Online appendix to "Preferences and the effectiveness of behavior-change interventions: Evidence from adoption of improved cookstoves in India"

Marc Jeuland Subhrendu K. Pattanayak Jie-Sheng Tan Soo
Faraz Usmani*

*Jeuland (Email: marc.jeuland@duke.edu; Tel: (+1) (919) 613-4395; Mailing address: Box 90239; Duke University, Durham, NC, 27708, USA) is Associate Professor of Public Policy and Global Health at Duke University and Research Fellow at the RWI–Leibniz Institute for Economic Research, Essen, Germany. Pattanayak is Professor of Public Policy, Global Health, Environment and Economics at Duke University. Tan Soo is Assistant Professor of Public Policy at the Lee Kuan Yew School of Public Policy, National University of Singapore. Usmani is a Postdoctoral Fellow at the Charles H. Dyson School of Applied Economics and Management at Cornell University.

Appendix A Why design policies that are responsive to preference heterogeneity?

A.1 Introduction

We begin by developing a stylized model to motivate the idea that preferences and choices will have an influence on the outcomes for potential beneficiaries of a new policy. Let policy τ be represented by a vector of normalized preference scores along n criteria,

$$\boldsymbol{\Lambda}^{\tau} = \left(\lambda_1^{\tau}, \ldots, \lambda_n^{\tau}\right)'. \tag{A.1}$$

Each $\lambda_k^{\tau} \in \boldsymbol{\Lambda}^{\tau}$ can be thought of as denoting a specific policy attribute, such as pecuniary cost, effectiveness, or the type of provisioning institution. Household i will choose to participate in (or adopt) a policy A if its perceived utility (u) from doing so is at least equal to that derived from non-participation (0). More formally, adoption occurs if:

$$u_i\left(\lambda_1^{A}, \ldots, \lambda_n^{A}\right) \geqslant u_i\left(\lambda_1^{0}, \ldots, \lambda_n^{0}\right). \tag{A.2}$$

If preferences over policy attributes vary across households, the "adoption condition" outlined in equation (A.2) may only hold for some targeted beneficiaries, such that others choose not to participate when A is the only policy option that is available. This result of imperfect adoption or compliance with an intervention is the usual result in real-world interventions. In this case, inclusion of an additional, sufficiently distinct policy B in the choice set may better cater to the preferences of those households j for whom equation (A.2) does not hold. That is:

$$u_j\left(\lambda_1^{B}, \ldots, \lambda_n^{B}\right) \geqslant u_j\left(\lambda_1^{0}, \ldots, \lambda_n^{0}\right) > u_j\left(\lambda_1^{A}, \ldots, \lambda_n^{A}\right). \tag{A.3}$$

When two options A and B are included (relative to A or B alone), overall participation will then be higher, unless option B also strictly dominates option A. The underlying premise of our experiment is that this is not likely to be the case, despite frequent arguments to the contrary—as evidenced, for example, by literature describing the idea that households move consistently up an ordinal energy ladder (Hosier and Dowd, 1987).

A.2 Simulation set-up

To demonstrate the extent to which uptake of a policy may differ as a function of such heterogeneity, we rely on numerical simulations based on the logic of equation (A.3) for an illustrative case where relevant policies only have two distinct attributes. As indicated in Equation (A.1), each $\lambda_i^\tau \in \mathbf{\Lambda}^\tau$ denotes the normalized performance of policy τ for attribute $i \in (1, \ldots, n)$. Instead of exogenously specifying the rate at which agents trade off one attribute against all others, we simply assume that if $\lambda_i^\tau \geqslant \lambda_i^{\tau'}$ for all $i \in (1, \ldots, n)$ then

$$\mathbf{\Lambda}^\tau \succsim_j \mathbf{\Lambda}^{\tau'} \tag{A.4}$$

for all agents j and policies τ, τ'. In addition, if $\lambda_i^\tau \geqslant \lambda_i^{\tau'}$ for all $i \in (1, \ldots, n)$ and $\lambda_i^\tau > \lambda_i^{\tau'}$ for any $i \in (1, \ldots, n)$ then

$$\mathbf{\Lambda}^\tau \succ_j \mathbf{\Lambda}^{\tau'}. \tag{A.5}$$

In other words, agent j weakly prefers policy τ to τ' if each of τ's attribute "scores" is weakly larger than the corresponding value for policy τ', and strictly prefers τ if at least one such value is strictly larger.

We next assume that $n = 2$, that is, each policy is a bundle of only two attributes.[1] Let $\mathbf{\Theta}_c^j = \left(\theta_1^j, \theta_2^j\right)'$ represent the least preferred policy that agent j in community c would be willing to participate in (or adopt), given these two attributes. This policy represents a "threshold." That is, agent j will adopt any policy whose attribute scores are higher than (or equal to) those of the threshold policy; if this is not the case, the policy under consideration will be rejected. Thus, $\mathbf{\Theta}_c^j$ embodies agent j's preferences over policy attributes, which we assume are normalized so as to fall within the interval $(-1, 1)$. To account for heterogeneity over preferences for these attributes, we model the threshold policy for each agent as a draw from a truncated bivariate normal distribution, bounded from above and below by 1 and -1, respectively. That is,

$$\mathbf{\Theta}_c^j = \begin{bmatrix} \theta_1^j \\ \theta_2^j \end{bmatrix} \sim \mathcal{N}\left(\boldsymbol{\mu}_c, \mathbf{\Sigma}_c\right) \tag{A.6}$$

[1] This is purely for illustrative clarity, and the simulation exercise that follows readily generalizes to larger values of n.

where

$$\boldsymbol{\mu}_c = \begin{bmatrix} \mu_1^c \\ \mu_2^c \end{bmatrix}, \tag{A.7}$$

$$\boldsymbol{\Sigma}_c = \begin{bmatrix} \sigma_1^2 & \sigma_{1,2} \\ \sigma_{1,2} & \sigma_2^2 \end{bmatrix} \tag{A.8}$$

$$f\left(\boldsymbol{\Theta}_c^j, \boldsymbol{\mu}_c, \boldsymbol{\Sigma}_c\right) = \begin{cases} \dfrac{\exp\left\{-\frac{1}{2}\left(\boldsymbol{\Theta}_c^j - \boldsymbol{\mu}_c\right)' \boldsymbol{\Sigma}_c^{-1}\left(\boldsymbol{\Theta}_c^j - \boldsymbol{\mu}_c\right)\right\}}{\int_{-1}^{1} \exp\left\{-\frac{1}{2}\left(\boldsymbol{\Theta}_c^j - \boldsymbol{\mu}_c\right)' \boldsymbol{\Sigma}_c^{-1}\left(\boldsymbol{\Theta}_c^j - \boldsymbol{\mu}_c\right)\right\} \mathrm{d}\boldsymbol{\Theta}_c^j} & \text{for } -1 < \theta_1^j, \theta_2^j < 1 \\ 0 & \text{otherwise.} \end{cases} \tag{A.9}$$

