of interventions directly. Since our intervention is designed principally to increase uptake of cleaner cooking technologies, our main outcome of interest is the purchase rate of intervention ICS. To investigate heterogeneity in this rate across treated households located in NGO and non-NGO villages separately, we estimate the following specification:

$$Y_{ij} = \beta_0 + \beta_1 (TREATMENT_j) + \beta_2 (NGO_j) + \beta_3 (TREATMENT_j \times NGO_j) + \sum_n \beta_n X_{i,n} + \nu_{ij}, \quad (19)$$

where Y_{ij} is a binary variable that equals 1 if household i in hamlet j purchased at least one of the two intervention ICS offered during intervention activities and 0 if it did not; $TREATMENT_j$ is a binary variable that equals 1 if hamlet j is randomly assigned to the treatment group and 0 if it is assigned to the control group; NGO_j is a binary variable that equals 1 if hamlet j is located in an NGO village and 0 if it is in a non-NGO village; $X_{i,n}$ represents a set of household-level controls; and ν_{ij} is a normally distributed error term. Our coefficient of interest is β_3, which sheds light on the additional impact of the randomized ICS-promotion intervention on purchase rates in the NGO stratum of villages.

[19] The market price of the electric stove was approximately INR 1,000 (USD 20 in 2012) while that of the biomass stove was INR 1,400 (USD 28). The highest rebate amount, therefore, was around USD 6–8, depending on stove type.

Figure 1: Study design and timeline

Notes. This figure presents an overview of our intervention design and implementation timeline. Random household-level rebates were provided as a percentage of the price of the stove; the market price of the electric stove was approximately INR 1,000, while that of the biomass stove was INR 1,400. USD 1 $\approx$ INR 50 in 2012.
Source: authors' illustration.

4.2 Difference-in-difference-in-differences specification

Although the villages across the NGO and non-NGO strata are matched on 16 different community-level characteristics (Table 1), one may still be concerned that unobservable community-level differences drive either the selection of NGOs into certain villages, the selection of households into villages in the NGO stratum, or both. Identification may be threatened, for instance, by an NGO-stratum-specific factor that affects households' responsiveness to the intervention. To address this concern, we leverage the multiple rounds of our survey in a difference-in-difference-in-differences ('triple-differences') specification.

Specifically, we estimate the following model:

$$\begin{aligned}
Y_{ijt} = & \beta_4 + \beta_5 (POST_1) + \beta_6 (POST_2) \\
& + \beta_7 (TREATMENT_j \times POST_1) + \beta_8 (TREATMENT_j \times POST_2) \\
& + \beta_9 (NGO_j \times POST_1) + \beta_{10} (NGO_j \times POST_2) \\
& + \beta_{11} (TREATMENT_j \times NGO_j \times POST_1) + \beta_{12} (TREATMENT_j \times NGO_j \times POST_2) \\
& + \gamma_i + \nu_{ijt},
\end{aligned} \tag{20}$$

where Y_{ijt} represents ICS-related adoption, use, or impact for household i in hamlet j in survey round t. In Equation (20), $POST_1$ and $POST_2$ represent binary variables that are equal to 1 if data for the relevant observation were collected during the first and second follow-up survey rounds, respectively, and 0 otherwise; these variables capture time trends over our multiple survey rounds. We also include household fixed-effects (represented by γ_i) to control for unobserved household-level differences.[20] Our coefficients of interest are now β_{11} and β_{12}, which represent the additional impact of our intervention for treated households located in NGO villages relative to treated households located in non-NGO villages during the midline and endline survey rounds, respectively.

It is worth noting that the fully interacted triple-differences specification outlined in Equation (20) considerably relaxes our identifying assumptions. Identification would only be threatened by a confluence of factors—say, if hamlets in the treatment arm were located closer to urban areas; if NGO villages exhibited a greater degree of rural-to-urban migration (unobserved by us); *and* if the ICS-promotion intervention spanned a period that entailed a seasonal return of said (relatively cash-rich) urban migrants back to their homes, resulting in a time-varying shock specific to treated NGO-stratum hamlets that positively influenced households' purchase of ICS technologies. While certainly possible, we contend that this is unlikely in practice. Random allocation of hamlets to the intervention and control arms should preclude community-level characteristics in treated and untreated hamlets from differing significantly. In addition, our matching approach controls for a host of observed community-level differences between NGO- and non-NGO villages; the inclusion (and interaction) of survey-round time trends accounts for changes over time in hamlets in the intervention arm as a whole as well as in the NGO stratum specifically; and household fixed-effects soak up time-invariant unobservable differences.[21]

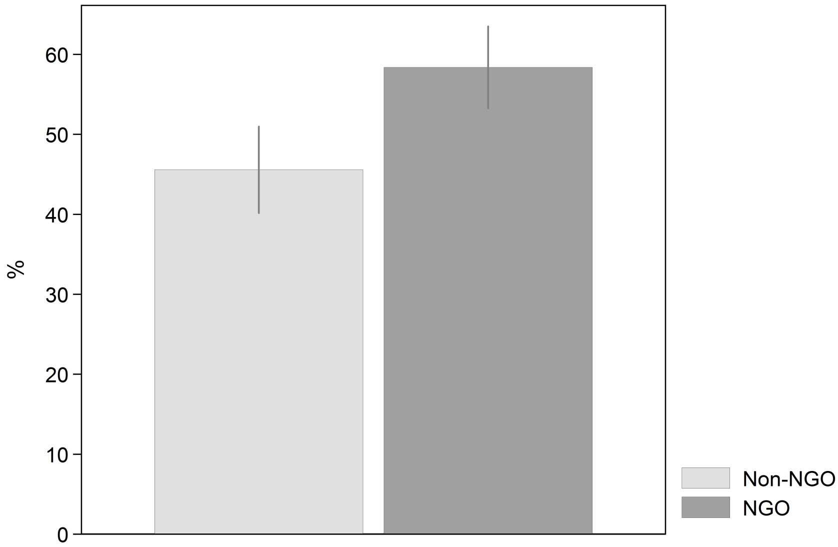
5 Results and discussion

We now turn to a discussion of the results of our empirical analyses. Our main results focus on heterogeneity in the effectiveness of the ICS-promotion intervention across matched NGO and non-NGO settings. Although our intervention is not designed to evaluate the impact of ICS in real-world settings, as part of secondary analyses we also investigate heterogeneity in impacts of ICS promotion on households' energy-use and time-allocation patterns across NGO and non-NGO communities.

[20]Collinearity of the $TREATMENT_j$ and the NGO_j binary variables—not included separately in Equation (20)—with the household-specific binary variables implies that γ_i also captures any differences that may exist across households in treated and untreated hamlets, and in NGO and non-NGO villages.

[21]Because we evaluate the impact of our ICS-promotion intervention using primary data collected during one baseline (pre-intervention) and two follow-up (post-intervention) survey rounds, we are unable to verify whether pre-intervention trends for our ICS-related adoption, use, and impact outcomes of interest are parallel across NGO and non-NGO villages. Instead, we test for differences in pre-trends across NGO and non-NGO villages for a host of community-level characteristics that are likely to be correlated with our outcomes of interest using the 2001 and 2011 rounds of the Indian Census. We find no evidence of differences in pre-trends across NGO and non-NGO villages over this period (Table B3).

Figure 2: Mean intervention ICS purchase rates in treated hamlets in NGO and non-NGO villages



Notes: this figure plots the share of households in NGO and non-NGO villages that purchased at least one intervention ICS in response to the ICS-promotion intervention as a percentage of all treated households in the respective stratum. Error bars represent 95 per cent confidence intervals for the means.
Source: authors' illustration based on household-level ICS purchase rates during intervention activities (see Figure 1).

5.1 Effectiveness of the intervention across NGO and non-NGO villages

Household-level purchase of intervention ICS technologies

Our primary outcome of interest is household-level purchase of ICS technologies promoted during the ICS-promotion intervention. Figure 2 highlights the mean purchase rate for households located in treated hamlets in NGO and non-NGO villages. On average, nearly 60 per cent of treated households in NGO villages purchased at least one of the two intervention ICS. In non-NGO villages, the corresponding figure is approximately 45 per cent. We next evaluate this difference more rigorously in a linear regression framework. Table 3 presents our results. We find that the promotion campaign is extremely effective at encouraging the uptake of the intervention stoves; as shown in column (1), over half of targeted households purchase at least one of the two promoted ICS technologies. However, when we disaggregate our results by NGO and non-NGO villages following Equation (19) in column (2), we find that the purchase rate is nearly 13 percentage points (28 per cent) higher in treated hamlets located in NGO villages—a statistically significant and positive 'NGO effect'.[22] These results are robust to the inclusion of household-level controls that were found to be unbalanced at baseline in Table 2, as shown in column (3).

[22] In Figure A2 we present results from the application of an approach inspired by randomization-based inferential procedures (Athey and Imbens 2017) to village-level NGO stratum allocation. Our approach relies on randomly assigning villages to placebo NGO and non-NGO strata, and re-estimating Equation (19); this process is repeated 1,000 times to obtain a distribution of placebo 'NGO effect' estimates. If the effect we observe was due to the chance selection of the villages in our NGO and non-NGO strata, we would expect to observe our actual estimate located near the middle of this distribution. Instead, we find that only 3 per cent of these placebo estimates are greater in magnitude than our actual estimates. In addition, in Appendix C, rather than characterizing NGO and non-NGO villages using a binary variable, we characterize the village-specific 'intensity' of NGO activity based on two different measures: (1) the number of projects/initiatives the NGO has implemented in a particular village; and (2) the number of years it has been active in a particular village. We use these two measures to separately re-estimate Equation (19) for heterogeneity in rates of ICS purchase and find that these analyses provide evidence that is consistent with our main results.

Table 3: Effect of promotion on intervention ICS purchase in matched NGO/non-NGO villages

	(1)	(2)	(3)
	$\mathbb{1}$ (Purchased intervention ICS)		
$TREATMENT_j$	0.52***	0.46***	0.45***
	(0.029)	(0.049)	(0.048)
NGO_j		0.00***	−0.017*
		(0.00)	(0.010)
$TREATMENT_j \times NGO_j$		0.13**	0.13**
		(0.058)	(0.060)
Constant	−0.00***	−0.00***	−0.079*
	(0.00)	(0.00)	(0.043)
Mean dep. (control)	0.00	0.00	0.00
Observations	943	943	943
Household-level controls	No	No	Yes
Adjusted R^2	0.23	0.24	0.24

Notes: the outcome variable is an indicator that equals 1 if household i in hamlet j purchased at least one of the two ICS promoted during the intervention. Column (1) presents aggregated results; results are disaggregated by NGO and non-NGO villages (as shown in Equation 19) in column (2). Baseline household-level controls for household size, number of children under five, awareness of existence of cleaner stoves and fuels, and total traditional-fuel collection time per day are included in column (3). 'Traditional fuel' includes crop residue, dung, fuelwood, leaves, and household waste (trash); missing observations for total time spent collecting traditional fuel for 32 households are replaced with the sample mean value. Standard errors (in parentheses) are clustered at the hamlet level. $^* p < 0.10$, $^{**} p < 0.05$, $^{***} p < 0.01$.

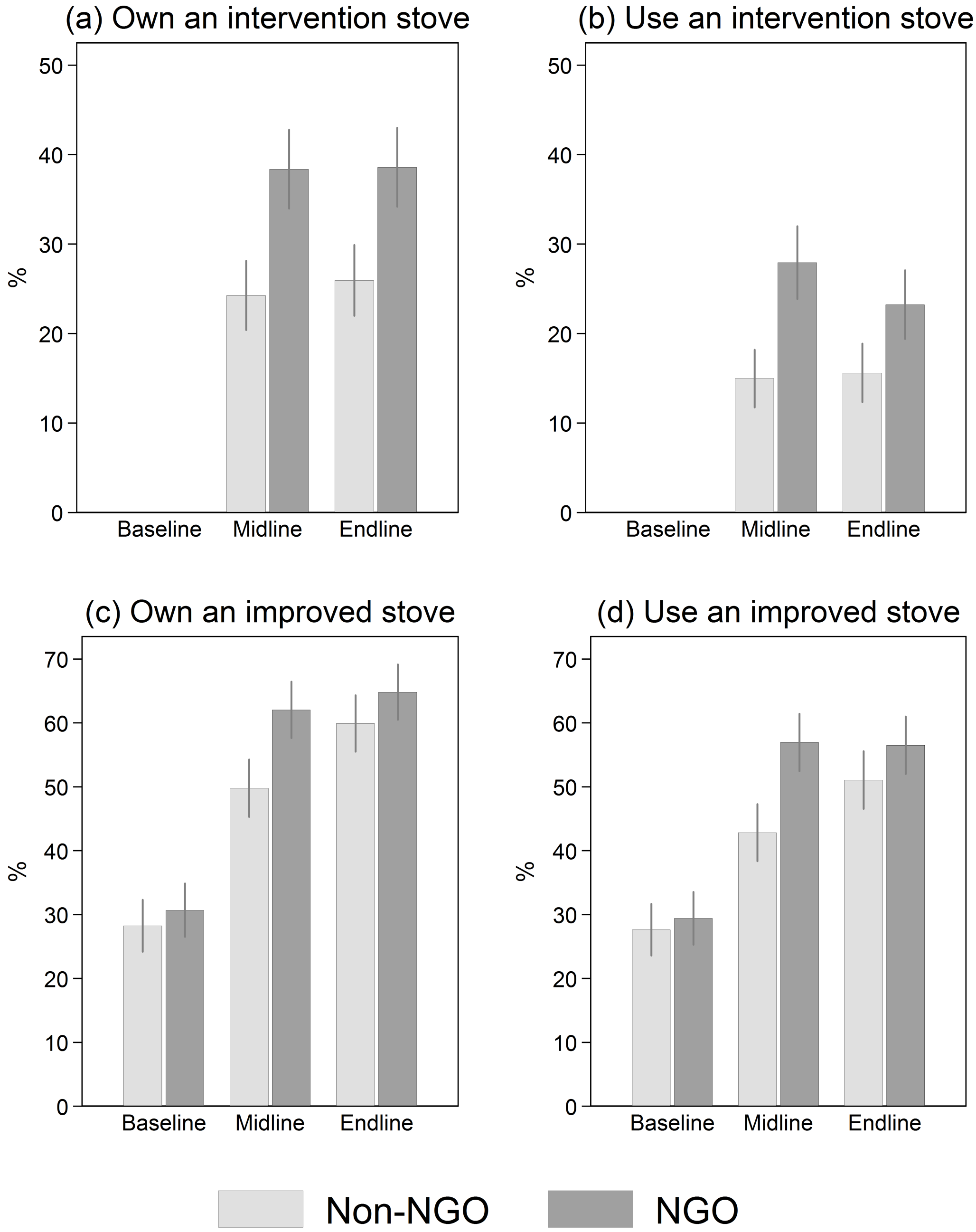
Source: authors' calculations based on household-level ICS purchase rates during intervention activities (see Figure 1).

