Technology Diffusion and Social Networks: Evidence from a Field Experiment in Uganda*

Jiehua Chen† **Macartan Humphreys**‡ **Vijay Modi**§

March 15, 2010

Abstract

We examine the dynamics of technology diffusion in rural Africa. Our focus is on the effectiveness of decentralized marketing to encourage the adoption of energy efficient woodstoves in southwest Uganda. Identifying the effects of a dissemination scheme is rendered difficult by the possibility of spillover effects—that areas that do not receive direct encouragement are nevertheless affected by the intervention. A novel randomization scheme is used to allocate "ambassadors" to communities in a way that can allow for the identification of direct and indirect effects. We provide here initial results from a process that is still underway. Although we find broad adoption of the technology and successful marketing by vendors, our initial results uncover no evidence either for positive direct or indirect effects.

First draft, not for citation or circulation

*We thank colleagues at the Millennium Village Project in Ruhiiri, especially David Siriri, John Okorio, Keneth Rwabinumi, and Maria Gabriela Gonzalez. Thanks to Adam Harris for playing a lead role in implementing the main survey and to Peter van der Windt and Caroline Peters for excellent research support in New York.

†Department of Statistics, Columbia University. Email: jc3288@columbia.edu

‡Corresponding Author. Department of Political Science, Columbia University. Email: mh2245@columbia.edu

§Department of Mechanical Engineering, Columbia University. Email: modi@columbia.edu

1 Introduction

We examine the dynamics of technology diffusion in rural Africa. Our focus is on the effectiveness of decentralized approach to encourage the adoption of energy efficient woodstoves in southwest Uganda. Our interest is in determining the extent to which technology adoption is aided by the presence of local and reliable information sources and to understand how information and technology adoption subsequently spreads. Our results are largely negative in that we do not find evidence that the presence of local disseminators spurs takeup, despite the fact that there is a broad interest in the technology, and disseminators are socially well established and incentivized to encourage take up. Our design helps us to rule out the possibility that the non-result is due simply to the difficulty of identifying positive effects in the presence of spillovers. To speak to this question this research addresses a set of substantive and methodological challenges.

1.1 The substantive challenge

The substantive motivation of this project is to examine the potential for using the Gold Standard or CDM to spur the dissemination of improved woodstoves in East Africa. The primary question is: Can a market for stoves be sustained at a scale and price that allows consumers to benefit from carbon credits? The second question is: What is an optimal way to generate such a market in conditions under which technological diffusion is often slow?

The improved stoves were developed at the Aprovecho Research Center and use about 40% less fuel and reduce emissions by 40-50% while reducing green house gas (GHG) emissions an estimated 40% or about 1-2 tons per year from Laboratory tests. Earlier work has shown that these stoves are preferred over a locally made stove and a three stone fire by 95% of users in this study area of Uganda. The stoves were also tested in the field under local kitchen conditions by cooking Matoke, the common local food, with local biomass fuel available, using paired tests (same food, same amount, same fuel mix, same pot) to compare with cooking on a three-stone fire. These results confirmed savings of 38%.

A key motivation for the project is to examine the potential for poor households to benefit from carbon credits through the Clean Development Mechanism though adoption of stoves. We estimate that with a market size on the order of 50,000 households, households could produce savings that could translate into a $5 subsidy for a $15 stove. A key question then is whether such a subsidy could generate a market for stoves at the subsidized price of $10.

Adoption of this technology could come with economic, health, and environmental benefit. Economic benefits may include savings on purchases, savings on labor time for gathering fuel wood, and savings on labor time for food preparation; these latter savings are likely especially important for women. Health benefits could include reduced respiratory infections and eye diseases. Environmental benefits could include a reduction in the destruction of local forests and a reduction in carbon dioxide emissions.

We have begun our assessment of the market for these stoves at a project site in Ruhiira in the Isingiro District of southwestern Uganda. The region is the site of one of the "Millennium Villages" and as such is subject to numerous interventions. Compared to otherwise similar sites, populations in Ruhiira are likely more exposed to new technologies especially in the areas of agricultural production. The region is a poor area with an estimated annual per capita income of $250. Fuel wood in Ruhiira is extremely scarce; clearing of forests to open land for cropping is estimated to have left only 5% of the land with tree cover (Ruhiira Wood Supply Report). As a result there is a serious shortage of fuel wood; women and children spend many hours searching for fuel wood mainly from tree stumps. Some households are not able to prepare two meals a day, not because there is a lack of food, but because there is a lack of fuel wood to cook the food. Biomass is the main source of fuel for cooking in the region. An energy survey conducted in 2007 showed that 99% of cooking was done with fuel wood and crop residue. 95% of fuel wood is collected, the remainder is purchased.