In equation (A.9), $f(\cdot)$ represents the density of a truncated normal distribution that has support only over the interval $(-1, 1)$. More crucial are (i) $\boldsymbol{\mu}_c$, in equation (A.7), which represents a vector of community-specific means for the two policy attributes; and (ii) $\boldsymbol{\Sigma}_c$, in equation (A.8), the community-specific covariance matrix associated with agents' preferences over the two attributes. We specify $\boldsymbol{\mu}_c$ as itself consisting of independent draws from a uniform distribution over the full range of possible policy attributes:

$$\mu_1^c, \mu_2^c \sim \mathcal{U}(-1, 1). \tag{A.10}$$

To obtain $\boldsymbol{\Sigma}_c$, we must randomly generate (positive-definite) community-specific covariance matrices. To do so, we first generate a random correlation matrix ($\boldsymbol{R}$) via the method proposed by Joe (2006) and operationalized by Qiu and Joe (2015). We then randomly generate two additional values:

$$\sigma_1', \sigma_2' \sim \mathcal{U}(0, 1) \tag{A.11}$$

and use the resulting vector to construct $\boldsymbol{\Sigma}_c$, as follows:

$$\boldsymbol{\Sigma}_c = \operatorname{diag}\left(\sigma_1', \sigma_2'\right)' \boldsymbol{R} \operatorname{diag}\left(\sigma_1', \sigma_2'\right)'. \tag{A.12}$$

This specification allows us to model heterogeneity in preferences for policy attributes hierarchically:

1. Heterogeneity *across* communities emerges via random generation of community-specific means and covariance matrices in (A.10) and (A.12), respectively, which ensures that the distribution of agents' preferences over the two policy attributes—including how agents see the relationship between the two attributes—is unique to each simulated community; while

2. Heterogeneity *within* communities emerges via random draws from the truncated bivariate normal distribution in (A.9), conditional on the community-specific means and covariance matrices obtained in the first step.

From a policy perspective, this specification also ensures that we remain relatively agnostic about the "correct" underlying distribution of preferences over policy attributes. Indeed, states often implement largescale social-welfare programs in a diversity of settings with distinct geographic, socioeconomic, and cultural characteristics—including within the same country.[2] There is no reason to believe *a priori* that preferences are distributed identically across starkly different contexts.

A.3 Evaluating a single-choice policy

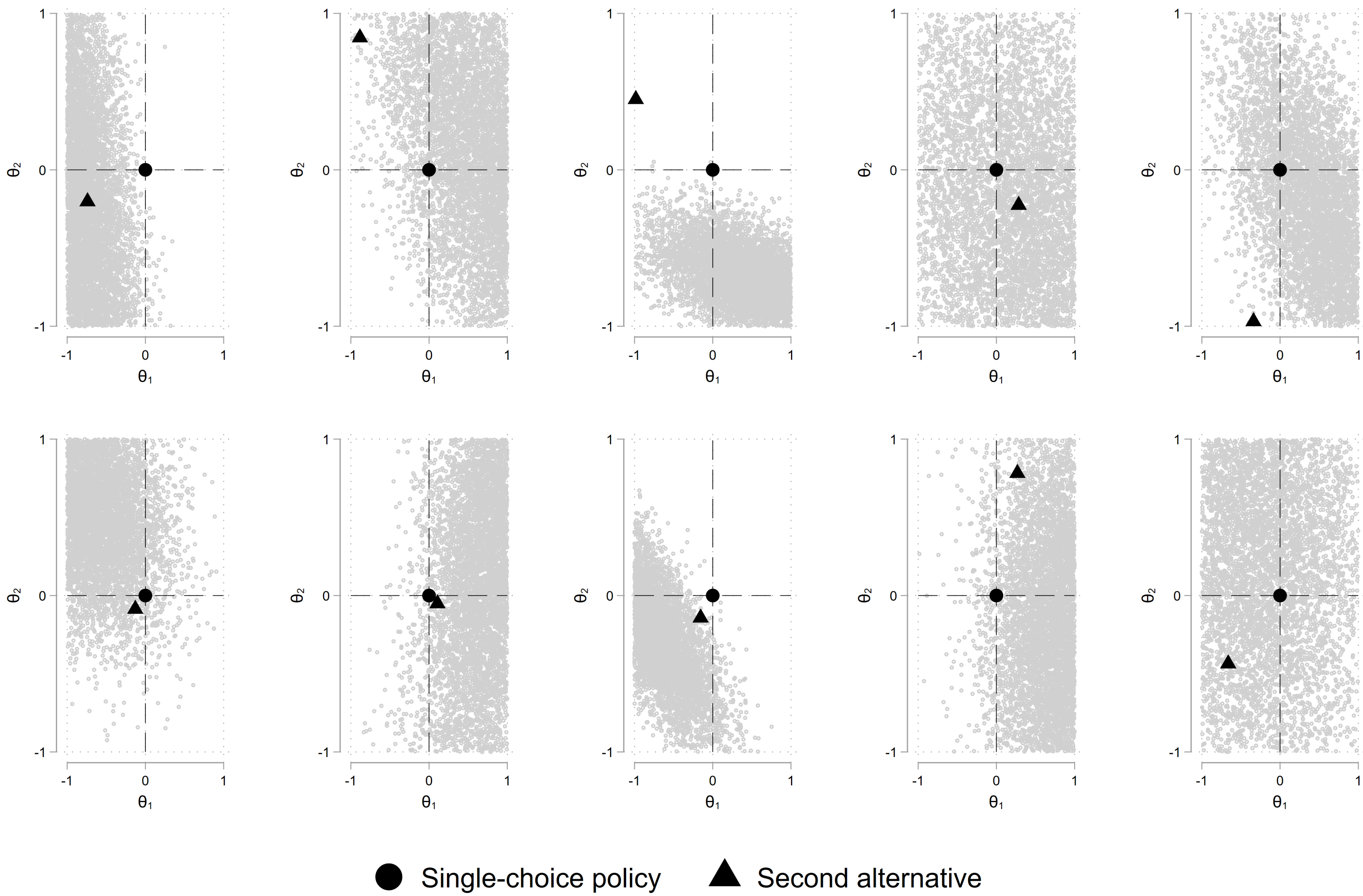
We first generate $\boldsymbol{M}^s = \left(\boldsymbol{\mu}^s_{1,c}, \ldots, \boldsymbol{\mu}^s_{1000,c}\right)$ and $\boldsymbol{E}^s = \left(\boldsymbol{\Sigma}^s_{1,c}, \ldots, \boldsymbol{\Sigma}^s_{1000,c}\right)$ using the distributions and procedures outlined in equations (A.10), (A.11), and (A.12). Conditional on each $\boldsymbol{\mu}^s_{i,c} \in \boldsymbol{M}^s$ and each $\boldsymbol{\Sigma}^s_{i,c} \in \boldsymbol{E}^s$, we then sample agents' preferences over policy attributes $\boldsymbol{\Theta}^s = (\boldsymbol{\theta}^s_1, \ldots, \boldsymbol{\theta}^s_{1000})$ for each community $i \in (1, \ldots, 1000)$. Each $\boldsymbol{\theta}^s_i \in \boldsymbol{\Theta}^s$ is a (5000×2) matrix, representing the threshold policy of 5,000 simulated agents in 1,000 simulated communities.