Ownership and use of intervention ICS

This apparent 'NGO effect' is not limited to the initial purchase decision. Panels (a) and (b) of Figure 3 highlight that ownership and reported use of intervention ICS remains higher in NGO villages over multiple survey rounds. To rigorously evaluate differences in these trends, we separately estimate the triple-differences specification outlined in Equation (20) for ownership and use of intervention ICS during the midline (month 3) and endline (month 15) follow-up surveys. Our results are shown in Table 4, which also includes results from estimating a double-differences specification without NGO-specific interactions for comparison. While approximately 40 per cent of treated households report owning an intervention ICS and 30 per cent report having used it recently (columns 1 and 3, respectively), reported ownership and use by treated households in NGO villages during the first follow-up (in columns 2 and 4) are nearly 16 percentage points higher. This represents an increase in the size of the treatment effect of between 50 and 80 per cent relative to ownership and use by households in treated non-NGO hamlets. By the endline follow-up survey (conducted approximately 15 months after the start of the intervention), the difference in intervention ICS ownership rates across treated NGO and non-NGO hamlets remains positive, but is no longer statistically significant.[23] Similarly, although the difference in reported intervention ICS use rates between the two treated NGO and non-NGO groups remains positive by the endline survey, it is no longer statistically significant.

[23]That said, we are unable to reject that the difference between the two coefficients—for ownership of intervention ICS by households in treated NGO hamlets at midline and at endline—is statistically zero.

Figure 3: ICS ownership and use in NGO and non-NGO villages



Notes: panels (a) and (c) plot the share of households that own at least one of the two intervention ICS or an improved ICS, respectively, as a percentage of all (treated and untreated) sample households in the indicated stratum of villages. Panels (b) and (d) plot the subset of these households that report having used these devices in the week prior to the survey as a percentage of all (treated and untreated) sample households in the indicated stratum of villages. 'Improved stove' includes stoves fuelled by biogas, electricity, LPG, kerosene, and commercially available efficient biomass cookstoves; we also include the two ICS promoted as part of the promotion intervention in this definition. Error bars represent 95 per cent confidence intervals for the means. Baseline survey activities occurred approximately one year before the intervention; midline and endline surveys occurred approximately 3 and 15 months, respectively, after the intervention (see Figure 1).
Source: authors' illustration based on data collected during baseline, midline, and endline survey activities (see Figure 1).

Table 4: Effect of promotion on intervention ICS adoption in matched NGO/non-NGO villages

	(1)	(2)	(3)	(4)
	$\mathbb{1}$ (Owns intervention ICS)		$\mathbb{1}$ (Uses intervention ICS)	
$POST_1$	0.012*	0.014	0.0039	0.0068
	(0.0064)	(0.0091)	(0.0038)	(0.0066)
$POST_2$	0.039**	0.027*	0.027**	0.020*
	(0.016)	(0.016)	(0.011)	(0.011)
$TREATMENT_j \times POST_1$	0.41***	0.33***	0.29***	0.21***
	(0.030)	(0.047)	(0.026)	(0.034)
$TREATMENT_j \times POST_2$	0.39***	0.34***	0.23***	0.20***
	(0.034)	(0.053)	(0.027)	(0.042)
$NGO_j \times POST_1$		−0.0046		−0.0068
		(0.013)		(0.0066)
$NGO_j \times POST_2$		0.027		0.016
		(0.034)		(0.025)
$TREATMENT_j \times NGO_j \times POST_1$		0.16***		0.16***
		(0.058)		(0.048)
$TREATMENT_j \times NGO_j \times POST_2$		0.098		0.061
		(0.070)		(0.056)
Mean dep. (baseline non-NGO control)	0.00	0.00	0.00	0.00
Observations	2,829	2,829	2,829	2,829
Adjusted R^2	0.56	0.56	0.34	0.35
Household fixed-effects	Yes	Yes	Yes	Yes

Notes: the outcomes variables are ownership (columns 1–2) and use (columns 3–4) of at least one of the two intervention ICS as measured during baseline, and midline and endline follow-up surveys (conducted approximately 3 and 15 months after the start of the intervention, respectively). The outcome variable for ownership is an indicator that equals 1 if household i in hamlet j reports owning at least one of the two ICS promoted during the intervention in survey round t; for use, it is an indicator that equals 1 if the household—conditional on ownership—reports having used at least one of these two ICS in the past week. The even-numbered columns present the results of estimating the triple-differences specification outlined in Equation (20). The accompanying odd-numbered columns present results from estimating a double-differences specification without NGO-specific interactions for comparison. Standard errors (in parentheses) are clustered at the hamlet level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Source: authors' calculations based on data collected during baseline, midline, and endline survey activities (see Figure 1).

Ownership and use of improved stoves

We next separately estimate Equation (20) for reported ownership and use of all improved stoves—and not only the two ICS promoted during the intervention. Here, we aim to evaluate the effectiveness of our intervention more broadly.[24] Indeed, although we limited promotion activities to two specific ICS technologies, the primary goal of such an intervention—were it to be implemented as part of national or regional policies—would arguably be to increase reliance on cleaner, more efficient cooking technologies more generally. Understanding how institutional actors influence intervention effectiveness at this higher level is, thus, important.

As shown in Table 5, we find that the differential adoption and use patterns associated with uptake of the intervention ICS in treated NGO hamlets (highlighted in Tables 3 and 4) do not appear to be related to adoption and use of improved stoves more broadly. Specifically, while the intervention is effective at increasing ownership and use of improved stoves in treated hamlets (columns 1 and 3, respectively), we detect no differential effect of the intervention between treated hamlets in NGO and non-NGO villages (columns 2 and 4). In fact, the negative coefficient for both of our triple-differences estimates hints at an underlying 'technology substitution' effect—households that choose to purchase intervention ICS in treated NGO hamlets are those that might have adopted an improved stove in the absence of the intervention anyway.[25]

At the same time, our results also highlight other NGO-related dynamics. Recall that the first and second follow-up surveys occurred approximately 3 and 15 months after the start of ICS-promotion activities, respectively. Given the concerted efforts of the Indian government in recent years to enhance access to cleaner cooking fuels and technologies—and LPG, in particular (Barnwal 2017; Kumar et al. 2016)—we would expect ownership of improved stoves to naturally increase over this period. In Equation (20), we control for these time trends by including dummy variables for both follow-up survey rounds. As part of our triple-differences specification, we also interact these variables with the NGO-village dummy, which allows us to separately control for NGO-village-specific time trends. With these controls, we find that only three months after the intervention, ownership and use rates of improved stoves in control hamlets in NGO villages (represented by the coefficient for $NGO_j \times POST_1$) are nearly 15 percentage points higher.

In the Appendix, we attempt to shed light on these two phenomena by separately analysing trends in ownership of the three most important types of improved stoves in our setting (LPG, electric, and efficient biomass stoves) across treatment/control and NGO/non-NGO communities.[26] Figure A3 suggests that changes in ownership of improved stoves in both NGO and non-NGO control hamlets are almost entirely driven by changes in ownership of LPG stoves. This trend is confirmed when we estimate Equation (20) for only LPG stoves, indicated by the near equality in magnitude of the coefficients on the $NGO_j \times POST_1$ interaction term in columns (1) and (2) of Table B5. In addition, we find that our intervention indeed appears to have *lowered* ownership of LPG stoves in treated NGO hamlets initially relative to treated non-NGO hamlets—as indicated by the negative coefficient on the $TREATMENT_j \times NGO_j \times POST_1$ triple interaction term—in line with our 'technology substitution' hypothesis.

[24]Recall that in addition to the two intervention ICS, 'improved stoves' include kerosene burners, LPG and biogas stoves, as well as other electric or efficient biomass alternatives. In our study area, LPG stoves were the main improved alternative, owned by approximately one-third of households at baseline.

[25]Ownership and reported use of improved stoves at baseline (pre-intervention) are balanced across NGO and non-NGO hamlets (Figure 3, panels (c) and (d)), and baseline ownership of improved stoves does not predict subsequent purchase of an intervention ICS (Table B4), suggesting that intervention ICS are typically being purchased by households that did not already have an improved stove.

[26]Recall that one of our two intervention ICS is an electric stove while the other is an efficient biomass stove.

Table 5: Effect of promotion on improved-stove adoption in matched NGO/non-NGO villages

	(1)	(2)	(3)	(4)
	$\mathbb{1}$ (Owns improved stove)		$\mathbb{1}$ (Uses improved stove)	
$POST_1$	0.016	−0.048	0.0039	−0.054
	(0.036)	(0.042)	(0.034)	(0.041)
$POST_2$	0.14***	0.12***	0.12***	0.095*
	(0.039)	(0.044)	(0.040)	(0.048)
$TREATMENT_j \times POST_1$	0.34***	0.38***	0.29***	0.30***
	(0.047)	(0.066)	(0.042)	(0.056)
$TREATMENT_j \times POST_2$	0.26***	0.28***	0.18***	0.20***
	(0.048)	(0.065)	(0.047)	(0.061)
$NGO_j \times POST_1$		0.15**		0.14**
		(0.068)		(0.063)
$NGO_j \times POST_2$		0.040		0.067
		(0.083)		(0.085)
$TREATMENT_j \times NGO_j \times POST_1$		−0.10		−0.045
		(0.092)		(0.080)
$TREATMENT_j \times NGO_j \times POST_2$		−0.047		−0.059
		(0.10)		(0.098)
Mean dep. (baseline non-NGO control)	0.36	0.36	0.35	0.35
Observations	2,829	2,829	2,829	2,829
Adjusted R^2	0.56	0.57	0.52	0.52
Household fixed-effects	Yes	Yes	Yes	Yes

Notes: the outcome variables are ownership (columns 1–2) and use (columns 3–4) of all improved stoves as measured during baseline, and midline and endline follow-up surveys (conducted approximately 3 and 15 months after the start of the intervention, respectively). The outcome variable for ownership is an indicator that equals 1 if household i in hamlet j reports owning at least one improved stove in survey round t; for use, it is an indicator that equals 1 if the household—conditional on ownership—reports having used such a device in the past week. The even-numbered columns present the results of estimating the triple-differences specification outlined in Equation (20). The accompanying odd-numbered columns present results from estimating a double-differences specification without NGO-specific interactions for comparison. 'Improved stove' includes stoves fuelled by biogas, electricity, LPG, kerosene, and commercially available efficient biomass cookstoves; we also include the two ICS promoted as part of the promotion intervention in this definition. Standard errors (in parentheses) are clustered at the hamlet level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Source: authors' calculations based on data collected during baseline, midline, and endline survey activities (see Figure 1).

The mechanisms that lead to households in control NGO communities adopting LPG stoves at these elevated rates are unclear. It may be the case that opportunities for households living in different hamlets to interact are higher in NGO villages. This could lead to information spillovers across treated and control hamlets within NGO villages that induce households in the latter to adopt relatively readily available LPG stoves. It could also be the case that our NGO partner—having gained experience implementing improved-stove interventions—began to enable households in its control communities to avail of the Indian government's LPG-promotion schemes. Indeed, it is only one year later at the second follow-up that we see non-NGO control communities (represented by the identical coefficient on $POST_2$ in columns 1 and 2 of Table B5) beginning to catch up in terms of ownership of LPG stoves, consistent with the expected spread of improved LPG technologies over time. While we are unable to speak definitively about these underlying dynamics, our stratified study design does allow us to see how our intervention may have crowded out improved energy-technology acquisition in treated hamlets with prior interaction with the NGO. Because the socioeconomic and environmental benefits associated with an intervention like ours ultimately depend on community-level uptake of improved stoves broadly, this NGO-specific crowding out also has implications for intervention effectiveness.