Initial tests demonstrated interest in the stove among participants. However the diffusion of many simple technologies, such as fertilizers to improve yields, mosquito nets to contain malaria, and condoms to limit the transmission of AIDS, have been surprisingly slow to diffuse in Africa. Part of the reason for the slow diffusion may lie in the ways that information about the effectiveness of these technologies is transmitted through social networks (see for example Isham (2002)). Limited information transmission can slow the spread of a technology that can have real benefits for end users. One approach to overcome this hurdle is to use a marketing approach in which technologies are "seeded" throughout the site in a way that maximizes demonstration effects.

For the present research we hypothesize that a mechanism—the seeding of stove ambassadors—that provides users with detailed information about the stoves from trusted sources is an effective way to set up the market. The study employs a randomized design to test this hypothesis.

1.2 The estimation challenge

Estimating the effectiveness of employing stove ambassadors as decentralized information hubs for generating a market is a methodologically challenging task. The key problem is that although one can directly determine the placement of these ambassadors the effects of these ambassadors may be felt to a large extent in areas beyond the areas in which they are placed. More generally, in the presence of spillovers the treatment of interest is often not the same as the treatment that is directly under the control of the researcher and in particular treatment assignment may be related to covariates.

We are operating in an environment in which it is difficult to isolate units to prevent interference and in which, moreover, the level of interference is itself a quantity of interest. A complication arises in the measurement of spillover effects however because even if assignment to treatment is random, "assignment to spillovers" may not be. To illustrate the problem, suppose that there are only four units and one of them is to be assigned to treatment.

Unit	Location	$D_\emptyset$	$y(D_\emptyset)$	D_1	$y(D_1)$	D_2	$y(D_2)$	D_3	$y(D_3)$	D_4	$y(D_4)$
A	1	0	0	1	3	0	1	0	0	0	0
B	2	0	0	0	3	1	3	0	3	0	0
C	3	0	0	0	0	0	3	1	3	0	3
D	4	0	0	0	0	0	0	0	1	1	3
$\bar{y}_{\text{treated}}$			-		3		3		3		3
$\bar{y}_{\text{untreated}}$			0		1		4/3		4/3		1
$\bar{y}_{\text{neighbors}}$			-		3		2		2		3
$\bar{y}_{\text{pure control}}$			0		0		0		0		0
ATT (direct effect)			-		3		3		3		3
ATT (indirect effect)			-		3		2		2		3

Table 1: Potential outcomes for four units for a null treatment plus four possible treatment profiles, D_1-D_4. In each case D_i represents an allocation to treatment and $y_j(D_i)$ is the potential outcome for (row) unit j given (column) allocation i.

Assume that potential outcomes are as described in Table 1. In this example unit j is allocated to treatment under profile D_j and D_j is selected with probability $\frac{1}{4}$. The table lists the potential outcomes for each of these profiles, in addition it provides the potential outcome for the situation in which *no* unit were treated (these can be thought of as the outcomes of a 'uniformity trial' (Rosenbaum, 2007)). In this example units that are adjacent to treated units are affected by the treatment of their neighbor. Units are not however affected by the treatment status of nonadjacent units. In this case the average direct treatment effect is 3 (indeed it is 3 for all units). The "indirect treatment effects" — specifically the effect for an untreated unit of having one neighbor treated — is 1 for units 1 and 4 and 3 for units 2 and 3; the average indirect treatment effect is 2. Importantly for locational reasons, units 2 and 3 are exposed to spillover effects in 2 schemes whereas units 1 and 4 are only exposed in 1 scheme.

Consider now an analysis that ignored spillover effects. Such an analysis would estimate a direct effect of $\frac{(3-1)+(3-4/3)+(3-4/3)+(3-1)}{4} = \frac{11}{6} < 3$. Because of the spillovers, this approach would underestimate the true treatment effects.

Consider next an analysis that takes into account the fact that adjacent (but only adjacent) units are subject to spillovers but that assumes that exposure to spillovers, like exposure to treatment, is random. In this case the expected estimate of the direct effect would be $\frac{(3-0)+(3-0)+(3-0)+(3-0)}{4} = 3$. Thus the estimate of direct effects is unbiased. The expected estimated of *indirect* effects would be $\frac{(3-0)+(2-0)+(2-0)+(3-0)}{4} = 2.5$. This estimate is biased. The reason is that although each profile is equally likely, because of their location, the sensitive units, 2 and 3, are more likely to be exposed to spillovers than the less sensitive units, 1 and 4.

Consider now two approaches to ensure that indirect effects are measured correctly.