Recall that an agent prefers one policy to another if its attributes strictly dominate those of the alternative. Without loss of generality, let $\boldsymbol{\Theta}^* = (0, 0)$ represent the attributes of the alternative available for a single-choice policy intervention.[3] With this in hand, we calculate the proportion of agents whose threshold policy is strictly dominated by the single-choice policy to obtain the policy's "adoption rate" in each of the 1,000 simulated communities. Figure A1—which shows the first ten communities (and their corresponding agents) generated in our simulations—provides an illustrative look at this process. The circular point at the origin marks the location of the single-choice policy relative to each agent's threshold policy alternative. The single-choice policy strictly dominates any threshold policy located in the third (bottom-left) quadrant; agents whose thresholds are located in this region adopt the single-choice policy during implementation.

[2] India's National Rural Employment Guarantee Scheme (NREGS)—which entitles every rural household across the country to up to 100 days of paid employment per year, thereby covering roughly eleven percent of the world's population—is a case in point (Niehaus and Sukhtankar, 2013).

[3] This may reflect, for instance, a scenario where a jobs-training program offers only software training for unemployed workers; prospective beneficiaries may either elect to participate (and receive this very specific type of training) or not.

Figure A1: First ten iterations of 1,000 simulated communities with single- and multi-choice policy interventions



A.4 Extending to the multi-choice setting

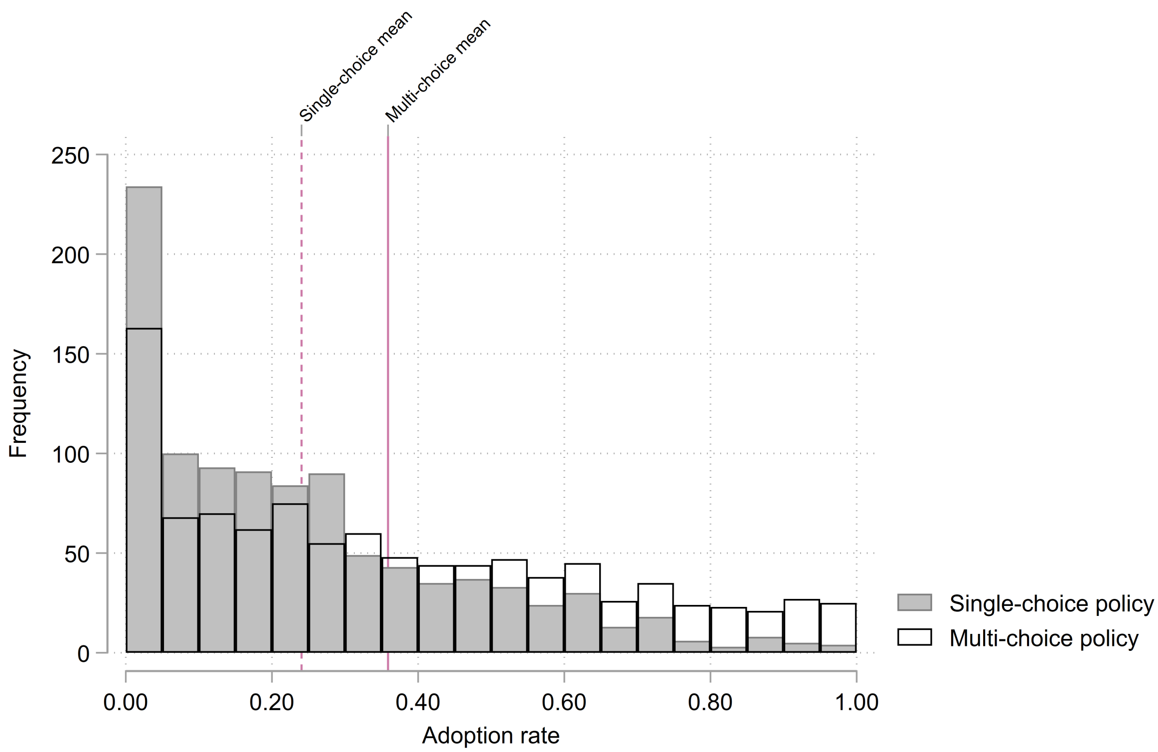
We next extend our simulations to evaluate how adoption rates change relative to the single-choice benchmark when a second alternative is included in the set of options available to potential beneficiaries. Note that given the setup of our model, any alternative located in the first (top-right) quadrant of the panels in Figure A1 would strictly dominate the single-choice alternative and at least weakly increase adoption rates. From a policy perspective, however, it is not always clear that such a "strictly better" alternative necessarily exists. Weak supply chains in remote, rural settings, for instance, limit the availability of alternatives available to policies that aim to enhance the adoption of technologies that deliver welfare benefits (Pattanayak et al., 2019). Policymakers and local implementers (such as non-governmental organizations) may also simply be unaware of the underlying community-level distribution of preferences and thus not be in a position to target policies with such precision. Indeed, there may not even exist viable pathways to glean information about these underlying preferences short of conducting large-scale surveys, an exercise that may quickly become infeasible as the number of relevant policy attributes under consideration rises. Finally, from a conceptual standpoint, certain policy attributes—such as the price and efficiency of a device offered as part of a policy to increase use of energy-efficient technologies—may simply be fundamentally at odds with each other. Assuming that upfront costs are increasing in efficiency, a device that can cater to agents' preferences for both higher efficiency and lower price fully is unlikely to be available in the short- to medium-term, given underlying technological constraints.