Table 6: Impact of promotion on fuel collection and use in matched NGO/non-NGO villages

	(1)	(2)	(3)	(4)
	Fuelwood use (kilograms per day)		Fuel-collection time (minutes per day)	
$POST_1$	4.42***	3.38***	3.95	–19.0
	(0.68)	(0.57)	(18.4)	(18.5)
$POST_2$	1.50**	1.44	–9.83	–30.7**
	(0.75)	(1.18)	(15.5)	(12.5)
$TREATMENT_j \times POST_1$	–2.01**	–0.61	–26.0	19.2
	(0.93)	(0.91)	(21.9)	(25.4)
$TREATMENT_j \times POST_2$	0.65	0.47	–14.4	12.8
	(1.08)	(1.53)	(17.0)	(15.9)
$NGO_j \times POST_1$		2.24**		53.3
		(1.12)		(33.2)
$NGO_j \times POST_2$		0.12		48.5*
		(1.46)		(29.0)
$TREATMENT_j \times NGO_j \times POST_1$		–2.93*		–95.9**
		(1.67)		(40.4)
$TREATMENT_j \times NGO_j \times POST_2$		0.34		–60.8*
		(2.12)		(32.1)
Mean dep. (baseline non-NGO control)	8.90	8.90	113.6	113.6
Observations	1,143	1,143	2,829	2,829
Adjusted R^2	0.19	0.20	0.011	0.029
Household fixed-effects	Yes	Yes	Yes	Yes

Notes: the outcome variables are daily fuelwood use (columns 1–2) and total time spent collecting fuels (columns 3–4) as measured during baseline, and midline and endline follow-up surveys (conducted approximately 3 and 15 months after the start of the intervention, respectively). Daily fuelwood use is derived from a 24-hour before–after household-level fuel-weighing test; we restrict the analysis in columns (1) and (2) to the subsample of households that participated in such tests in at least two of our three survey rounds ($N = 388$). The outcome variable for fuel-collection time is derived from self-reported data on time spent (per day, week, or month) collecting fuelwood, crop residue, leaves, dung, biomass pellets, kerosene, LPG, biogas, and—if relevant—any other fuel used by the household; missing observations for time spent collecting fuel for up to six households in each survey round are replaced with the survey-round-specific sample mean value. The even-numbered columns present the results of estimating the triple-differences specification outlined in Equation (20). The accompanying odd-numbered columns present results from estimating a double-differences specification without NGO-specific interactions for comparison. Standard errors (in parentheses) are clustered at the hamlet level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Source: authors' calculations based on data collected during baseline, midline, and endline survey activities (see Figure 1).

5.2 Heterogeneity in impacts across NGO and non-NGO villages

Finally, we turn to an evaluation of the socioeconomic and environmental impacts of the ICS promotion with an eye to investigating heterogeneity in impacts across NGO and non-NGO villages.[27] Specifically,

[27]We note that our intervention was designed to evaluate the effectiveness of tools to promote the sale, adoption, and use of ICS, and was not intended to evaluate ICS impacts in real-world settings. That said, our stratified design does allow us to investigate impact heterogeneity. Bensch and Peters (2015), Bensch et al. (2015), Beyene et al. (2015), Brooks et al. (2016), Hanna et al. (2016), Lewis et al. (2016), Meeks et al. (2018), and Somanathan and Bluffstone (2015) are some recent examples of evaluations of the economic, environmental, and health impacts of various non-traditional cooking technologies.

we separately estimate Equation (20) for an objective measure of daily fuelwood use and for reported time spent collecting fuels for household use per day.

Columns (1) and (2) of Table 6 present our results for an objective measure of fuelwood use, obtained from a subsample of households that were randomly selected for weight-based measurements of their solid fuel use over a 24-hour period.[28] As shown in column (1), the intervention appears to reduce fuelwood use by approximately 2 kg per day in the relatively short run (as indicated by the negative estimated coefficients for $TREATMENT_j \times POST_1$). However, our triple-differences estimate for the midline in column (2) reveals that this effect is almost entirely driven by reductions in fuelwood use by households in treated NGO hamlets. These household appear to use nearly 3 kg less fuelwood per day—evidence of the 'NGO effect' that is consistent with reported use of purchased intervention stoves. By the time of the endline survey one year later (that is, the coefficient on $TREATMENT_j \times NGO_j \times POST_2$), this effect appears to attenuate somewhat, and we no longer detect a significant difference between fuelwood use in NGO and non-NGO communities. That said, it is worth noting that we are unable to reject that the difference between the triple-differences estimates for the midline and endline survey rounds is zero.

Consistent with these fuelwood use patterns, in column (3) we find that while there is no detectable effect of the intervention on reported daily fuel-collection times for treatment hamlets in general, households in treated hamlets in NGO villages reported significant reductions in time spent collecting fuels for household use at the time of the midline and endline follow-up surveys (column 4). This effect appears to be driven primarily by reductions in time spent collecting fuelwood and other traditional fuels (Table B6). Once again, the two triple-differences estimates are statistically indistinguishable from each other.

6 Bayesian synthesis of the evidence

How do our findings contribute to the limited evidence on the roles NGOs play in shaping outcomes of interventions? To answer this question, our final set of analyses combines the insights from the implementer identity literature—which finds that NGO-led interventions are often more effective than comparable efforts by other actors—with the results of our matched-experimental study design in a simple Bayesian regression framework.

Suppose β_3 is a random variable that represents the true effect of the NGO on the outcomes of the ICS promotion intervention in our setting—the same term we use to indicate the coefficient on the $TREATMENT_j \times NGO_j$ interaction term in Equation (19). Our first step is to suitably characterize the prior evidence base for this parameter. For this, we focus on two existing randomized evaluations. First, we turn to Bold et al. (2016), who evaluate a nationwide education reform in Kenya that expanded funding for hiring 'contract teachers'. Such teachers are hired directly by schools—typically at wages that are below those offered to tenured public school teachers—to address teacher shortages; they are also not accorded the same tenure protections available to their civil-service colleagues. There are a number of pathways through which contract teachers might improve educational outcomes.[29] Yet, as in many other cases, the evidence on the efficacy of such programmes is from relatively small-scale

[28]Households were instructed to set aside an amount of fuelwood they expected to use over the next 24-hour period. This amount was weighed by the field team, which returned the next day to reweigh the remaining amount. We underscore that the relatively involved nature of the fuel-weight test limited the available sample size and, consequently, our ability to detect meaningful impacts for this outcome variable.

[29]Specifically, Bold et al. (2016) highlight three pathways: (1) contract teachers are typically hired from a waiting list of candidates for civil-service teaching appointments, and are thus similarly skilled and less expensive to hire; (2) a 'selection effect', whereby only relatively good teachers are retained over time as poor-quality teachers do not have their contracts extended; and (3) an 'incentive effect', whereby the lack of permanent contracts introduces dynamic incentives to increase teaching effort.

interventions. As part of their large-scale evaluation, Bold et al. (2016) randomly assign a nationwide sample of public schools to a control group, and one of two treatment groups that differ only in the identity of the implementer—in one, the implementation is led by an NGO while in the other it is led by the government. They find that 'an additional contract teacher in a school where the program is managed by the NGO increased test scores by roughly 0.18 standard deviations'. In contrast, the treatment effect is both smaller and statistically insignificant in the government-implementation schools. Digging deeper into mechanisms, they also find that contract teachers were actually hired and in place for at least 12 per cent more months in NGO-implementation schools over the course of the intervention.

We turn next to Cameron and Shah (2017), who evaluate the scale-up of a community-led total sanitation (CLTS) intervention in the Indonesian province of East Java. Nearly one billion people engage in open defecation, partly due to a lack of access to improved sanitation facilities. CLTS relies primarily on social pressures—rather than subsidies or grants—to encourage adoption of latrines and induce sustained behaviour-change.[30] In addition to evaluating the impact of CLTS relative to a control group, Cameron and Shah (2017) investigate heterogeneity in impacts across treatment communities in which implementation was carried out by different actors. Specifically, treated communities were nearly evenly split; implementation was led by government staff in one half and by local non-governmental 'resource agencies' in the other.[31] They find that households are over 5 percentage points more likely to build a latrine in communities in which NGOs triggered the CLTS intervention—a 42 per cent increase in the rate of latrine ownership relative to the mean in the control communities. Once again, the treatment effect is both smaller and statistically insignificant in the government-implementation communities.

While not perfectly analogous to our setting, a priori, the results in these two studies offer insights about the possible range of β_3—the 'NGO effect'. A 42 per cent increase in the size of the treatment effect—as found by Cameron and Shah (2017)—would translate in our setting into an increase in the ICS purchase rate by households in treated NGO hamlets of nearly 20 percentage points relative to the purchase rate across treated non-NGO hamlets. In contrast, the more conservative 12 per cent estimate from Bold et al. (2016) would imply a difference in the purchase rate between treated NGO and non-NGO hamlets of closer to 5 percentage points. The midpoint of these two estimates is 12.5 percentage points. This prior information suggests that we use a prior distribution $f(\beta_3)$ that assigns most of its probability to the interval $(0.05, 0.20)$, and that the expected value of β_3 under $f(\beta_3)$ be close to 0.125. We, therefore, represent our prior information about β_3 as follows:

$$\beta_3 \sim \text{Beta}(2, 14). \tag{21}$$

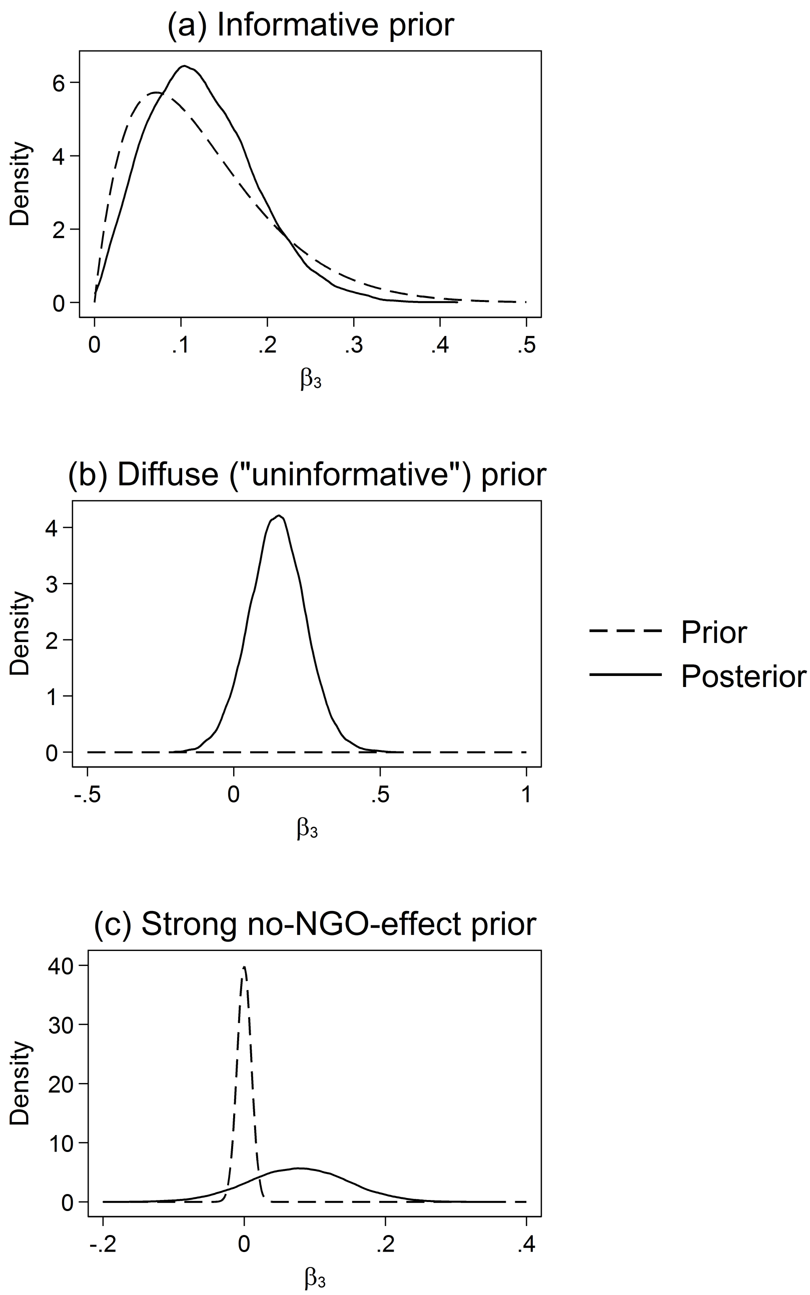
The density of this prior distribution—a beta distribution with shape parameters $\alpha = 2$ and $\beta = 14$—is represented by the dashed line in panel (a) of Figure 4. The expected value of β_3 under this prior is 0.125, and the most probable value is approximately 0.07, corresponding to the peak in the density function. Just under two-thirds of the area under the curve lies between 0.05 and 0.20. Importantly, this distribution only has positive support over the interval $[0, 1]$. Together, these characteristics implicitly capture the prior information that the 'NGO effect' is strictly positive but not excessively high.[32]

[30]Dickinson et al. (2015), for instance, describe a village-level CLTS intervention in India that entailed a 'walk of shame', defecation mapping, and faecal weighing—all designed to invoke an emotional response about the ubiquity of defecation sites in and around rural communities.