The first approach is to employ a randomization scheme that ensures that units are exposed to indirect and direct treatments in the same way. In this example however the only such scheme involves selecting unit 2 and unit 3 with 0.5 probability each. Moreover in general just as a scheme that ensures equal exposure to direct effects would generally not produce equal exposure to indirect effects, so a scheme that ensured equal exposure to indirect effects is not guaranteed to produce equal exposure to direct effects. Multilevel randomization schemes seek to address this problem in this way (Hudgens and Halloran, 2008); approaches of this form have been adopted by McConnell, Sinclair, and Green (2010) who provide an elegant illustration of the use of a related approach for identifying spillover effects arising in the context of political advertising campaigns, and by Ichino and Schuendeln

(2009), who examine the effect of registration monitoring on corruption in Ghana.

The second approach is to estimate indirect effects by matching only on units that have the same propensity to exposure to indirect effects. In this case under profiles D_1 and D_4 the average indirect effect $3 - 0 = 3$ can be estimated correctly for units 2 and 3 but no estimate can be made for units 1 and 4 since if either one is treated the other is a pure control and the units that are exposed to spillovers are not comparable to these units. A related approach using regression rather than matching, is employed by Miguel and Kremer (2004) and Oster and Thornton (2009). See Humphreys (2010) for a discussion of the biases that can result from regression in this context when treatment effects are heterogeneous. A related approach is to use inverse propensity score weighting, which is especially attractive when the propensities are known with certainty.

We use versions of both approaches in this study, using a randomization scheme that does not require the hierarchical structures used elsewhere to support the first approach and using a combination of inverse propensity score weighting and network data from surveys to generate strata to support the second approach. We emphasize however that a fundamental problem with both approaches is that they require knowledge of the structure of the spillover problem. Just as in the standard situation the SUTVA assumption requires that there are no spillovers, so too here, an analogous assumption is that spillovers exist only over adjacent units. We return to the implications of this kind of assumption in the conclusion.

For either approach there may be a problem which we will refer to as the '*checkerboard problem.*' If *all* control units are approximately equally exposed to spillovers from treated units then the data may exhibit insufficient variation to allow one to compare the effects of more or less exposed units. For any population there may exist some random assignments of treatment such that the resulting profile exhibits the checkerboard problem, and others where it does not. Estimating spillover effects however requires selecting the latter type of profile and not the former.

1.3 Outline

We proceed as follows. In the next section we provide a detailed description of the research design and the strategy employed to estimate spillover effects. We then turn to the main results, which are, in the main, negative. A final section discusses the results and examines a number of possible explanations underlying them. Throughout we focus narrowly on treatment effects and their interpretation; in the Appendix however we provide basic observational data on correlates of purchase, which can help inform interpretations of the patterns we discover.

2 Research Design

We study a situation in which a set of "stove ambassadors" are randomly selected from a set of potential stove ambassadors and entrusted with promoting the stove technology among rural populations living in the Millennium Villages area of South West Uganda. Ambassadors are incentivized to promote stoves, earning a commission for each sale for which they were noted as a referee, and are provided with equipment and training to run 'demonstrations' of the stoves locally. An initial shipment of 500 stoves arrived in the area at the start of this project in August 2009. These 500 stoves were sold out by November and an initial round of survey data collection was undertaken in December. A second shipment is due to arrive in March 2010. In this sense *the data and patterns reported here represent intermediary results from a marketing process that has not yet run its course.* This fact should be taken into account in assessing all reported results.

For analyzing direct and indirect effects our unit of analysis is the *local council area* (LC1) and we conceptualize direct treatment as the presence, for a potential buyer, of a '*stove ambassador*' in their LC1 area; indirect treatment is conceptualized as the presence of ambassadors in *neighboring* LC1s within the study area. We employ a scheme in which units are assigned to treatment in such a way that ensures substantial diversity in the exposure to this indirect treatment. We also examine the effects, for an individual, of knowing a given number of ambassadors, conditional upon the number of potential ambassadors one knows.

A number of features of the design require further explanation.

2.1 Study area

The study area is a collection of 90 local council areas (LC1s) that lie within the Millennium village project area of Ruhiira. These LC1s are the lowest administration unit in Uganda; they contain approximately 100 households and are administered by local committees headed by LC1 chairs. The area includes an estimated 9100 households. A map of these areas is provided in Figure 9 and information on the population distribution is given in Figure 10. As illustrated in Table 2 close social relations are largely contained within LC1s, a point to which we return later.

	I Buyers	II Non-buyers	III Ambassador	I − II	I − III
Share of friends	0.59	0.66	0.65	-0.07	-0.06
from own LC1	(141)	(171)	(90)	(0.2)	(0.36)
Share from	0.18	0.17	0.18	0.01	0.00
neighboring LC1	(141)	(171)	(90)	(0.8)	(0.99)

Table 2: Parochialism. Share of 5 friends living in own LC1 or in neighboring LC1. Number of respondents in parentheses. Column III provides estimates for potential ambassadors

2.2 Treatment

As described above, the simplest conceptualization of treatment treats the LC1 as the unit of analysis and exposes populations within an LC1 to the presence of a stove ambassador. There are however other ways to conceptualize the treatment. More continuously, one can think of treatment as being the social or geographic distance between a buyer and an ambassador or of the exposure of an individual to information about the technology whether directly through one (or multiple ambassadors) or indirectly through other potential buyers. Our design is structured to allow for analysis of a number of these different conceptualizations of treatment. Estimating different treatment effects may however require different assumptions regarding the nature of spillovers. In Section 3 we examine effects of proximity to ambassadors—exposure to which depends on geographic features of LC1s. In section 4 we look at the effects of being "socially connected" to ambassadors—the likelihood of which depends on the extent to which an individual is connected in general.