Thus, to reflect constraints and uncertainty inherent in policymakers' ability to pick the "right" alternative to include to maximize adoption, we specify a sampling model for the second alternative in a multi-choice policy intervention. Specifically, we assume that the two attributes of the second alternative, $\boldsymbol{\Theta}' = (\theta_1', \theta_2')$, are sampled independently from a uniform distribution over the interval $(-1, 1)$:

$$\theta_1', \theta_2' \sim \mathcal{U}\left(-1, 1\right), \tag{A.13}$$

which is analogous to randomly picking a second alternative to include in the intervention. We then draw $\boldsymbol{\Theta}'^s = \left(\boldsymbol{\theta}_1'^s, \ldots, \boldsymbol{\theta}_{1000}'^s\right)$ from the distribution outlined in (A.13) and repeat the analysis outlined above—with a different second alternative in each iteration, holding the initial alternative constant. In Figure A1, our randomly picked second policy alternative for the first ten sampled communities is represented by a triangle in each panel. Now, agents whose threshold stoves are strictly dominated by the original alternative *or* by the second alternative are assumed to adopt one of the two policy alternatives. In this way, we recalculate the adoption rate for each of the 1,000 simulated communities.

Figure A2: Histogram of adoption rates in single- and multi-choice policies



A.5 Simulation results

Our simulation results are presented in Figure A2. We see that the introduction of a second policy alternative—one that is randomly picked—leads to a considerable shift in the distribution of adoption rates in a multi-choice policy setting relative to the single-choice context. Specifically, the mean adoption rate across all 1,000 communities for the multi-choice setting (represented by the solid vertical line) is over 10 percentage points (40%) higher than the mean adoption rate in the single-choice scenario.[4] These shifts are not limited to a comparison of means. Our simulation results also show, for instance, that one of the two policy alternatives strictly dominate the median agent's threshold—that is, at least half of all agents adopt one of the two alternatives—in approximately 31% of simulated communities. In

[4]Note that as our analysis relies on random simulation of a second policy alternative, it may be understating the benefits of offering additional choices. Indeed, policymakers and promoters are likely to at least make an effort to target available alternatives in ways that ex ante suggest increase policy adoption. To approximate this, we omit all iterations in which the randomly simulated second policy alternative is in the bottom-left quadrant—where it has no influence on adoption rates as it itself is dominated by the original single-choice policy—and find that the mean adoption rate in a multi-choice setting is approximately 40% (over 15 percentage higher than that for the single-choice setting).

contrast, this is only the case in 14% of communities in a single-choice setting. Our simulation results, thus, suggest that in the presence of across- and within-community heterogeneity in preferences over policy attributes, policies that entail the promotion of multiple alternatives may greatly increase rates of adoption by expected beneficiaries.

A.6 Discussion

We can relate the general policy problem discussed above to the more specific example of adoption of preventive health behavior outlined by Pattanayak and Pfaff (2009). In their model, a household's adoption of an environmental health improvement depends on its contribution to utility given a range of relevant constraints, specifically on time and budget resources, and on the nature of the health production function. The decision to adopt is made based on a comparison of the private marginal benefits (monetized based on the contribution of income to utility), which come from direct contributions to utility as well as through reduced illness, and the marginal costs of the investment, which are comprised of the purchase and operating price of the change and the cost of obtaining knowledge about it. In that context, most existing research aims to enhance adoption by exogenously providing a price subsidy, easing maintenance cost, or supplying information, that is, by making changes to the marginal costs of behavior change (Beltramo et al., 2014; Hamoudi et al., 2012). Parameters related to marginal benefits are usually deemed less amenable to intervention and are thus frequently neglected, or considered to be of secondary importance.

Nonetheless, the inherent heterogeneity in the contribution of health or other implications of a technology to household utility can lead to very different adoption decisions and very different outcomes across households and settings (Brown et al., 2017; Heckman et al., 1997). In the environmental health sphere specifically, such heterogeneity in demand and/or outcomes has been demonstrated for a diverse set of technology promotion alternatives including for household water treatment (Brown et al., 2017; Jeuland et al., 2016; Yang et al., 2006), sanitation (Vasquez and Alicea-Planas, 2018), insecticide-treated bednets (Bonan et al., 2017), and even vaccines (Larson et al., 2014). Moreover, systematic analyses have previously described a set of common correlates of adoption behaviors, suggesting that such patterns are at least partially amenable to explanation (Lewis and Pattanayak, 2012). If this heterogeneity can be characterized and acted upon in a meaningful way, it could help enhance the effectiveness of efforts to promote many seemingly beneficial technologies. This paper develops and applies such an approach in the context of households' choice of cooking and heating technologies in low-income settings.

Appendix B Stove fact sheet (with translation)

Appendix C Additional tables of results

Table C1: Balance tests across rebate levels (treatment group only)

Variables	**Mean** Low rebate	**Mean** Medium rebate	**Mean** High rebate	**Normalized difference** (R1 vs. others)	**Normalized difference** (R2 vs. others)	**Normalized difference** (R3 vs. others)
Village has paved road	0.31	0.33	0.29	0.003	0.055	-0.059
Distance to doctor (km)	8.84	9.45	9.70	-0.091	0.023	0.069
Bank facility in village	0.33	0.31	0.31	0.042	-0.024	-0.018
Presence of NGO	0.49	0.52	0.56	-0.097	-0.016	0.100*
Household size	4.87	4.65	4.79	0.077	-0.084	0.016
Education: Head of household (years)	6.36	6.37	6.02	0.042	0.047	-0.067
Education: Primary cook (years)	4.59	5.05	4.62	-0.056	0.097	-0.048
Female head of household	0.28	0.20	0.26	0.102	-0.179**	0.028
Below poverty line household	0.56	0.57	0.55	0.006	0.034	-0.036
Scheduled caste/Scheduled tribe	0.22	0.29	0.27	-0.127**	0.100*	0.031
Household members cold/cough in past two weeks	0.063	0.077	0.091	-0.110	0.001	0.111
Relative wealth (1: Low to 6: High)	2.09	2.12	2.19	-0.072	-0.017	0.114
Household has taken loan in past year	0.12	0.16	0.21	-0.171***	0.011	0.186**
Household is able to save money	0.22	0.27	0.27	-0.123	0.064	0.036
Hours of electricity per day	17.7	16.5	17.0	0.138*	-0.113	-0.006
Log of total expenditure (INR/month)	8.40	8.44	8.44	-0.056	0.028	0.028
Number of cell phones owned	1.30	1.34	1.27	0.002	0.055	-0.054
Total rooms in house	4.68	4.73	4.70	-0.013	0.014	-0.001
Presence of toilet	0.84	0.85	0.83	0.003	0.029	-0.032
Owns/leases agricultural land	0.98	0.98	0.98	-0.001	0.006	-0.005
Most patient respondent	0.50	0.49	0.50	0.002	-0.016	0.014
Most risk-taking respondent	0.42	0.47	0.40	-0.022	0.126	-0.106
Household believes ICS/clean fuels are beneficial	0.29	0.31	0.33	-0.068	0.006	0.063
Believe smoke is unsafe	0.51	0.47	0.52	0.020	-0.097	0.038
Traditional stove ownership	0.98	0.99	0.97	0.043	0.088	-0.134*
Improved stove ownership	0.30	0.31	0.33	-0.046	0.001	0.046
Minutes traditional stove use (minutes/day)	284	281	279	0.025	-0.017	-0.033
Amount of solid fuel used (kg/day)	7.50	7.37	7.82	0.004	-0.042	0.079
Total fuel expenditure (INR/month)	262	243	237	0.036	-0.011	-0.025
N	255	259	248			