[31]While not technically randomized across implementers, the authors contend that no systematic process guided the selection of implementing teams. They also note that baseline village- and household-level characteristics are balanced across implementer arms.

[32]Admittedly, the choice of the specific beta prior outlined in Equation (21) is somewhat subjective as any number of alternative distributions would satisfy our outlined mean- and interval-related criteria. We note that the relative dearth of related evidence that we might draw upon to examine the distribution of estimates in more detail contributes to this subjectivity. Given this

Figure 4: Bayesian analysis of heterogeneity in ICS purchase across NGO and non-NGO hamlets



Notes: this figure plots prior and posterior distributions of the β_3 parameter, which represents the additional effect of the ICS promotion intervention on purchase rates in treated NGO hamlets relative to treated non-NGO hamlets. Posterior distributions are estimated via Markov Chain Monte Carlo (MCMC) simulation. Specifically, we ran 50,000 MCMC samples after a burn-in period of 10,000 iterations, with thinning every fifth iteration. In panel (a), $f(\beta_3) = \mathrm{Beta}(2,14)$; in panel (b), $f(\beta_3) = \mathcal{N}(0,10000)$; and in panel (c), $f(\beta_3) = \mathcal{N}(0,0.01)$. Diffuse priors are used for all other parameters in the underlying model for each panel (see Appendix D).

Source: authors' illustration based on procedure described in Appendix D and data on household-level ICS purchase rates collected during intervention activities (see Figure 1).

With this prior specified, we fit a Bayesian multilevel mixed-effects model broadly corresponding to the specification outlined in Equation (19) to investigate heterogeneity in purchase rates across treated NGO and non-NGO hamlets (described in detail in Appendix D).[33] Our results are shown in panel (a) of Figure 4, where the solid line approximates the posterior distribution we obtain for β_3. This distribution has less density in its tails and is more peaked, reflecting our updated beliefs given our specified prior and our data. The estimated posterior mean for β_3 is 0.124. The 90 per cent 'credible interval' (the range that has 90 per cent posterior probability to contain the true effect) is between 0.034 and 0.24. This represents a 7–50 per cent increase approximately in the size of the treatment effect, respectively.

To test the robustness of this result to prior specification, we repeat our analysis two additional times with alternatively specified prior distributions. First, we assume that:

$$\beta_3 \sim \mathcal{N}\left(\hat{\beta}_3^{\text{OLS}}, 10000\right). \tag{22}$$

This normal distribution (centred at our estimated coefficient for β_3 from Table 3) is highly diffuse ('uninformative') and has positive support over the real line; these characteristics reflect the a-priori beliefs of someone who only weakly suspects that the 'NGO effect' may be positive. Panel (b) of Figure 4 presents our results, in which the solid line represents the posterior distribution we obtain. The resulting posterior mean for β_3 is 0.15. The 90 per cent credible interval is $(-0.01, 0.31)$. In addition, the posterior probability that the 'NGO effect' is positive is 0.94.

Next, we assume that:

$$\beta_3 \sim \mathcal{N}(0, 0.01). \tag{23}$$

This distribution, in contrast, reflects relatively strong a-priori beliefs that NGOs do *not* influence the outcomes of the interventions they implement.[34] Panel (c) of Figure 4 presents our results. The resulting posterior mean for β_3 of 0.08 is expectedly diminished but remains positive. Indeed, the posterior probability that the 'NGO effect' is positive is 0.87—relatively unaffected by the considerably stronger prior centred on zero (compared to the highly diffuse prior in panel (b)). The 90 per cent credible interval is $(-0.04, 0.19)$.

Our analyses, thus, demonstrate that the evidence for a large and positive effect of the NGO on the effectiveness of the ICS-promotion intervention is relatively robust—even under strong prior distributional assumptions about the lack of such an effect. We also show how evidence from related research can be used to inform these distributional assumptions and guide causal inference.

7 Conclusion

Using data from an experimental intervention covering nearly 1,000 households across 97 geographically distinct hamlets in rural Uttarakhand, India, we highlight how NGOs influence the outcomes of applied interventions. We first develop a model of household decision-making grounded in transaction costs. We

constraint, we believe our specified prior distribution does a reasonable job of characterizing the existing evidence. In addition, we also test the robustness of our results to prior specification using both 'weak' and 'strong' prior distributional assumptions.

[33] Multilevel models account for hierarchical structures within the data. In our case, households are located within hamlets that are part of villages, which themselves are in distinct NGO and non-NGO strata. These level-specific effects also vary randomly across levels based on specified prior distributions. To facilitate comparability with our main analyses—where we cluster our estimated standard errors at the level of the hamlet—we restrict the nested structure to the level of the hamlet. In addition, we specify diffuse ('uninformative') priors for all random model parameters besides our parameter of interest.

[34] Arguably, this belief is implicit in the act of conducting applied research in partnership with NGOs without having in place a study design similar to ours, which explicitly attempts to identify NGO-related heterogeneity.

posit that NGOs lower transaction costs, and thus enhance the effectiveness of the interventions they implement. To empirically test our model's prediction, we use ex-ante propensity score matching to create a sample of observationally similar rural communities that are differentiated by prior exposure to a local development NGO. In partnership with this NGO, we then stratify an experimental intervention designed to promote ICS on this institutional variable to identify heterogeneity in adoption, use, and impacts.

We uncover a large, positive, and statistically significant 'NGO effect'—prior exposure to the NGO increases the effectiveness of the intervention by nearly 30 per cent. Specifically, in line with our model's predictions, ICS purchase rates for households in treated hamlets located in 'NGO villages' are 13 percentage points (28 per cent) higher than for households in treated hamlets located in matched 'non-NGO villages'. Using a difference-in-difference-in-differences ('triple-differences') specification, we find an even larger NGO effect on ICS use in these communities: households in such communities are up to 16 percentage points more likely to have used an ICS, representing a 50 per cent increase in the size of the treatment effect. Consistent with these patterns of ownership and use, households in villages with prior exposure to the NGO also exhibit reductions in daily fuelwood use and fuel-collection time; in contrast, we find no evidence of any impact of the intervention on energy-use patterns for treated households in non-NGO villages.

Although previous work has noted the presence of differential impacts across communities with and without NGO activity, to the best of our knowledge, we are the first to rigorously examine how the presence of an effective local NGO—and the specific institutional context that represents—directly influences household decision-making and ultimately determines the effectiveness of interventions. As such, our study begins to address a knowledge gap that has significant implications for the policy relevance of experimental research conducted in partnership with NGOs and other civil society organizations. Effective local organizations may be crucial for the implementation of environmental, health, and development interventions in remote, rural settings. Subsequent attempts to scale-up findings deemed effective in applied research conducted in partnership with such institutions into national or regional policies may prove much less successful than anticipated if their roles and contributions are insufficiently accounted for. Alternatively, promoters of scaled-up interventions could achieve greater success if they enlist the assistance of trusted local partners.

References

Aldashev, G. and C. Navarra. (2018) ‘Development NGOs: basic facts’. *Annals of Public and Cooperative Economics*, 89: 125–55. doi: 10.1111/apce.12188.

Aldashev, G., M. Limardi, and T. Verdier. (2015) ‘Watchdogs of the invisible hand: NGO monitoring and industry equilibrium’. *Journal of Development Economics*, 116: 28–42. doi: 10.1016/j.jdeveco.2015.03.006.

Allcott, H. (2015) ‘Site selection bias in program evaluation’. *Quarterly Journal of Economics*, 130(3): 1117–65. doi: 10.1093/qje/qjv015.

Anand, U. (2017). ‘India has 31 lakh NGOs, more than double the number of schools’. *The Indian Express*, August 2015. http://indianexpress.com/article/india/india-others/india-has-31-lakh-ngos-twice-the-number-of-schools-almost-twice-number-of-policemen (accessed on 11 July 2017).

Athey, S. and G. Imbens. (2017). ‘The econometrics of randomized experiments’. In *Handbook of Field Experiments*. New York: Elsevier,

Bailis, R., R. Drigo, A. Ghilardi, and O. Masera. (2015). ‘The carbon footprint of traditional woodfuels’. *Nature Climate Change*, 5(3): 266–72. doi: 10.1038/nclimate2491.

Banerjee, A., A.H. Amsden, R.H. Bates, J.N. Bhagwati, A. Deaton, and N. Stern. (2007). *Making Aid Work*. Cambridge, MA: MIT Press.

Banerjee, A., E. Duflo, N. Goldberg, D. Karlan, R. Osei, W. Pariente, J. Shapiro, B. Thuysbaert, and C. Udry. (2015). ‘A multifaceted program causes lasting progress for the very poor: evidence from six countries’. *Science*, 348(6236): 1260799. doi: 10.1126/science.1260799.

Banerjee, A., R. Banerji, J. Berry, E. Duflo, H. Kannan, S. Mukerji, M. Shotland, and M. Walton. (2017). ‘From proof of concept to scalable policies: challenges and solutions, with an application’. *Journal of Economic Perspectives*, 31(4): 73–102. doi: 10.1257/jep.31.4.73.

Barnwal, P. (2017). ‘Curbing leakage in public programs: evidence from India’s Direct Benefit Transfer policy’. https://bit.ly/2LxtcLK

Bengtsson, N. (2013). ‘Catholics versus Protestants: on the benefit incidence of faith-based foreign aid’. *Economic Development and Cultural Change*, 61 (3): 479–502. doi: 10.1086/669257.

Bensch, G. and J. Peters. (2015). ‘The intensive margin of technology adoption: experimental evidence on improved cooking stoves in rural Senegal’. *Journal of Health Economics*, 42: 44–63. doi: 10.1016/j.jhealeco.2015.03.006.

Bensch, G., M. Grimm, and J. Peters. (2015). ‘Why do households forego high returns from technology adoption? Evidence from improved cooking stoves in Burkina Faso’. *Journal of Economic Behavior & Organization*, 116: 187–205. doi: 10.1016/j.jebo.2015.04.023.

Berge, L.I.O., K. Bjorvatn, K.S. Juniwaty, and B. Tungodden. (2012). ‘Business training in Tanzania: from research-driven experiment to local implementation’. *Journal of African Economies*, 21(5): 808–27. doi: 10.1093/jae/ejs016.

Bernard, T., A. de Janvry, S. Mbaye, and E. Sadoulet. (2017). ‘Expected product market reforms and technology adoption by Senegalese onion producers’. *American Journal of Agricultural Economics*, 99(4): 1096–115. doi: 10.1093/ajae/aax033.

Beyene, A.D., R.A. Bluffstone, S. Dissanayake, Z. Gebreegziabher, P. Martinsson, A. Mekonnen, and M. Toman. (2015). Can improved biomass cookstoves contribute to REDD+ in low-income countries? Evidence from a controlled cooking test trial with randomized behavioral treatments. Policy Research Working Paper 7394. Washington, DC: World Bank.

Bold, T., M. Kimenyi, G. Mwabu, A. Ng'ang'a, and J. Sandefur. (2016). 'Experimental evidence on scaling up education reforms in Kenya'. https://www.tessabold.com/uploads/7/0/1/0/70101685/scaling_up_december_2016.pdf.

Brass, J.N. (2012). 'Why do NGOs go where they go? Evidence from Kenya'. *World Development*, 40(2): 387–401. doi: 10.1016/j.worlddev.2011.07.017.

Brooks, N., V. Bhojvaid, M. Jeuland, J.J. Lewis, O. Patange, and S.K. Pattanayak. (2016). 'How much do alternative cookstoves reduce biomass fuel use? Evidence from North India'. *Resource and Energy Economics*, 43: 153–71. doi: 10.1016/j.reseneeco.2015.12.001.

Cameron, L. and M. Shah. (2017). 'Scaling up sanitation: Evidence from an RCT in Indonesia'. Discussion Paper 10619. Bonn: IZA.

Crump, R.K., V.J. Hotz, G.W. Imbens, and O.A. Mitnik. (2009). 'Dealing with limited overlap in estimation of average treatment effects'. *Biometrika*, 96(1): 187–99. doi: 10.1093/biomet/asn055.

Deaton, A. (2010). 'Instruments, randomization, and learning about development'. *Journal of Economic Literature*, 48(2): 424–55. doi: 10.1257/jel.48.2.424.

Devarajan, S., S. Khemani, and M. Walton. (2013). 'Can civil society overcome government failure in Africa?' *The World Bank Research Observer*, 29(1): 20–47. doi: 10.1093/wbro/lkt008.