The treatment itself took the form of engaging a community member, selected by local authorities, to serve as a stove ambassador. Ambassadors received information on the benefits of the stoves and were trained in their use; they were provided with a free stove for demonstration purposes, were incentivized to promote the stoves through a scheme in which for every stove sold with them as a referrer they would gain 1,000 shillings (stoves are sold for 20,000 shillings), finally they were provided the opportunity to buy the stoves themselves at a subsidized price. Ambassadors were local figures, and in many cases were local council chairmen, but, importantly, they were not instructed to promote only in their area. Ambassadors were active in the period from September until stoves were sold out in November.

2.3 Stove Ambassadors

A *stove ambassador* is an individual who is selected by the project to act as a promoter of energy efficient woodstoves. In practice stove ambassadors are selected by chairmen of LC1 areas. In selecting ambassadors LC1 leaders were asked to identify people that (a) have time to do this kind of work (b) are social, have good literacy skills, and are able to be mobile to carry a stove around for demonstrations (c) be genuinely interested in making sure that good stoves are used and so should be especially interested in the needs of households and especially of those that gather fuel wood and cook food (d) be familiar with cooking and hence likely to be especially interested in the benefits of these stoves. On this final point we suggested that women are likely to be good candidates for the role and that potential ambassadors who are well linked in to their communities by playing a role in groups would be good candidates.

2.4 Assignment to treatment

Assignment to treatment is undertaken by using a randomization scheme to determine which 45 of 90 LC1s will receive ambassadors. In practice, rather than randomly sampling LC1 areas for treatment and then identifying ambassadors in each, we identify a *potential* ambassador in each of the 90 LC1 areas and then we randomly sample 45 of the 90 potential ambassadors. Given our interest in spillover effects described above we sought a scheme that satisfies a number of criteria. Specifically we seek a scheme in which:

a) We avoid the checkerboard problem

b) Units are randomly exposed to direct effects

c) There is covariate balance across units that do and do not receive direct treatment

d) Units are randomly exposed to indirect effects

e) There is covariate balance across units that are and are not exposed to indirect effects

The scheme we employ seeks to address all five of these features and does so with varying degrees of success. We avoid the checkerboard problem but still ensure that units are randomly exposed to direct effects and moreover that there is balance between units that are or are not directly treated. There is some variation in the extent to which units are exposed to indirect effects although this variation is minimized subject to a number of desiderata. Moreover we ensure that within similar pairs the probability of exposure to indirect effects is similar. We describe the scheme in more detail next.

2.4.1 Balance

Randomization ensures that possible confounding factors are balanced in expectation; however there is no guarantee of balance in realization, nor as noted above is there a guarantee that random assignment to treatment produces random assignment to indirect effects. To ensure that treated and untreated units are balanced ex post we first divided the set of 90 LC1s into 45 pairs, matched on the basis of characteristics of the potential ambassadors

Before carrying out the random sampling, we conducted surveys for the potential ambassadors in each village, so that we can divide the 90 villages into 45 pairs that are matched with respect to numbers of neighboring villages in the project area, share of LC1 boundary that lies within the project area, the potential ambassadors' gender (57% are female), whether or not they have a cell phone (57% do), whether or not they are a chair of the village (37% are), and whether or not they are in the village committee (73% are). We used optimal non-bipartite matching to generate pairs of potential ambassadors; this method chooses pairs to minimize differences on these dimensions within the pairs; thus typically in a given pair, the two ambassadors are of the same gender, have similar committee membership, LC1 characteristics and so on. We then restricted attention to randomization schemes in which one member from each of these pairs was selected, thus ensuring balance on these dimensions across the directly treated areas. We describe the balance obtained on these dimensions for the indirect treatment as well as for other covariates below.

2.4.2 Assignment probabilities

To produce a scheme that assigns units uniformly to indirect effects we assumed that spillover effects are geographically structured. In particular we assume that the strength of spillovers is a function of the *share of neighboring areas* that receive treatment. We define "sparse" villages as villages that share less than 1/3 of their boundary with villages that have ambassadors. Similarly, "dense" villages are villages that share more than 2/3 of their boundary with villages that have ambassadors. All remaining villages are denoted as "medium" villages.