Notes. Balance was also assessed by regressing each variable in the left-hand column on treatment status using OLS, clustering standard errors at the hamlet level. Significance of the coefficient for treatment status from these regressions is indicated in the three rightmost columns as follows: *** $p < 0.01$; ** $p < 0.05$; * $p < 0.1$. Rebate was assigned prior to the intervention; the means and comparisons above include only households that ended up receiving a sales offer (results among all households by rebate level are available upon request).

Table C2: Analysis of serial non-response and class 3 membership

Variables	**Serial non-respondent**	**Other respondent**	***N***
Household in class 3	543	169	712
Household not in class 3	0	351	351
N	543	520	1,063

Notes. Serial non-respondents are households who selected the traditional stove alternative in the DCE in all four choice tasks, no matter the attributes of the ICS options.

Table C3: Differential responses to rebate amount, by preference class

Variables	**(1) Visit 1 purchase**	**(2) Visit 1 purchase**	**(3) All purchases**	**(4) All purchases**
Treatment group (exposed to sales)	0.17***	0.18***	0.24***	0.24***
	(0.043)	(0.044)	(0.048)	(0.048)
Treatment × Class 1†	0.092	0.10	0.028	0.040
	(0.092)	(0.091)	(0.096)	(0.095)
Treatment × Class 2†	0.074	0.062	-0.007	-0.031
	(0.091)	(0.093)	(0.092)	(0.092)
Treatment × Rebate × Class 1	0.0019***	0.0019***	0.0020***	0.0020***
	(0.0003)	(0.0003)	(0.0003)	(0.0003)
Treatment × Rebate × Class 2	0.0016***	0.0017***	0.0019***	0.0019***
	(0.0003)	(0.0003)	(0.0003)	(0.0003)
Treatment × Rebate × Class 3	0.0014***	0.0014***	0.0013***	0.0013***
	(0.0002)	(0.0002)	(0.0002)	(0.0002)
Constant	0.00***	-0.13	-0.00**	-0.094
	(0.00)	(0.11)	(0.00)	(0.11)
Other controls‡	No	Yes	No	Yes
Observations	987	987	987	987
R^2	0.327	0.342	0.350	0.368

Notes. Linear probability model. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Standard errors in parentheses; standard errors clustered at the hamlet level. Only includes households who did not attrit from the final sample; results are qualitatively unchanged when all those with sales data are included. †'Class 1' and 'Class 2' are indicator variables denoting assignment to a latent classes 1 and 2, respectively. Class 3 is the omitted class. ‡The other controls include all of those from the complete model in Table 6 (e.g., the basic controls from Table 6, column 3 plus those indicated in the notes below Table 6). None of these were found to be significantly related to purchase; as shown they did not alter the sign or significance of the main results shown here (Column 2). Observations with missing values for these additional covariates are retained in these regressions by filling in median values; results are not sensitive to this procedure.

Table C4: Stove choice among households exposed to sales intervention, including all sales (marginal effects)

Variables	**(1) Basic**		**(2) + Rebate and controls**		**(3) + Rebate-class interactions**	
	Biomass ICS	Electric ICS	Biomass ICS	Electric ICS	Biomass ICS	Electric ICS
Rebate amount (INR)			0.0006***	0.0012***	0.0005***	0.0010***
			(0.0001)	(0.0002)	(0.0002)	(0.0003)
Class 1†	-0.11**	0.26***	-0.077	0.26***	-0.034	0.064
	(0.055)	(0.053)	(0.056)	(0.062)	(0.13)	(0.10)
Class 2†	0.022	0.010**	0.019	0.074	-0.014	-0.030
	(0.027)	(0.049)	(0.030)	(0.056)	(0.087)	(0.12)
Rebate × Class 1					0.0004	0.0020***
					(0.006)	(0.0006)
Rebate × Class 2					0.0007***	0.0015***
					(0.0003)	(0.0004)
Other controls‡	No	No	Yes	Yes	Yes	Yes
Observations	716	716	716	716	716	716
Pseudo-R^2	0.018	0.018	0.131	0.131	0.135	0.135

Notes. Multinomial logit model using all purchases; we report marginal effects at the mean of the sample covariates. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Standard errors in parentheses; standard errors are clustered at the hamlet level. Only includes treatment group households who did not attrit from the final sample; results are qualitatively unchanged when all those with sales data are included. †Class 1 and Class 2 are indicator variables denoting assignment to latent classes 1 and 2, respectively. Class 3 is omitted. ‡The other controls include all of those from the complete model in Table 6 (e.g., the basic controls from Table 6, column 3 plus those indicated in the notes below Table 6). Few of these were found to be significantly related to purchase; as shown they did not alter the sign or significance of the main results shown here. Observations with missing values for these additional covariates are retained in these regressions by filling in median values; results are not sensitive to this procedure.

Table C5: Stove use conditional on purchase, by latent class

Variables	**(1) Basic** Daily use	**(2) + Rebate** Daily use	**(3) + Controls** Daily use	**(4) Electric ICS** Daily use	**(5) Biomass ICS** Daily use
Rebate amount (INR)		0.0007**	0.0008***	0.0004	0.0011***
		(0.0003)	(0.0002)	(0.0003)	(0.0004)
Electricity supply (hours/day)			0.010	0.013**	0.015
			(0.0072)	(0.0063)	(0.013)
Respondent reports high price of firewood			0.0040	0.0085	0.24**
			(0.054)	(0.065)	(0.091)
Class 1†	0.0040	-0.0066	0.040	0.046	0.49**
	(0.073)	(0.077)	(0.082)	(0.082)	(0.23)
Class 2†	-0.031	-0.048	-0.034	-0.066	0.36***
	(0.064)	(0.062)	(0.064)	(0.073)	(0.10)
Constant	0.55***	0.40***	0.30	0.17	0.23
	(0.035)	(0.069)	(0.21)	(0.23)	(0.33)
Other controls‡	No	No	Yes	Yes	Yes
Observations	373	373	373	265	105
R^2	0.0007	0.024	0.114	0.113	0.424

Notes. Linear probability model using all households that purchased a stove. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Standard errors in parentheses; standard errors are clustered at the hamlet level. †Class 2 and Class 3 are variables denoting the probability of assignment to latent classes 2 and 3, respectively. Class 1 probability is omitted (all probabilities sum to 1). ‡The other controls include all of those from the complete model in Table 6 (e.g., the basic controls from Table 6, column 3 plus those indicated in the notes below Table 6). Few of these were found to be significantly related to purchase; as shown they did not alter the sign or significance of the main results shown here (columns 3-5). Observations with missing values for these additional covariates are retained in these regressions by filling in median values; results are not sensitive to this procedure.