Dickinson, K.L., S.R. Patil, S.K. Pattanayak, C. Poulos, and J.-C. Yang. (2015). 'Nature's call: impacts of sanitation choices in Orissa, India'. *Economic Development and Cultural Change*, 64 (1): 1–29. doi: 10.1086/682958.

Edlin, A.S., and C. Shannon. 1998 'Strict monotonicity in comparative statics'. *Journal of Economic Theory*, 81(1): 201–19. doi: 10.1006/jeth.1998.2405.

Foster, A.D., and M.R. Rosenzweig. (2010). 'Microeconomics of technology adoption'. *Annual Review of Economics*, 2(1): 395–424. doi: 10.1146/annurev.economics.102308.124433.

Fruttero, A., and V. Gauri. (2005). 'The strategic choices of NGOs: location decisions in rural Bangladesh'. *Journal of Development Studies*, 41(5): 759–87. doi: 10.1080/00220380500145289.

Glennerster, R. (2015). 'A movement grows up: running randomized evaluations – a practical guide.' http://runningres.com/blog/2015/7/13/a-movement-grows-up (accessed 24 April 2016).

Grant, L.E., and K.K. Grooms. (2017). 'Do nonprofits encourage environmental compliance?' *Journal of the Association of Environmental and Resource Economists*, 4(S1): S261–S288. doi: 10.1086/692508.

Grossman, G., M. Humphreys, and G. Sacramone-Lutz. (2016). 'Information technology and political engagement: Mixed evidence from Uganda'. www.researchgate.net/publication/303818782_Information_Technology_and_Political_Engagement_Mixed_Evidence_from_Uganda

Grossman, M. (1972). 'On the concept of health capital and the demand for health'. *Journal of Political Economy*, 80(2): 223–55. doi: 10.1086/259880.

Hanna, R., E. Duflo, and M. Greenstone. (2016). 'Up in smoke: the influence of household behavior on the long-run impact of improved cooking stoves'. *American Economic Journal: Economic Policy*, 8(1): 80–114. doi: 10.1257/pol.20140008.

Henderson, J.V., and Y.S. Lee. (2015). 'Organization of disaster aid delivery: spending your donations'. *Economic Development and Cultural Change*, 63(4): 617–64. doi: 10.1086/681277.

Herman, M.L., G.L. Head, T.E. Fogarty, and P.M. Jackson. (2003). *Managing Risk in Nonprofit Organizations: A Comprehensive Guide*. New York: Wiley.

Holloway, G., C. Nicholson, C. Delgado, S. Staal, and S. Ehui. (2000). 'Agroindustrialization through institutional innovation: transaction costs, cooperatives and milk-market development in the East-African highlands'. *Agricultural Economics*, 23(3): 279–88. doi: 10.1111/j.1574-0862.2000.tb00279.x.

Ito, T. (2009). 'Caste discrimination and transaction costs in the labor market: evidence from rural North India'. *Journal of Development Economics*, 88(2): 292–300. doi: 10.1016/j.jdeveco.2008.06.002.

Jack, W. and T. Suri. (2014). 'Risk sharing and transactions costs: evidence from Kenya's mobile money revolution'. *American Economic Review*, 104(1): 183–223. doi: 10.1257/aer.104.1.183.

Jayachandran, S., J. de Laat, E.F. Lambin, C.Y. Stanton, R. Audy, and N.E. Thomas. (2017). 'Cash for carbon: a randomized trial of payments for ecosystem services to reduce deforestation'. *Science*, 357(6348): 267–73. doi: 10.1126/science.aan0568.

Jeuland, M. and S.K. Pattanayak. (2012). Benefits and costs of improved cookstoves: assessing the implications of variability in health, forest and climate impacts. *PLoS One*, 7(2): e30338. doi: 10.1371/journal.pone.0030338.

Jeuland, M., S.K. Pattanayak, and R. Bluffstone. (2015). 'The economics of household air pollution'. *Annual Review of Resource Economics*, 7(1): 81–108. doi: 10.1146/annurev-resource-100814-125048.

Jones, D., D. Molitor, and J. Reif. (2018). 'What do workplace wellness programs do? Evidence from the Illinois workplace wellness study'. Working Paper 24229. Cambridge, MA: NBER.

King, G., E. Gakidou, N. Ravishankar, R.T. Moore, J. Lakin, M. Vargas, M.M. Téllez-Rojo, J.E.H. Ávila, M.H. Ávila, and H.H. Llamas. (2007). 'A "politically robust" experimental design for public policy evaluation, with application to the Mexican Universal Health Insurance program'. *Journal of Policy Analysis and Management*, 26(3): 479–506. doi: 10.1002/pam.20279.

Kranton, R.E. (1996). 'Reciprocal exchange: A self-sustaining system'. *American Economic Review*, 86(4): 830–51.

Kumar, P., R.K. Rao, and N.H. Reddy. (2016). 'Sustained uptake of LPG as cleaner cooking fuel in rural India: role of affordability, accessibility, and awareness'. *World Development Perspectives*, 4: 33–37. doi: 10.1016/j.wdp.2016.12.001.

Lewis , J.J. and S.K. Pattanayak. (2012). 'Who adopts improved fuels and cookstoves? A systematic review'. *Environmental Health Perspectives*, 120(5): 637–45. doi: 10.1289/ehp.1104194.

Lewis, J.J., J.W. Hollingsworth, R.T. Chartier, E.M. Cooper, W.M. Foster, G.L. Gomes, P.S. Kussin, J.J. MacInnis, B.K. Padhi, P. Panigrahi, C.E. Rodes, I.T. Ryde, A.K. Singha, H.M. Stapleton, J. Thornburg, C.J. Young, J.N. Meyer, and S.K. Pattanayak. (2016). 'Biogas stoves reduce firewood use, household air pollution, and hospital visits in Odisha, India'. *Environmental Science & Technology*, 51(1): 560–69. doi: 10.1021/acs.est.6b02466.

Lin, L., S.K. Pattanayak, E.O. Sills, and W.D. Sunderlin. (2012). 'Site selection for forest carbon projects'. In A. Angelsen, M. Brockhaus, W.D. Sunderlin, and L. Verchot (eds), *Analysing REDD+: Challenges and Choices*. Bogor: Center for International Forestry Research (CIFOR).

MacLean, L.M., J.N. Brass, S. Carley, A. El-Arini, and S. Breen. (2015). 'Democracy and the distribution of NGOs promoting renewable energy in Africa'. *Journal of Development Studies*, 51(6): 725–42. doi: 10.1080/00220388.2014.989994.

Maurya, N.K. (2014). 'A critical evaluation of the state finances of the Uttarakhand governments: 2002–03 to 2011–12'. Technical Report. Uttar Pradesh: Fourteenth Finance Commission, Government of India.

Meeks, R., K.R.E. Sims, and H. Thompson. (2018). 'Waste not: can household biogas deliver sustainable development?'. *Environmental and Resource Economics*. doi: 10.1007/s10640-018-0224-1.

Micklewright , J., and A. Wright. (2003). 'Private donations for international development'. Discussion Paper 2003/82. Helsinki: UNU-WIDER.

Miguel, E., and M. Kremer. (2004). 'Worms: identifying impacts on education and health in the presence of treatment externalities'. *Econometrica*, 72(1): 159–217. doi: 10.1111/j.1468-0262.2004.00481.x.

Mobarak, A.M., P. Dwivedi, R. Bailis, L. Hildemann, and G. Miller. (2012). 'Low demand for nontraditional cookstove technologies'. *Proceedings of the National Academy of Sciences*, 109(27): 10815–20. doi: 10.1073/pnas.1115571109.

Niehaus, P. and S. Sukhtankar. (2013). 'The marginal rate of corruption in public programs: evidence from India'. *Journal of Public Economics*, 104: 52–64. doi: 10.1016/j.jpubeco.2013.05.001.

Pattanayak, S.K. and A. Pfaff. (2009). 'Behavior, environment, and health in developing countries: evaluation and valuation'. *Annual Review of Resource Economics*, 1(1): 183–217. doi: 10.1146/annurev.resource.050708.144053.

Pattanayak, S.K., J.-C. Yang, K.L. Dickinson, C. Poulos, S.R. Patil, R.K. Mallick, J.L. Blitstein, and P. Praharaj. (2009). 'Shame or subsidy revisited: social mobilization for sanitation in Orissa, India'. *Bulletin of the World Health Organization*, 87(8): 580–87. doi: 10.2471/blt.08.057422.

Pattanayak, S.K., R.A. Kramer, and J.R. Vincent. (2017). 'Ecosystem change and human health: Implementation economics and policy'. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 372(1722): 20160130. doi: 10.1098/rstb.2016.0130.

Peters, J., J. Langbein, and G. Roberts. (2018). 'Generalization in the tropics: development policy, randomized controlled trials, and external validity'. *The World Bank Research Observer*, 33(1): 34–64. doi: 10.1093/wbro/lkx005.

Ravallion, M. (2009). 'Evaluation in the practice of development'. *The World Bank Research Observer*, 24(1): 29–53. doi: 10.1093/wbro/lkp002.

Rosenbaum, P.R. and D.B. Rubin. (1984). 'Reducing bias in observational studies using subclassification on the propensity score'. *Journal of the American Statistical Association*, 79(387): 516–24. doi: 10.1080/01621459.1984.10478078.

Rosenthal, J., K. Balakrishnan, N. Bruce, D. Chambers, J. Graham, D. Jack, L. Kline, O. Masera, S. Mehta, I.R. Mercado, G. Neta, S. Pattanayak, E. Puzzolo, H. Petach, A. Punturieri, A. Rubinstein, M. Sage, R. Sturke, A. Shankar, K. Sherr, K. Smith, and G. Yadama. (2017). 'Implementation science to accelerate clean cooking for public health'. *Environmental Health Perspectives*, 125(1): A3–A7. doi: 10.1289/ehp1018.

Schaner, S. (2016). 'The cost of convenience?' *Journal of Human Resources*, 52(4): 919–45. doi: 10.3368/jhr.52.4.0815-7350r1.

Sharma, B.P., S.K. Pattanayak, M. Nepal, P. Shyamsundar, and B.S. Karky. (2015). 'REDD+ impacts: Evidence from Nepal'. Working Paper 95-15. Kathmandu: South Asian Network for Development and Environmental Economics.

Somanathan, E. and R. Bluffstone. (2015). 'Biogas: clean energy access with low-cost mitigation of climate change'. *Environmental and Resource Economics*, 62(2): 265–77. doi: 10.1007/s10640-015-9961-6.

T. Suri. (2011). 'Selection and comparative advantage in technology adoption'. *Econometrica*, 79(1): 159–209. doi: 10.3982/ecta7749.

Venkataraman, C. (2005). 'Residential biofuels in South Asia: Carbonaceous aerosol emissions and climate impacts'. *Science*, 307(5714): 1454–56. doi: 10.1126/science.1104359.

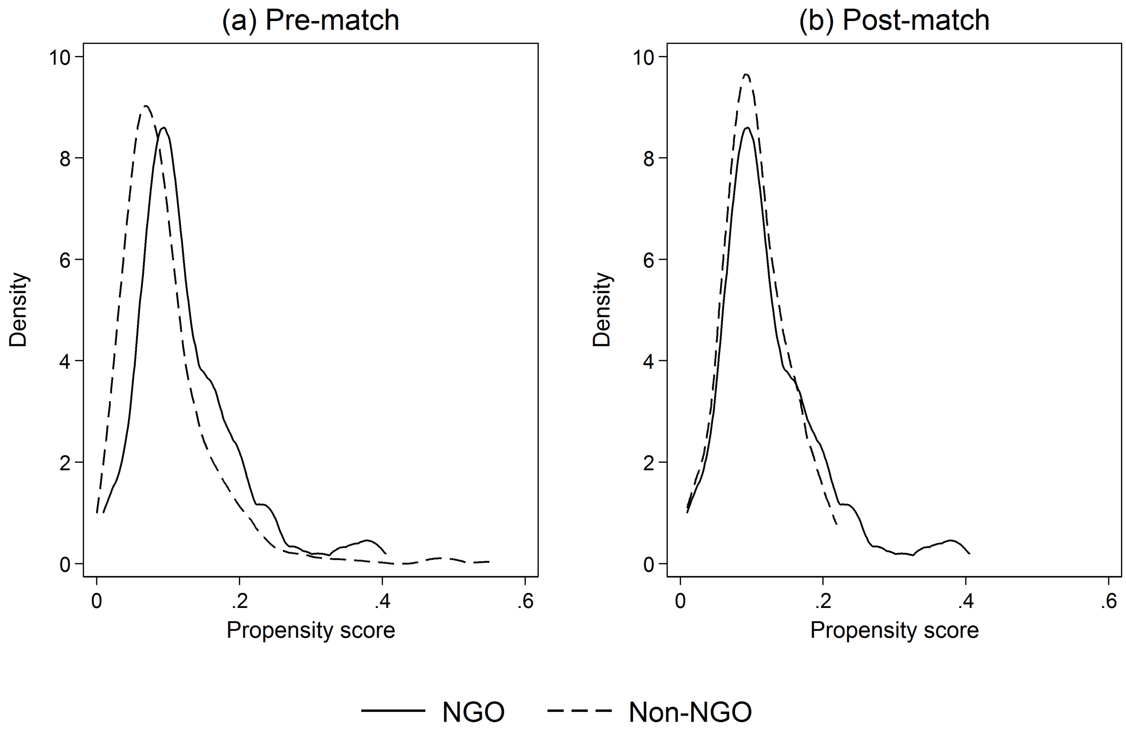
Vivalt, E. (2017). 'How much can we generalize from impact evaluations?' http://evavivalt.com/wp-content/uploads/How-Much-Can-We-Generalize.pdf.