Our random sampling algorithm is designed to ensure that there are both enough "sparse" villages and enough "dense" villages, so that we have enough power to estimate the indirect effect and to ensure that individual units are assigned with equal probability to direct and indirect treatments (equal across units, not across treatments). Of course these categories are an artifact and the ensuring data can be examined with respect to any other definition of exposure.

The random sampling procedure is defined as follows:

Step 1: Restriction. We randomly select 45 of the 90 potential ambassadors, with one ambassador drawn from each pair. We then calculate the proportions of villages in "dense" areas, "medium" areas and "sparse" areas among the selected 45 villages and the remaining 45 villages.

Table 3: Village Proportions

		Indirect		
		Sparse (S=0)	Medium (S=1)	Dense (S=2)
Direct	Yes (T=1)	0.18	0.14	0.18
	No (T=0)	0.18	0.14	0.18

We then admit a selected profile if the proportion of units in each treatment category are close to the target values in Table 3. The proportions are said to be "close" if the sum of squares of the differences between the calculated proportions from the draw in each of the treatment cells and the targets in Table 3 is smaller than 0.001. The numbers in Table 3 are chosen to minimize the checkerboard problem and ensure that we have enough "sparse" and "dense" villages, and for each kind of village, selected and not-selected numbers are balanced. Figure 1 provides an illustration of 6 profiles that meet these criteria. In each case one can see that the selection structure involves clustering although the nature of the clustering varies greatly across profiles.

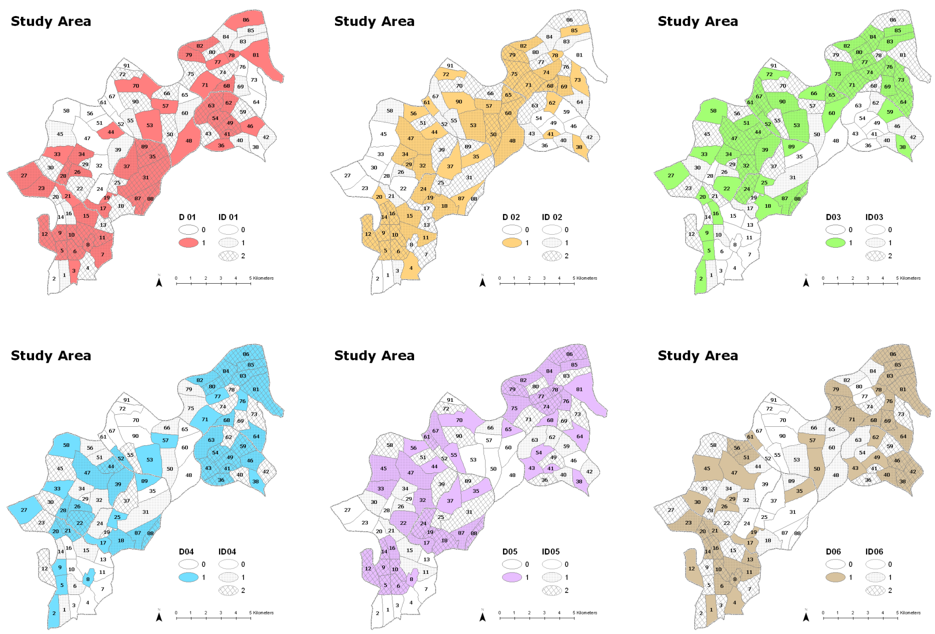

Figure 1: A sample of the 6440 profiles that are admitted to the restricted randomization

This procedure is repeated 20 million times. Of these 20 million draws we found 6,440 "admissible" profiles (none of which is repeated).[1] Ideally, we want each village within each pair to have exactly the same probabilities to be the "dense", "medium" and "sparse" village, and they have exactly 1/2 probability to be selected to have an ambassador. However, the villages have certain geographic characteristics which limits our ability to ensure that

[1]There are 10^{26} ways to choose 45 units from 90, there are $2^{45} \approx 3.5 \times 10^{15}$ ways to do so, restricting choices to one treated unit from each pair

assignment probabilities to each category are uniform.

Step 2: Optimization. We use convex optimization to reduce this imbalance and ensure that each unit has similar probabilities of being exposed directly and indirectly (thus while we are guaranteed to have assignment levels close to those in Table 3 on average, we in fact would like these assignment probabilities for every unit). The loss function we employ sums (a) the quadratic difference between ambassadors in each pair of ambassadors in the assignment probabilities to each of the six treatment combinations (b) the quadratic difference in probabilities, for each unit, in the probability of assignment to direct treatment conditional on assignment to an indirect treatment, and (c) the quadratic difference between the probability of selecting a given profile and the average probability of selecting a profile. The first element maintains balance within pairs in indirect assignment probabilities, the second forces individual assignment probabilities close to those in Table 3, and the third ensures that the probabilities assigned to different profiles are not heavily clustered on a small number of profiles. The latter is important for employing randomization inference.