Appendix D Scale-adjusted latent class logit

The conditional probability of respondent i choosing alternative j in choice task t, given that she belongs to preference class c and scale group s is:

$$\mathbb{P}\left[C_t^i = j \mid s\right] = \frac{\exp\left(\lambda_s \beta_c X_{jt} + \lambda_s \beta_{0,c} p_{jt}\right)}{\sum_{k=0}^{K} \exp\left(\lambda_s \beta_c X_{kt} + \lambda_s \beta_{0,c} p_{kt}\right)}, \tag{D.1}$$

where β_c is a vector of coefficients for the attributes of the alternatives and λ_s is a scale parameter for the scale group s. This scale parameter λ_s is also a measurement of the uncertainty in estimation of the β coefficients (Magidson and Vermunt, 2007).

Next, the unconditional probability that respondent i belong to class c and scale s are respectively:

$$\mathbb{P}\left[\text{respondent } i \text{ is class } c\right] = \left(\frac{\exp\left(\alpha_c Z_i\right)}{\sum_{c=1}^{C} \exp\left(\alpha_c Z_i\right)}\right) \tag{D.2}$$

$$\mathbb{P}\left[\text{respondent } i \text{ is scale } s\right] = \left(\frac{\exp\left(\gamma_s Z_i\right)}{\sum_{s=1}^{S} \exp\left(\gamma_s Z_i\right)}\right). \tag{D.3}$$

Depending on the researcher's objective, Z_i can be a vector of observable characteristics such as education and income or just a vector of constants. If Z_i contains individual characteristics, this probability expression is expressed as a multinomial form where coefficients are relative to that for the excluded class. In our case, Z_i is replaced with a constant; hence these constants can be interpreted as the overall class and scale share in equations (D.2) and (D.3), respectively.

Combining these expressions, the probability that respondent i chooses alternative j in task t is:

$$\mathbb{P}\left[C_i^t = j\right] = \tag{D.4}$$
$$\sum_{c=1}^{C}\sum_{s=1}^{S}\left(\frac{\exp\left(\alpha_c Z_i\right)}{\sum_{c=1}^{C} \exp\left(\alpha_c Z_i\right)}\right)\left(\frac{\exp\left(\gamma_s Z_i\right)}{\sum_{s=1}^{S} \exp\left(\gamma_s Z_i\right)}\right)\left(\frac{\exp\left(\lambda_s \beta_c X_{jt} + \lambda_s \beta_{0,c} p_{jt}\right)}{\sum_{k=0}^{K} \exp\left(\lambda_s \beta_c X_{kt} + \lambda_s \beta_{0,c} p_{kt}\right)}\right).$$

If we multiply this expression by the probabilities from the other choice tasks, we recover the probability for choosing a particular sequence of alternatives:

$$\mathbb{P}\left[C_i = \left(C^i_{j1}, \ldots, C^i_{jT}\right)\right] = \tag{D.5}$$
$$\prod_{t=1}^{T}\left[\sum_{c=1}^{C}\sum_{s=1}^{S}\left(\frac{\exp\left(\alpha_c Z_i\right)}{\sum_{c=1}^{C}\exp\left(\alpha_c Z_i\right)}\right)\left(\frac{\exp\left(\gamma_s Z_i\right)}{\sum_{s=1}^{S}\exp\left(\gamma_s Z_i\right)}\right)\left(\frac{\exp\left(\lambda_s\beta_c X_{jt} + \lambda_s\beta_{0,c}p_{jt}\right)}{\sum_{k=0}^{K}\exp\left(\lambda_s\beta_c X_{kt} + \lambda_s\beta_{0,c}p_{kt}\right)}\right)\right].$$

This expression is then summed over the entire sample to arrive at the log likelihood:

$$\ln L = \sum_{n=1}^{N}\ln\left\{\mathbb{P}\left[C_i = \left(C^i_{j1}, \ldots, C^i_{jT}\right)\right]\right\}. \tag{D.6}$$

The preference parameters β_c, $S-1$ scale parameters λ_s, and class α_c and scale shares γ_s can be recovered from this log likelihood function either by maximum likelihood estimation or through application of the EM algorithm (Bhat, 1997; Greene and Hensher, 2003; Train, 2008). To allow for the identification of the scale parameters λ_s, one of them is normalized to take a value of one in order for all other scale parameters to be identified.

Using Bayes theorem and the parameters from the likelihood function, we can estimate the conditional probabilities that a respondent belongs to each class c or each scale s or the combination of either given her selection of alternative j:

$$\mathbb{P}\left[\text{respondent } i \text{ is class } c \mid C^i_t = j\right] = \tag{D.7}$$
$$\sum_{s=1}^{S}\frac{\mathbb{P}\left[\text{respondent } i \text{ is class } c \text{ and scale } s\right] \times \mathbb{P}\left[C^i_t = j \mid \text{respondent } i \text{ is class } c \text{ and scale } s\right]^{dijt}}{\mathbb{P}\left[C^i_t = j\right]}$$

$$\mathbb{P}\left[\text{respondent } i \text{ is scale } s \mid C^i_t = j\right] = \tag{D.8}$$
$$\sum_{c=1}^{C}\frac{\mathbb{P}\left[\text{respondent } i \text{ is class } c \text{ and scale } s\right] \times \mathbb{P}\left[C^i_t = j \mid \text{respondent } i \text{ is class } c \text{ and scale } s\right]^{dijt}}{\mathbb{P}\left[C^i_t = j\right]}$$

$$\mathbb{P}\left[\text{respondent } i \text{ is class } c \text{ and scale } s \mid C^i_t = j\right] = \tag{D.9}$$
$$\frac{\mathbb{P}\left[\text{respondent } i \text{ is class } c \text{ and scale } s\right] \times \mathbb{P}\left[C^i_t = j \mid \text{respondent } i \text{ is class } c \text{ and scale } s\right]^{dijt}}{\mathbb{P}\left[C^i_t = j\right]}.$$