Werker, E., and F.Z. Ahmed. (2008). 'What do nongovernmental organizations do?'. *Journal of Economic Perspectives*, 22(2): 73–92. doi: 10.1257/jep.22.2.73.

Westfall, P.H. and S.S. Young. 1993 *Resampling-Based Multiple Testing: Examples and Methods for p-Value Adjustment*. New York: Wiley-Interscience.

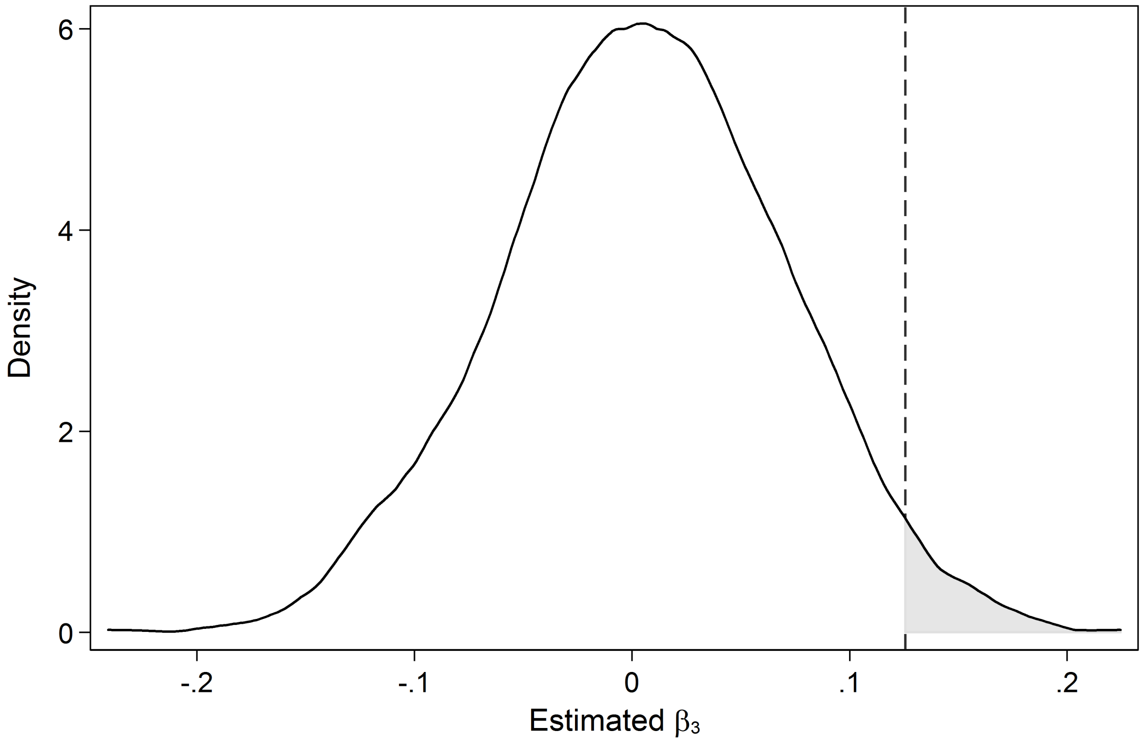
A Additional figures

Figure A1: Distribution of predicted propensity scores of NGO and non-NGO villages



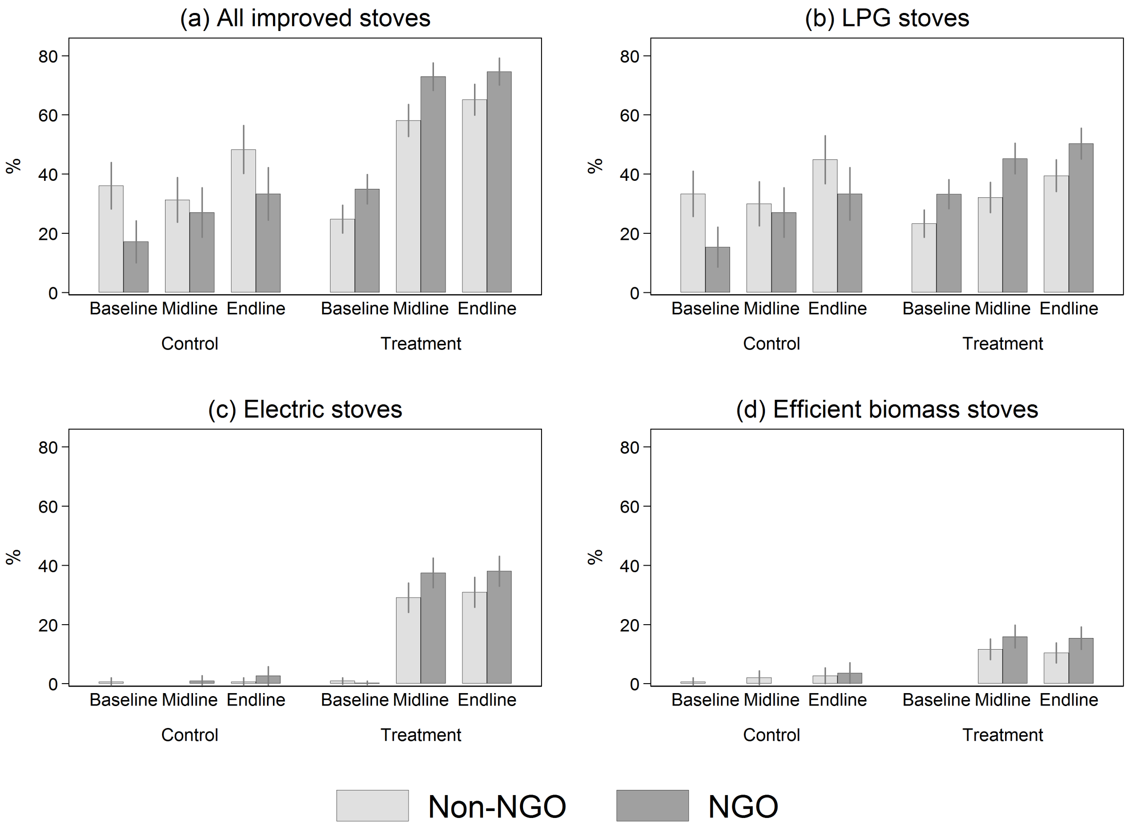
Notes: this figure presents the distribution of predicted propensity scores using the model outlined in column (3) of Table B1 before (panel (a)) and after (panel (b)) the propensity-score matching exercise. Prior to matching, we restrict our sample to villages in nine sub-districts of Bageshwar and Nainital districts of the state of Uttarakhand for implementation-related logistical reasons; the distribution of propensity scores for all villages in these sub-districts ($N_{\text{NGO}} = 97$ and $N^{\text{Unmatched}}_{\text{Non-NGO}} = 536$) is shown in panel (a). In panel (b), the distribution of propensity scores for only those non-NGO villages that are matched to at least one NGO village ($N^{\text{Matched}}_{\text{Non-NGO}} = 74$) is shown.
Source: authors' illustration based on calculations described above and in Section 3.1, and data from the 2001 round of the Census of India.

Figure A2: Randomization-based inferential procedure applied to village-level stratum allocation



Notes: this figure plots the distribution of 1,000 estimated β_3 coefficients from a randomization inference procedure (Athey and Imbens 2017) applied to village-level NGO stratum allocation to estimate Equation (19). We randomly assign each village in the sample to placebo NGO and non-NGO strata, and estimate the specification presented in column (3) of Table 3 to obtain a placebo 'NGO effect' estimate for heterogeneity in purchase of intervention ICS. This procedure is repeated 1,000 times to obtain a distribution of placebo effects. The vertical line indicates the magnitude of our actual estimated 'NGO effect'. Approximately 3 per cent of placebo estimates are larger than the actual estimated effect (shaded area).
Source: authors' illustration based on calculations described above and in footnote 22, and data on household-level ICS purchase rates collected during intervention activities (see Figure 1).

Figure A3: Trends in ownership of selected improved stoves



Notes: this figure presents mean ownership rates of all improved stoves as well as LPG, electric and improved biomass variants separately for treatment/control and NGO/non-NGO communities during each survey round. 'Improved stove' in panel (a) includes stoves fuelled by biogas, electricity, LPG, kerosene, and commercially available efficient biomass cookstoves; we also include the two ICS promoted as part of the promotion intervention in this definition. Electric stoves (panel (c)) and efficient biomass stoves (panel d) include the respective ICS promoted as part of the promotion intervention. Error bars represent 95 per cent confidence intervals for the means. Baseline survey activities occurred approximately one year before the intervention; midline and endline surveys occurred approximately 3 and 15 months, respectively, after the intervention (see Figure 1).
Source: authors' illustration based on data collected during baseline, midline, and endline survey activities (see Figure 1).

B Additional tables

Table B1: Propensity-score estimation using logistic regression

Village-level characteristic	(1) $\mathbb{1}$ (NGO village)	(2) $\mathbb{1}$ (NGO village)	(3) $\mathbb{1}$ (NGO village)
Area (km^2)	0.00055	0.000096	0.000096
	(0.00058)	(0.00011)	(0.00011)
Area^2	−0.000000025		
	(0.00000010)		
Total population	0.00078	−0.00058*	−0.00060**
	(0.00047)	(0.00030)	(0.00030)
Scheduled Caste population (proportion)	0.26	1.06***	1.06***
	(0.37)	(0.28)	(0.28)
Scheduled Tribe population (proportion)	7.02	−4.41	−4.59
	(4.84)	(5.07)	(5.14)
Population density	−0.057*		
	(0.030)		
Number of primary schools	0.41**	0.59***	0.58***
	(0.18)	(0.12)	(0.12)
Number of middle schools	−0.16	0.080	0.11
	(0.29)	(0.22)	(0.22)
Number of secondary schools	−0.058		
	(0.60)		
$\mathbb{1}$ (Medical facilities)	−0.66**		
	(0.33)		
Number of health centres	0.63	0.60	0.57
	(0.87)	(0.62)	(0.63)
Number of primary health centres	−0.44	−0.23	−0.20
	(0.96)	(0.73)	(0.73)
Number of telephone connections	−0.099**	−0.093*	−0.094*
	(0.046)	(0.055)	(0.056)
$\mathbb{1}$ (Bus services)	−0.50	0.65***	0.62***
	(0.31)	(0.24)	(0.24)
$\mathbb{1}$ (Credit societies)	0.31	0.72**	0.71**
	(0.47)	(0.34)	(0.34)
$\mathbb{1}$ (Approach to village: paved road)	−0.29	−0.43*	−0.41*
	(0.30)	(0.23)	(0.23)
Distance from nearest town (km)	−0.017**	−0.016***	−0.015***
	(0.0081)	(0.0045)	(0.0045)
Forest area (hectares)	0.00019	0.00026	0.00036
	(0.00035)	(0.00033)	(0.00034)
$\mathbb{1}$ (Tap water)			−0.0054
			(0.31)
$\mathbb{1}$ (Electricity for all purposes)			−0.0058
			(0.34)
Constant	−19.2	−2.63***	−2.62***
	(594.4)	(0.20)	(0.33)
Observations	1,960	1,965	1,903
Pseudo R^2	0.51	0.079	0.077
Sub-district fixed effects	Yes	No	No

Notes: this table presents results from logistic regressions of an indicator for whether our partner NGO had operated in village i in the past—represented by $\mathbb{1}$ (NGO village)—on a set of village-level characteristics from the 2001 round of the Indian Census. Standard errors in parentheses. Our final model is shown in column (3). For this model, we initially restrict our sample to all Census-designated villages in the Bageshwar and Nainital districts of the state of Uttarakhand with non-zero or non-missing values for total population. We then exclude six villages where pretesting activities occurred with an alternative NGO partner (details available upon request). The final estimation sample for the model presented in column (3), thus, consists of all remaining villages with non-missing values for the village-level characteristics used for estimation. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Source: authors' calculations based on NGO programme data and the 2001 round of the Census of India.