This procedure then leave us with a well defined probability distribution over selection profiles which we use to ultimately select one profile; the distribution however can be used to calculate propensities precisely and to implement randomization inference.

Figure 2 shows the effects of this optimization procedure for reducing imbalance. On the horizontal is the square of the difference in probability of assignment to each of the indirect treatments between units of a matched pair before optimization (multiplied by 100), and the vertical axis is the square of the differences after optimization (multiplied by 100). All points fall well below the diagonal line, indicating strong improvement of balance due to the optimization step.

Figure 3 shows how the restricted randomization addresses our overall ability to solve the checkerboard problem while maintaining limited dispersion in assignment probabilities. The upper panel shows the assignment probabilities that would obtain if we selected without restrictions, the lower panel shows the assignment probabilities that obtain with our restrictions. The most important feature, which can be seen from the histograms, is that there is a shift to the right in the probability of assignment to the dense (and also to the sparse) indirect treatments. Without restrictions we would have approximately 15% of our data in these categories, in expectation; with restrictions we have 36%. We also see that although the distribution of as-

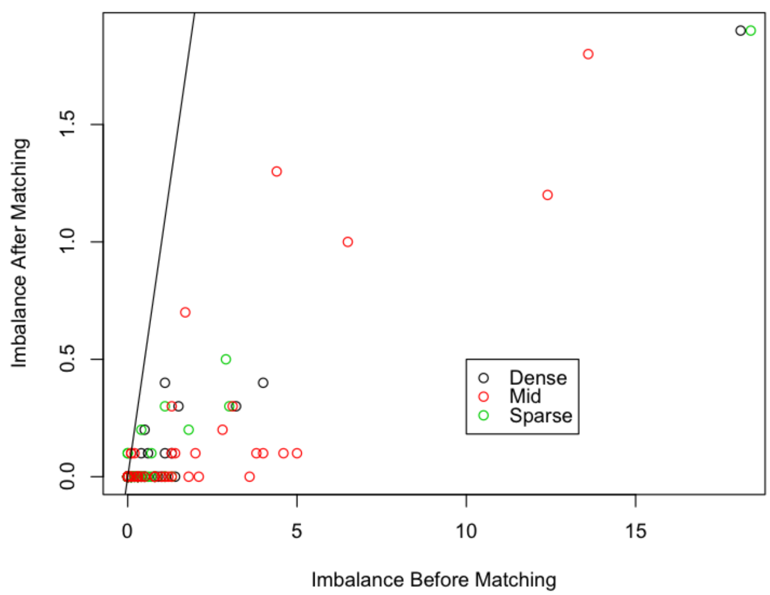

Figure 2: Balance Check after Optimization

signment probabilities is relatively tight, it is less tight under the restricted randomization than under the unrestricted randomization. Thus solving the checkerboard problem has come at a cost of increasing the heterogeneity of assignment probabilities. The figures on the right show however that within matched pairs assignment to indirect treatment is very similar, with probabilities typically within one or two percentage points (average difference between probabilities for each approach are shown in the bottom right of these figures; as can be seen these distances under the restricted randomization scheme with optimization are approximately half the distance under the non-restricted randomization).

Finally we note that in principle the restricted randomization scheme can introduce a correlation structure to our data. It is possible for example that to generate the target distribution, multiple units tend to move together into or out of treatment. Figure 4 illustrates the correlations introduced for direct effects; the figure shows a histogram of the entries of the 90 by 90 matrix of correlations for pairs of units of assignment to treatment over the 6,440 profiles, given the weighting scheme. The correlations, it can be seen, are distributed relatively tightly around zero. The cluster at 1 corresponds to the diagonal of the correlation matrix, the cluster at -1 corresponds to the collection of matched pairs; the symmetry of the histograms follows from the use of matched pairs.

2.5 Selection

The profile ultimately selected from the induced probability distribution is given in Figure 5. By design, half of the units in this profile are directly exposed to treatment, and the indirect exposure is very close to our target, given in Table 3.

2.6 Balance again

In addition to ensuring balance on key covariates ex ante we now can also inspect balance on covariates that were not part of our matching scheme. Table 2.6 suggests that treatment and control groups are indeed balanced on a set of key covariates.