Appendix E Results using predicted probabilities of class membership (rather than dichotomous membership)

Table E1: Stove purchase by latent class, using class probabilities

Variables	**(1)** **Basic** Visit 1 purchase	**(2)** **+ Rebate** Visit 1 purchase	**(3)** **+ All controls** Visit 1 purchase	**(4)** **Basic** With later purchases	**(5)** **+ All controls** With later purchases	**(6)** **Basic** Stove use at follow-up§	**(7)** **+ All controls** Stove use at follow-up§
Treatment group (exposed to sales)	0.42***	0.15***	0.16***	0.47***	0.20***	n.a.	n.a.
	(0.036)	(0.040)	(0.040)	(0.038)	(0.043)		
Treatment × Rebate amount (INR)		0.0015***	0.0015***		0.0015***		0.0011***
		(0.0002)	(0.0002)		(0.0002)		(0.0001)
Age of household head			-0.0004		-0.0007		-0.0009
			(0.001)		(0.001)		(0.001)
Female household head			-0.024		-0.024		-0.076
			(0.032)		(0.033)		(0.047)
Education of household head			-0.0008		-0.0000		0.0005
			(0.004)		(0.004)		(0.005)
Perception of relative wealth			0.010		0.0086		-0.022
			(0.017)		(0.017)		(0.023)
Hours of electricity supply per day			-0.0027		-0.0024		0.0032
			(0.003)		(0.003)		(0.004)
Respondent reports high price of firewood			-0.032		-0.057**		-0.036
			(0.030)		(0.027)		(0.036)
Treatment × Class 1†	0.26***	0.21***	0.23***	0.24***	0.22***	0.15*	0.19**
	(0.076)	(0.073)	(0.070)	(0.077)	(0.072)	(0.078)	(0.080)
Treatment × Class 2†	0.14**	0.11*	0.098	0.12*	0.067	0.042	0.021
	(0.065)	(0.063)	(0.064)	(0.064)	(0.061)	(0.054)	(0.051)
Constant	0.00***	-0.00***	-0.13	0.00***	-0.087	0.26***	-0.030
	(0.00)	(0.00)	(0.11)	(0.00)	(0.11)	(0.029)	(0.13)
Other controls‡	No	No	Yes	No	Yes	No	Yes
Observations	987	987	987	987	987	716	716
R^2	0.219	0.323	0.338	0.243	0.364	0.007	0.130

Notes. Linear probability model. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Standard errors in parentheses; standard errors clustered at the hamlet level. The analytical sample is the final sample, after sample attrition (results are qualitatively identical for the full sample with sales intervention data, however). †'Class 1' and 'Class 2' are variables denoting the predicted probabilities of membership in latent classes 1 and 2, respectively. Class 3 probability is omitted. ‡The other controls that are not displayed include all but the covariate "respondent is primary cook" shown in Table 5, which was omitted because the ultimate purchasing behavior should not depend on who participated in the DCE during the baseline survey. None of these controls were found to be significantly related to purchase; as shown they did not alter the sign or significance of the main results (Columns 3, 5 and 7). Observations with missing values for these additional covariates are retained in these regressions by filling in median hamlet-level values; results are not sensitive to this procedure. §The use outcome is only analyzed for households who were in the intervention group; hence the smaller sample size.

Table E2: Differential responses to rebate amount (first sales visit only), using preference class probabilities

Variables	**(1) Visit 1 purchase**	**(2) Visit 1 purchase**	**(3) All purchases**	**(4) All purchases**
Treatment group (exposed to sales)	0.18***	0.19***	0.25***	0.25***
	(0.046)	(0.047)	(0.052)	(0.051)
Treatment × Class 1†	0.047	0.068	-0.061	-0.035
	(0.14)	(0.14)	(0.15)	(0.14)
Treatment × Class 2†	0.065	0.045	-0.031	-0.064
	(0.11)	(0.11)	(0.11)	(0.11)
Treatment × Rebate × Class 1	0.0023***	0.0023***	0.0026***	0.0026***
	(0.0004)	(0.0005)	(0.0004)	(0.0004)
Treatment × Rebate × Class 2	0.0016***	0.0016***	0.0019***	0.0019***
	(0.0004)	(0.0004)	(0.0004)	(0.0004)
Treatment × Rebate × Class 3	0.0013***	0.0013***	0.0012***	0.0012***
	(0.0002)	(0.0002)	(0.0002)	(0.0002)
Constant	0.00***	-0.13	-0.00**	-0.089
	(0.00)	(0.11)	(0.00)	(0.11)
Other controls‡	No	Yes	No	Yes
Observations	987	987	987	987
R^2	0.324	0.339	0.351	0.369

Notes. Linear probability model. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Standard errors in parentheses; standard errors clustered at the hamlet level. Only includes households who did not attrit from the final sample; results are qualitatively unchanged when all those with sales data are included. †'Class 1' and 'Class 2' are denote the predicted probability of membership in latent classes 1 and 2, respectively. Class 3 is the omitted probability. ‡The other controls include all of those from the complete model in Table 6 (e.g., the basic controls from Table 6, column 3 plus those indicated in the notes below Table 6). None of these were found to be significantly related to purchase; as shown they did not alter the sign or significance of the main results shown here (Column 2). Observations with missing values for these additional covariates are retained in these regressions by filling in median values; results are not sensitive to this procedure.

Table E3: Stove choice among households exposed to sales intervention including all sales, using predicted class probabilities (marginal effects)

Variables	(1) **Basic** Biomass ICS purchase	(2) **Basic** Electric ICS purchase	(3) **+ Rebate and Controls** Biomass ICS purchase	(4) **+ Rebate and Controls** Electric ICS purchase	(5) **+ Rebate and Controls** Biomass ICS use	(6) **+ Rebate and Controls** Electric ICS use
Rebate amount (INR)			0.00066***	0.0010***	0.0004***	0.0005**
			(0.0001)	(0.0002)	(0.0001)	(0.0001)
Class 1†	-0.085	0.32***	-0.080	0.34***	-0.023	0.14**
	(0.069)	(0.078)	(0.065)	(0.076)	(0.038)	(0.058)
Class 2†	0.057*	0.083	0.063	0.045	0.091**	-0.027
	(0.032)	(0.060)	(0.038)	(0.068)	(0.038)	(0.037)
Hours of electricity supply per day			-0.003	-0.001	0.001	0.004
			(0.004)	(0.006)	(0.002)	(0.003)
Respondent reports high price of firewood			0.023	-0.078*	0.015	0.001
			(0.024)	(0.041)	(0.022)	(0.029)
Other controls‡	No	No	Yes	Yes	Yes	Yes
Observations	716	716	716	716	716	716
Pseudo-R^2	0.018	0.018	0.131	0.131	0.269	0.327