Table B2: Comparison of NGO and non-NGO villages using selected 2011 Census variables

	(1) $\mathbb{1}$ (NGO village)	(2) p value	(3) Adjusted p value	(4) R^2	(5) N
Number of dwelling rooms (%)					
No exclusive room	1.22	0.259	0.933	0.035	38
One	−8.70	0.054*	0.658	0.099	38
Two	−1.80	0.715	0.999	0.0037	38
Three	4.63	0.178	0.891	0.050	38
Four	−1.03	0.782	0.999	0.0022	38
Five	1.56	0.408	0.979	0.019	38
Six or more	4.13	0.306	0.941	0.029	38
Household size					
One	−2.67	0.010**	0.320	0.17	38
Two	−2.57	0.025**	0.487	0.13	38
Three	−2.36	0.129	0.857	0.063	38
Four	−2.69	0.146	0.873	0.058	38
Five	0.89	0.648	0.999	0.0059	38
Six to eight	7.87	0.011**	0.334	0.16	38
Nine or greater	1.52	0.174	0.891	0.051	38
Tap water from treated source (%)	−1.99	0.879	0.999	0.00065	38
Main source of lighting (%)					
Electricity	12.1	0.004***	0.197	0.21	38
Kerosene	−11.1	0.004***	0.201	0.21	38
Type of fuel used for cooking (%)					
Fuelwood	−0.33	0.958	0.999	0.000079	38
LPG	0.58	0.924	0.999	0.00025	38
Electricity	−0.058	0.324	0.946	0.027	38
Number of households availing of banking services	2.04	0.711	0.999	0.0039	38
Asset ownership (%)					
Radio	−1.37	0.803	0.999	0.0017	38
Television	9.72	0.145	0.873	0.058	38

Notes: column (1) presents the estimated β_1 coefficients for the specified Census outcome variable from a regression model of the form: $Y_i = \beta_0 + \beta_1 \cdot \mathbb{1}(\text{NGO village}) + \nu_i$, where Y_i represents a village-level characteristic for village i in the 2011 Census round, $\mathbb{1}(\text{NGO village})$ represents an indicator for whether our partner NGO had operated in village i in the past, and ν_i represents a normally distributed error component. Column (2) shows the corresponding p value—derived from heteroscedasticity robust standard errors—associated with each estimated coefficient. Column (3) shows p values obtained using the free step-down resampling methodology of Westfall and Young (1993), as operationalized by Jones et al. (2018). The unit of analysis is the Census-designated village $(N = 38)$. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Source: authors' calculations based on NGO programme data and the 2011 round of the Census of India.

Table B3: Pre-trends for selected village-level variables from the 2001 and 2011 Census

Village-level characteristic	(1)	(2)	(3)	(4)
	$\mathbb{1}$ (Census 2011) $\times$ $\mathbb{1}$ (NGO village)		R^2	N
Number of households	5.05	(13.0)	0.037	76
Total population	34.3	(66.4)	0.032	76
Total population (females)	13.8	(32.2)	0.014	76
Total population (males)	20.5	(35.6)	0.051	76
Total population (Scheduled Caste/Scheduled Tribe)	11.8	(58.6)	0.013	76
Number of primary schools	−0.00	(0.21)	0.017	76
Number of other educational facilities	0.16	(0.44)	0.026	76
Number of primary health centres	0.053	(0.053)	0.040	76
Number of community health workers	−0.00	(0.074)	0.027	76
$\mathbb{1}$ (Tap water)	0.11	(0.072)	0.081	76
$\mathbb{1}$ (Tubewell)	−0.053	(0.053)	0.040	76
$\mathbb{1}$ (Bus services)	0.11	(0.17)	0.097	76
$\mathbb{1}$ (Electricity for agricultural use)	−0.26	(0.17)	0.092	76
$\mathbb{1}$ (Electricity for domestic use)	−0.053	(0.089)	0.050	76
$\mathbb{1}$ (Approach to village: paved road)	0.11	(0.21)	0.16	76
$\mathbb{1}$ (Post office)	0.11	(0.16)	0.14	76
Total irrigated land area (hectares)	2.78	(5.52)	0.043	76
Total unirrigated land area (hectares)	1.44	(12.7)	0.011	76

Notes: column (1) presents the estimated β_3 coefficients for the specified Census outcome variable from a regression model of the form: $Y_{it} = \beta_0 + \beta_1 \cdot \mathbb{1}\,(\text{Census 2011}) + \beta_2 \cdot \mathbb{1}\,(\text{NGO village}) + \beta_3\,[\mathbb{1}\,(\text{Census 2011}) \times \mathbb{1}\,(\text{NGO village})] + \nu_{it}$, where Y_{it} represents a village-level characteristic for village i in Census round t, $\mathbb{1}$ (Census 2011) represents an indicator for the 2011 Census round (the 2001 Census round is the omitted category), $\mathbb{1}$ (NGO village) represents an indicator for whether our partner NGO had operated in village i in the past, and ν_{it} represents a normally distributed error term. Heteroscedasticity robust standard errors in parentheses shown in column (2). * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Source: authors' calculations based on NGO programme data, and the 2001 and 2011 rounds of the Census of India.

Table B4: Baseline improved-stove ownership does not predict intervention-stove purchase

	(1)	(2)
	$\mathbb{1}$ (Purchased intervention ICS)	
$TREATMENT_j$	0.51***	0.50***
	(0.036)	(0.037)
$\mathbb{1}$ (Owns an improved stove at baseline)	0.00	0.0015
	(–)	(0.0073)
$TREATMENT_j \times \mathbb{1}$ (Owns an improved stove at baseline)	0.051	0.042
	(0.057)	(0.057)
Constant	−0.00***	−0.11**
	(0.00)	(0.045)
Mean dep. (control)	0.00	0.00
Observations	943	943
Adjusted R^2	0.23	0.24
Household-level controls	No	Yes

Notes: the outcome variable is an indicator that equals 1 if household i in hamlet j purchased at least one of the two ICS promoted during the intervention. Baseline household-level controls for household size, number of children under five, awareness of existence of cleaner stoves and fuels, and total traditional-fuel collection time per day are included in column (3). 'Traditional fuel' includes crop residue, dung, fuelwood, leaves, and household waste (trash); missing observations for total time spent collecting traditional fuel for 32 households are replaced with the sample mean value. Standard errors (in parentheses) are clustered at the hamlet level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Source: authors' calculations based on data collected during baseline survey activities and household-level ICS purchase rates during intervention activities (see Figure 1).

Table B5: Comparing impacts on ownership of improved stoves and LPG stoves

	(1) $\mathbb{1}$ (Owns improved stove)	(2) $\mathbb{1}$ (Owns LPG stove)
$POST_1$	−0.048	−0.034
	(0.042)	(0.037)
$POST_2$	0.12***	0.12***
	(0.044)	(0.042)
$TREATMENT_j \times POST_1$	0.38***	0.12***
	(0.066)	(0.045)
$TREATMENT_j \times POST_2$	0.28***	0.046
	(0.065)	(0.050)
$NGO_j \times POST_1$	0.15**	0.15**
	(0.068)	(0.060)
$NGO_j \times POST_2$	0.040	0.065
	(0.083)	(0.082)
$TREATMENT_j \times NGO_j \times POST_1$	−0.10	−0.12*
	(0.092)	(0.070)
$TREATMENT_j \times NGO_j \times POST_2$	−0.047	−0.056
	(0.10)	(0.090)
Mean dep. (baseline non-NGO control)	0.36	0.33
Observations	2,829	2,829
Adjusted R^2	0.57	0.65
Household fixed-effects	Yes	Yes

Notes: the outcome variable in column (1) is an indicator that equals 1 if household i in hamlet j reports owning at least one improved stove in survey round t; the results reported in column (1) are identical to those reported in column (2) of Table 5. Similarly, as in Table 5, 'improved stove' includes stoves fuelled by biogas, electricity, LPG, kerosene, and commercially available efficient biomass cookstoves; we also include the two ICS promoted as part of the promotion intervention in this definition. In column (2), the outcome variable is an indicator that equals 1 if household i in hamlet j reports owning an LPG stove. Standard errors (in parentheses) are clustered at the hamlet level. $^{*}\ p < 0.10$, $^{**}\ p < 0.05$, $^{***}\ p < 0.01$.

Source: authors' calculations based on data collected during baseline, midline, and endline survey activities (see Figure 1).

Table B6: Comparing impacts on fuel-collection time for all fuels and traditional fuels

	(1)	(2)
	Fuel-collection time (minutes per day)	
	All fuels	Traditional fuels
$POST_1$	-19.0	-11.8
	(18.5)	(19.0)
$POST_2$	-30.7^{**}	-42.0^{***}
	(12.5)	(11.3)
$TREATMENT_j \times POST_1$	19.2	15.9
	(25.4)	(24.8)
$TREATMENT_j \times POST_2$	12.8	14.0
	(15.9)	(14.7)
$NGO_j \times POST_1$	53.3	50.1
	(33.2)	(31.7)
$NGO_j \times POST_2$	48.5^{*}	53.3^{**}
	(29.0)	(24.9)
$TREATMENT_j \times NGO_j \times POST_1$	-95.9^{**}	-86.5^{**}
	(40.4)	(38.8)
$TREATMENT_j \times NGO_j \times POST_2$	-60.8^{*}	-54.5^{*}
	(32.1)	(28.2)
Mean dep. (baseline non-NGO control)	113.6	104.2
Observations	2,829	2,829
Adjusted R^2	0.029	0.038
Household fixed-effects	Yes	Yes

Notes: the outcome variable for fuel-collection time in column (1) is derived from self-reported data on time spent (per day, week, or month) collecting fuelwood, crop residue, leaves, dung, biomass pellets, kerosene, LPG, biogas, and—if relevant—any other fuel used by the household; the results in column (1) are identical to those presented in column (4) of Table 6. In column (2), fuel-collection time is restricted to only traditional fuels (fuelwood, crop residue, leaves, and dung). Standard errors (in parentheses) are clustered at the hamlet level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Source: authors' calculations based on data collected during baseline, midline, and endline survey activities (see Figure 1).

C Redefining 'NGO village'

There is often considerable spatial heterogeneity in the scale and scope of an NGO's activities. The same NGO can be deeply invested in the welfare of one particular community while at the same time only superficially involved with another. If this is the case in our setting, our characterization of NGO and non-NGO villages using a binary variable may be too crude. We, therefore, turn to two additional ways of defining the level of an NGO's involvement with each of our study villages: (1) the number of activities it is leading in a particular community; and (2) the number of years since it first began operating there.

We obtain information on spatial variation in the NGO's portfolio of activities through reviews of its annual reports, newsletters, and other promotional materials.[35] Specifically, we rely on these materials to identify which activities occur in which villages, and when the NGO's operations first began there. This presents challenges as these promotional documents are typically not sufficiently detailed to allow us to comprehensively construct both measures. Recall that our sample contains 38 villages, of which half are NGO villages. In total, we are able to construct a detailed breakdown of the NGO's activities in 12 of its 19 villages; we are able to identify the NGO's commencement year in an equal number of villages. A count of projects among these villages with non-missing implementation data reveals that the NGO leads just over four initiatives (with a minimum of one and maximum of seven) in each of its villages. Similarly, immediately prior to the start of the intervention, the NGO has been operating for just over 15 years in the average village, ranging from four years in the newest village to 25 years in the oldest one.

Using these additional measures (i.e. a count of the number of the NGO's active projects, and the overall age of its engagement in each village) we can investigate heterogeneity in purchase of intervention ICS—now across villages with relatively different 'intensities' of NGO activity. We account for uncertainty introduced by missing data through a simulation-based bootstrap. Specifically, let $\mathbf{NGO} = \mathbf{NGO}_{\text{obs}} \cup \mathbf{NGO}_{\text{miss}}$, where $\mathbf{NGO}$ represents our data (on the village-specific count of NGO projects or the age of its engagement there) and $\mathbf{NGO}_{\text{obs}}$ and $\mathbf{NGO}_{\text{miss}}$ represent non-overlapping observed and missing components of it, respectively. For each bootstrap simulation $n \in N$, we then proceed as follows:

1. Randomly generate:

   (a) $\mathbf{NGO}_{\text{miss}}^{\text{count},n} \overset{\text{iid}}{\sim} \text{unif}(0,7)$

   (b) $\mathbf{NGO}_{\text{miss}}^{\text{age},n} \overset{\text{iid}}{\sim} \text{unif}(0,25)$

2. Construct:

   (a) $\mathbf{NGO}^{\text{count},n} = \mathbf{NGO}_{\text{obs}} \cup \mathbf{NGO}_{\text{miss}}^{\text{count},n}$

   (b) $\mathbf{NGO}^{\text{age},n} = \mathbf{NGO}_{\text{obs}} \cup \mathbf{NGO}_{\text{miss}}^{\text{age},n}$

3. Randomly sample hamlets (with replacement) and estimate the specification outlined in Equation (19):

[35] Recall that the NGO leads activities related to agriculture and forestry (promotion of sustainable agricultural practices, sustainable fodder cultivation, and promotion of culinary herbs), health (local hospitals/clinics), education (local schools), village-level groups (self-help groups, youth groups, and vocational cooperatives), and water management (watershed renewal and spring-water recharge).