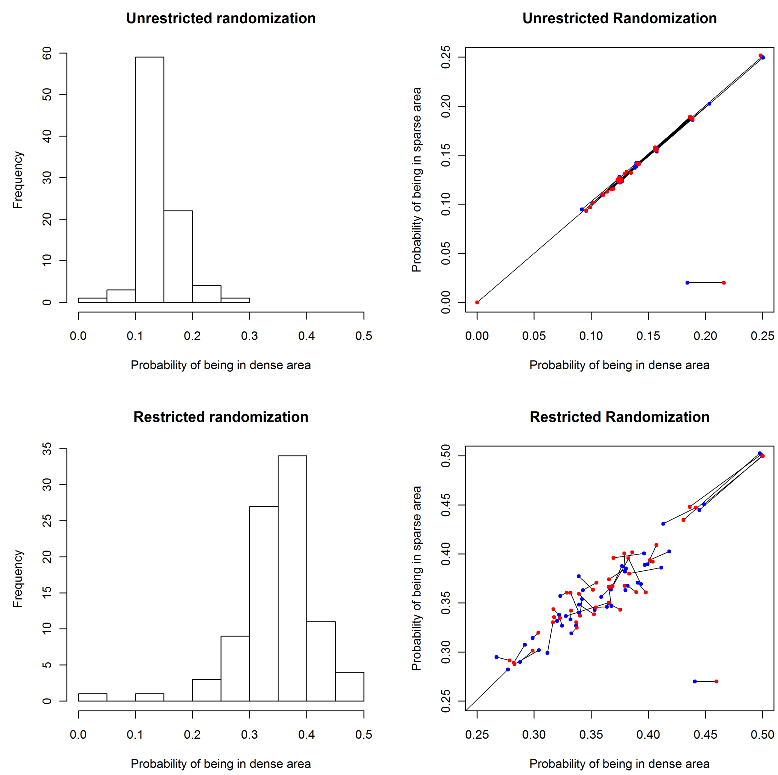

Figure 3: Restrictions. The upper panels show propensities for an unrestricted design, the lower for our restricted design. Propensities to dense and sparse indrect treatments are shifted to the right by the restriction. Propensities become somehwat more dispersed overall but average differences between matched pairs (indicated by segments on lower right of right hand side panels) are narrowed.

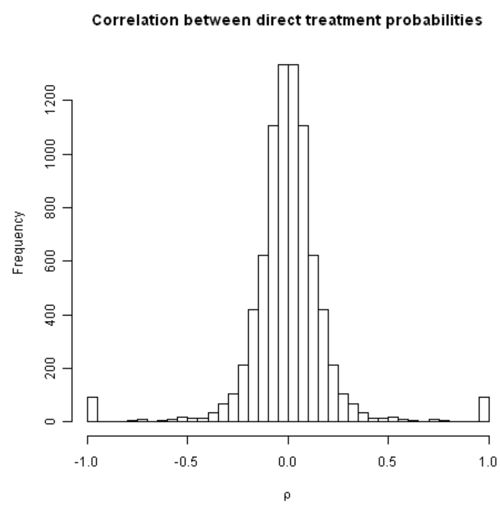

Figure 4: Correlations in assignment to treatment.