Notes. Columns 1-4 report results from a multinomial logit model using initial purchase decision only; we report marginal effects at the mean of the sample covariates; columns 5 and 6 are linear probability models for daily use of the biomass and electric intervention stoves, respectively. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Standard errors in parentheses; standard errors clustered at the hamlet level. The analytical sample is the final sample, after sample attrition (results are qualitatively identical for the full sample with sales intervention data, except that class 2 households are significantly more likely to adopt the biomass ICS with that larger sample). †Class 1 and Class 2 are denote the predicted probabilities of membership in latent classes 1 and 2, respectively. Class 3 probability is omitted. ‡The other controls include all of those from the complete model in Table 6 (e.g., the basic controls from Table 6, column 3 plus those indicated in the notes below Table 6). Few of these were found to be significantly related to purchase; as shown they did not alter the sign or significance of the main results shown here (columns 3-6). Observations with missing values for these additional covariates are retained in these regressions by filling in median values; results are not sensitive to this procedure.

References

Beltramo, Theresa, David I. Levine, and Garrick Blalock, "The Effect of Marketing Messages, Liquidity Constraints, and Household Bargaining on Willingness to Pay for a Nontraditional Cook-stove," 2014. Center for Effective Global Action (CEGA) Working Paper № 35. (Cit. on p. 9).

Bhat, Chandra R., "An Endogenous Segmentation Mode Choice Model with an Application to Intercity Travel," *Transportation Science*, feb 1997, *31* (1), 34–48. (Cit. on p. 17).

Bonan, J., P. LeMay-Boucher, and M. Tenikue, "Increasing anti-malaria bednet take-up using information and distribution strategies: evidence from a field trial in Senegal," *Journal of Development Effectiveness*, 2017, *9* (4), 543–562. (Cit. on p. 9).

Brown, Joe, Amar Hamoudi, Marc Jeuland, and Gina Turrini, "Seeing, believing, and behaving: Heterogeneous effects of an information intervention on household water treatment," *Journal of Environmental Economics and Management*, nov 2017, *86*, 141–159. (Cit. on p. 9).

Greene, William H. and David A. Hensher, "A latent class model for discrete choice analysis: contrasts with mixed logit," *Transportation Research Part B: Methodological*, sep 2003, *37* (8), 681–698. (Cit. on p. 17).

Hamoudi, A., M. Jeuland, S. Lombardo, S. Patil, S. K. Pattanayak, and S. Rai, "The Effect of Water Quality Testing on Household Behavior: Evidence from an Experiment in Rural India," *American Journal of Tropical Medicine and Hygiene*, jul 2012, *87* (1), 18–22. (Cit. on p. 9).

Heckman, J. J., J. Smith, and N. Clements, "Making The Most Out Of Programme Evaluations and Social Experiments: Accounting For Heterogeneity in Programme Impacts," *The Review of Economic Studies*, oct 1997, *64* (4), 487–535. (Cit. on p. 9).

Hosier, Richard H. and Jeffrey Dowd, "Household fuel choice in Zimbabwe," *Resources and Energy*, dec 1987, *9* (4), 347–361. (Cit. on p. 2).

Jeuland, M., J. Orgill, A. Shaheed, G. Revell, and J. Brown, "A matter of good taste: investigating preferences for in-house water treatment in peri-urban communities in Cambodia," *Environment and Development Economics*, 2016, *21* (3), 291–317. (Cit. on p. 9).

Joe, Harry, "Generating random correlation matrices based on partial correlations," *Journal of Multivariate Analysis*, nov 2006, *97* (10), 2177–2189. (Cit. on p. 4).

Larson, H.J., C. Jarrett, E. Eckersberger, D.M. Smith, and P. Paterson, "Understanding vaccine hesitancy around vaccines and vaccination from a global perspective: a systematic review of published literature," *Vaccine*, 2014, *32* (19), 2150–2159. (Cit. on p. 9).

Lewis, Jessica J. and Subhrendu K. Pattanayak, "Who Adopts Improved Fuels and Cookstoves? A Systematic Review," *Environmental Health Perspectives*, feb 2012, *120* (5), 637–645. (Cit. on p. 9).

Magidson, Jay and Jeroen K. Vermunt, "Removing the Scale Factor Confound in Multinomial Logit Choice Models to Obtain Better Estimates of Preference," in "Proceedings of the Sawtooth Software Conference" Sawtooth Software 2007, pp. 139–174. (Cit. on p. 16).

Niehaus, Paul and Sandip Sukhtankar, "The marginal rate of corruption in public programs: Evidence from India," *Journal of Public Economics*, aug 2013, *104*, 52–64. (Cit. on p. 5).

Pattanayak, S.K., M. Jeuland, J.J. Lewis, F. Usmani, N. Brooks, V. Bhojvaid, A. Kar, L. Lipinski, L. Morrison, O. Patange, N. Ramanathan, I.H. Rehman, R. Thadani, M. Vora, and V. Ramanathan, "Experimental evidence on promotion of electric and improved biomass cookstoves," *Proceedings of the National Academy of Sciences*, 2019, *116* (27), 13282–13287. (Cit. on p. 7).

Pattanayak, Subhrendu K. and Alexander Pfaff, "Behavior, Environment, and Health in Developing Countries: Evaluation and Valuation," *Annual Review of Resource Economics*, oct 2009, *1* (1), 183–217. (Cit. on p. 9).

Qiu, Weiliang and Harry Joe, *clusterGeneration: Random Cluster Generation (with Specified Degree of Separation)* 2015. R package version 1.3.4. (Cit. on p. 4).

Train, Kenneth E., "EM Algorithms for nonparametric estimation of mixing distributions," *Journal of Choice Modelling*, 2008, *1* (1), 40–69. (Cit. on p. 17).

Vasquez, W.F. and J. Alicea-Planas, "Unbundling household preferences for improved sanitation: A choice experiment from an urban settlement in Nicaragua," *Journal of Environmental Management*, 2018, *218*, 477–485. (Cit. on p. 9).

Yang, Jui-Chen, Subhrendu K. Pattanayak, F. Reed Jonson, Carol Mansfield, Caroline van den Berg, and Kelly Jones, "Unpackaging Demand For Water Service Quality: Evidence From Conjoint Surveys In Sri Lanka," 2006. World Bank Policy Research Working Paper № 3817. (Cit. on p. 9).