(a)

$$Y_{ij}^{\text{count},n} = \beta_0^{\text{count},n} + \beta_1^{\text{count},n} \left(TREATMENT_j\right) + \beta_2^{\text{count},n} \left(\mathbf{NGO}_j^{\text{count},n}\right) + \\ + \beta_3^{\text{count},n} \left(TREATMENT_j \times \mathbf{NGO}_j^{\text{count},n}\right) + \nu_{ij}$$

(b)

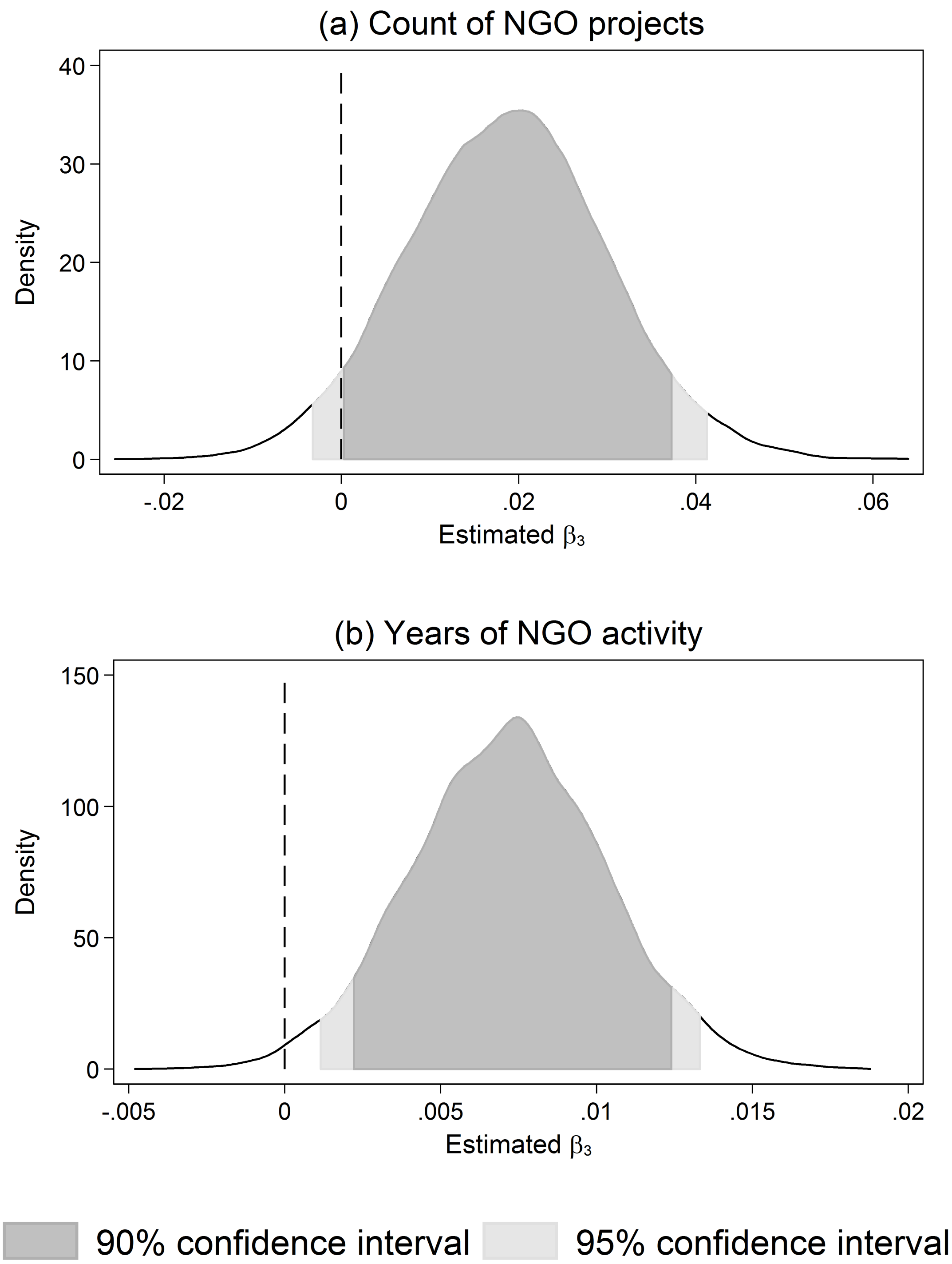
$$Y_{ij}^{\text{age},n} = \beta_0^{\text{age},n} + \beta_1^{\text{age},n} \left(TREATMENT_j\right) + \beta_2^{\text{age},n} \left(\mathbf{NGO}_j^{\text{age},n}\right) + \\ + \beta_3^{\text{age},n} \left(TREATMENT_j \times \mathbf{NGO}_j^{\text{age},n}\right) + \nu_{ij}$$

4. Save the estimated regression coefficients:
   (a) $\hat{\beta}_3^{\text{count},n}$
   (b) $\hat{\beta}_3^{\text{age},n}$

In other words, for villages lacking data on the count of projects, we replace missing count observations with random draws from a uniform distribution over the interval lying between the minimum (i.e. zero in non-NGO villages) and maximum (seven) number of projects observed in each village in our data. For villages lacking data on the age of NGO engagement, we similarly replace missing observations with random draws from a uniform distribution over the interval lying between the minimum (zero in non-NGO villages) and maximum (25) ages in our data. Having replaced these missing observations, we sample hamlets (with replacement) from our study sample to generate a bootstrapped sample, and separately estimate the specification outlined in Equation (19) by replacing our NGO binary variable with the count of NGO activities or the age of NGO engagement. For each specification, we repeat this process 10,000 times to obtain a distribution of the estimated coefficient for the $TREATMENT_j \times NGO_j$ interaction term. We note that this is a relatively conservative approach to dealing with the uncertainty surrounding missing observations. It is almost certainly the case that the NGO has active projects in each of its villages, and has operated in them for at least a few years. Nevertheless, the uniform distributions we use to replace missing observations have positive support over zero (the value these variables are assigned for non-NGO villages). In addition, our randomly generated replacements never exceed the maximum value observed in the (non-missing) data.

Figure C1 presents our results. As shown in panel (a), we find that purchase rates by households in treated NGO hamlets are, on average, approximately 2 percentage points higher for every additional project that the NGO leads in that particular village. This result is statistically significant at the 10 per cent level, as measured by the 90 per cent confidence interval of the distribution of our simulated regression coefficients. Similarly, panel (b) shows that every additional year of the NGO's presence in a village resulted in an increase in rates of ICS purchase by households in treated NGO hamlets by just under 1 percentage point. Broadly, these results serve as a robustness check for our main result (Table 3). They are also consistent with our model and show that NGO activity—defined in a variety of ways—influences the effectiveness of interventions.

Figure C1: Heterogeneity in ICS purchase rates based on alternative definitions of ‘NGO village’



Notes: this figure plots the distribution of 10,000 β_3 coefficients obtained from estimating Equation (19) using a simulation-based bootstrap approach. This coefficient represents the additional impact of the ICS promotion intervention on ICS purchase rates in treated hamlets located in villages with relatively higher levels of NGO activity. In panel (a), the NGO activity variable is a count of the number of active projects being implemented by the NGO in each village; in non-NGO villages, this variable equals 0. In panel (b), the NGO activity variable is the number of years since the NGO first began operating in each village; once again, in non-NGO villages, this variable equals 0. The outcome variable is an indicator that equals 1 if household i in hamlet j purchased at least one of the two ICS promoted during the intervention.
Source: authors’ illustration based on calculations described in Appendix C and data on household-level ICS purchase rates collected during intervention activities (see Figure 1).

D Bayesian analysis

We complement our main analyses using Bayesian methods. Standard frequentist approaches assume that underlying statistical parameters are fixed. Conditional on these fixed parameters, the data are one realization of infinitely many samples, and can be combined with assumptions about large-sample approximations (e.g. asymptotic normality) for inference. In contrast, Bayesian techniques assume that the true values of parameters are random variables, and assign distributions to these parameters based on a-priori information. Conditional on the observed data and the specified prior distribution, a posterior distribution for each parameter can be estimated and used for inference.

D.1 Likelihood

We specify our likelihood function for the data as follows:

$$Y_{ij} \sim \mathcal{N}\left(\beta_0 + \beta_1\left(TREATMENT_j\right) + \beta_2\left(NGO_j\right) + \beta_3\left(TREATMENT_j \times NGO_j\right) + \sum_{j=1}^{97} u_j \gamma_j, \sigma_\nu^2\right), \tag{D.1}$$

where Y_{ij} is a binary variable that equals 1 if household i in hamlet j purchased at least one of the two intervention ICS offered during intervention activities and 0 if it did not; $TREATMENT_j$ is a binary variable that equals 1 if hamlet j is randomly assigned to the treatment group and 0 if it is assigned to the control group; and NGO_j is a binary variable that equals 1 if hamlet j is located in an NGO village and zero if it is in a non-NGO village. In addition, we include hamlet-specific random effects, represented by each of the γ_j terms with coefficients u_j.

D.2 Priors

Model I: Informative prior for β_3 and diffuse priors for other parameters

$$\begin{gathered}
\beta_0 \sim \mathcal{N}\left(0, 10000\right) \\
\beta_1 \sim \mathcal{N}\left(\hat{\beta}_1^{\mathrm{OLS}}, 10000\right) \\
\beta_2 \sim \mathcal{N}\left(\hat{\beta}_2^{\mathrm{OLS}}, 10000\right) \\
\beta_3 \sim \mathrm{Beta}\left(2, 14\right) \\
u_j \sim \mathcal{N}\left(0, \sigma_\gamma^2\right) \\
\frac{1}{\sigma_\nu^2} \sim \Gamma\left(0.01, 0.01\right) \\
\frac{1}{\sigma_\gamma^2} \sim \Gamma\left(0.01, 0.01\right)
\end{gathered}$$

Model II: Diffuse priors for all parameters

$$\begin{gathered}
\beta_0 \sim \mathscr{N}(0, 10000) \\
\beta_1 \sim \mathscr{N}\left(\hat{\beta}_1^{\mathrm{OLS}}, 10000\right) \\
\beta_2 \sim \mathscr{N}\left(\hat{\beta}_2^{\mathrm{OLS}}, 10000\right) \\
\beta_3 \sim \mathscr{N}\left(\hat{\beta}_3^{\mathrm{OLS}}, 10000\right) \\
u_j \sim \mathscr{N}\left(0, \sigma_\gamma^2\right) \\
\frac{1}{\sigma_\nu^2} \sim \Gamma(0.01, 0.01) \\
\frac{1}{\sigma_\gamma^2} \sim \Gamma(0.01, 0.01)
\end{gathered}$$

Model III: Strong no-NGO-effect prior for β_3 and diffuse priors for other parameters

$$\begin{gathered}
\beta_0 \sim \mathscr{N}(0, 10000) \\
\beta_1 \sim \mathscr{N}\left(\hat{\beta}_1^{\mathrm{OLS}}, 10000\right) \\
\beta_2 \sim \mathscr{N}\left(\hat{\beta}_2^{\mathrm{OLS}}, 10000\right) \\
\beta_3 \sim \mathscr{N}(0, 0.01) \\
u_j \sim \mathscr{N}\left(0, \sigma_\gamma^2\right) \\
\frac{1}{\sigma_\nu^2} \sim \Gamma(0.01, 0.01) \\
\frac{1}{\sigma_\gamma^2} \sim \Gamma(0.01, 0.01)
\end{gathered}$$

D.3 Results

Posterior distributions of the parameters in the two models are estimated via Markov Chain Monte Carlo (MCMC) simulation. Specifically, we ran 50,000 MCMC samples after a burn-in period of 10,000 iterations, with thinning every fifth iteration. Table D1 presents our results.

Table D1: Markov Chain Monte Carlo results

	(1) Posterior mean	(2) Posterior standard deviation	(3) 95% credible interval
(a) Model I			
$TREATMENT_j$	0.46	0.055	$[0.34, 0.56]$
NGO_j	0.018	0.063	$[-0.11, 0.14]$
$TREATMENT_j \times NGO_j$	0.12	0.062	$[0.02, 0.26]$
Constant	-0.007	0.048	$[-0.10, 0.09]$
(b) Model II			
$TREATMENT_j$	0.44	0.064	$[0.32, 0.57]$
NGO_j	0.0004	0.082	$[-0.17, 0.16]$
$TREATMENT_j \times NGO_j$	0.15	0.097	$[-0.04, 0.34]$
Constant	0.0002	0.053	$[-0.10, 0.11]$
(c) Model III			
$TREATMENT_j$	0.48	0.056	$[0.37, 0.59]$
NGO_j	0.05	0.066	$[-0.08, 0.19]$
$TREATMENT_j \times NGO_j$	0.08	0.069	$[-0.06, 0.21]$
Constant	-0.02	0.048	$[-0.12, 0.07]$

Notes: the outcome variable is an indicator that equals 1 if household i in hamlet j purchased at least one of the two ICS promoted during the intervention. Columns (1) and (2) present the mean and standard deviations, respectively, for the MCMC sample. Column (3) presents the 95 per cent credible interval for the MCMC sample. Estimates for 97 hamlet-specific random effects not reported for brevity. As in Table 3, $N = 943$ households.

Source: authors' calculations based on procedure described in Appendix D and data on household-level ICS purchase rates collected during intervention activities (see Figure 1).