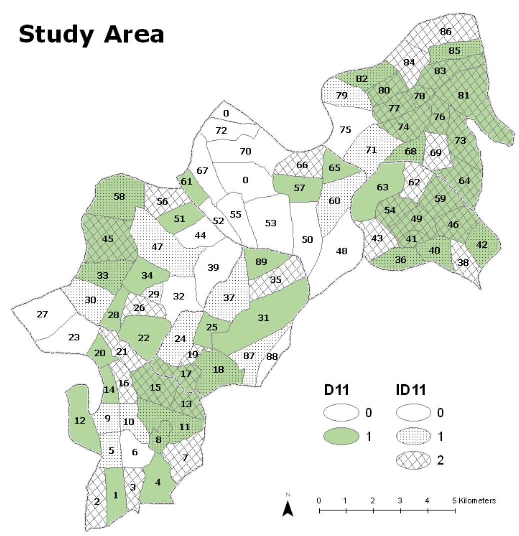

Figure 5: Selected Profile

	Direct				Not Direct			
	Dense	Medium	Sparse	All	Dense	Medium	Sparse	All
LC1 Level Features								
# Neighboring LC1s	5.11	4.18	5.06	4.87	4.31	4.92	4.76	4.64
	(0.96)	(1.47)	(1.73)	(1.42)	(1.45)	(1.44)	(1.64)	(1.51)
Share of boundary in MV	0.92	0.82	0.87	0.87	0.84	0.95	0.83	0.87
	(0.16)	(0.20)	(0.21)	(0.19)	(0.21)	(0.08)	(0.20)	(0.18)
# Households in LC1	79.89	104.73	108.12	96.00	91.31	115.67	114.06	106.40
	(27.56)	(57.82)	(26.21)	(38.29)	(31.32)	(43.63)	(40.57)	(39.21)
# Vendors in LC1	0.28	0.09	0.12	0.18	0.06	0.08	0.24	0.13
	(0.46)	(0.30)	(0.34)	(0.39)	(0.25)	(0.29)	(0.44)	(0.34)
Ambassador level features								
Gender	0.33	0.36	0.56	0.42	0.31	0.50	0.53	0.44
	(0.49)	(0.50)	(0.51)	(0.50)	(0.48)	(0.52)	(0.51)	(0.50)
Cell phone	0.44	0.55	0.31	0.42	0.44	0.42	0.47	0.44
	(0.51)	(0.52)	(0.48)	(0.50)	(0.51)	(0.51)	(0.51)	(0.50)
Chair or not	0.44	0.27	0.44	0.40	0.25	0.42	0.59	0.42
	(0.51)	(0.47)	(0.51)	(0.50)	(0.45)	(0.51)	(0.51)	(0.50)
LC1 Committee	0.22	0.27	0.31	0.27	0.25	0.08	0.41	0.27
	(0.43)	(0.47)	(0.48)	(0.45)	(0.45)	(0.29)	(0.51)	(0.45)
Buyer population features								
Household size	5.77	7.17	6.12	6.34	7.25	5.32	5.26	5.39
	(2.62)	(2.38)	(2.77)	(2.65)	(1.71)	(2.11)	(2.26)	(2.19)
Literacy	8.08	6.50	5.00	6.04	5.25	4.32	6.28	5.36
	(5.79)	(3.96)	(4.01)	(4.50)	(3.50)	(3.72)	(5.11)	(4.52)
Number of Assets	2.31	2.89	3.06	2.86	2.25	2.35	2.49	2.42
	(1.44)	(1.53)	(1.12)	(1.32)	(0.96)	(1.25)	(1.27)	(1.24)
Social Connectivity	1.46	1.72	1.45	1.53	1.50	1.29	1.38	1.35
	(0.52)	(0.46)	(0.51)	(0.50)	(0.58)	(0.46)	(0.49)	(0.48)
N (LC1s)	18	11	16	45	16	12	17	45

Table 4: **Balance**. Cell entries are the mean values of row variables for each treatment condition. Standard deviations are in parentheses.

2.7 Measurement

We benefit from a number of data sources. Prior to selection of ambassadors all potential ambassadors completed a short survey that provides information on their characteristics as well as their linkages to their communities. We refer to this data as the *ambassador survey data*. At the point of sale, vendors were asked to collect basic data on the buyers, including the LC1 from which they are from, and, if applicable, the ambassador that encouraged them to purchase; we refer to this as *the vendor data*. Vendor data was used together with a complete listing of households present in the MV area in order to generate a sampling frame for a buyer and non-buyer survey. From this survey we collected data on approximately 330 households, just under half of which were buyers. We used stratified random sampling using LC1s as strata, requiring at least two-non buyers in each LC1, and predetermining every second survey to be addressed either to the head of household or to the spouse. We refer to the ensuing data as the *household survey data*.

2.8 Estimation

Although our randomization scheme sought to ensure that treatment propensities were similar for all units, as seen in figure 3, there is nevertheless variation across pairs of units. This variation creates the risk that estimated effects from simple differences in means reflect underlying features that determine this variation. To prevent any biases of this form in Section 3 we estimate differences using inverted propensity score weighting—where we know propensities for each unit in each cell exactly (Lunceford and Davidian, 2004); for comparison we report the unweighted results in the appendix.

Following (Fisher, 1935) and others we use our randomization as the basis for inference when conducting statistical tests. To generate significance tests of the sharp null hypothesis of no effect we use our randomization scheme to estimate the probability (given different possible realizations of the lottery but keeping outcomes fixed) that we would observe estimated effects as larger or larger (in magnitude) if the true effect were zero. The use of the known distribution for generating test values appears especially important since the restricted randomization procedure introduces a correlation structure in assignment to treatment.

3 Community Level Analysis

Our first results examine outcomes at the community level and employ the assumption that spillover effects operate along geographic lines such that sales should be higher in treated areas than in untreated areas and should be higher in areas that are dense with ambassadors compared to areas with sparse ambassadors. We begin by looking at the fundamental issue which is the rate of sale of stoves in each LC1 as a function of whether or not they were directly or indirectly exposed to the efforts of the ambassadors.

Table 5 uses complete sales data gathered from vendors to measure the outcome of interest. The results provide no evidence for a positive direct or indirect effect of ambassadors. Strikingly where 'significant' results exist

	Dense	Medium	Sparse	All	Dense − Sparse
Direct	0.027	0.028	0.025	0.030	0.002
	(18)	(11)	(16)	(45)	(0.45)
Not direct	0.016	0.058	0.046	0.032	-0.030
	(16)	(12)	(17)	(45)	(0.02)
All	0.021	0.043	0.036	0.031	-0.015
	(34)	(23)	(33)	(90)	(0.11)
Direct − Not direct	0.011	-0.030	-0.021	-0.002	0.033
	(0.28)	(0.06)	(0.09)	(0.40)	(0.13)

Table 5: **Sales**. Cell entries are the average probability of purchase in a given LC1, where averages are calculated using inverse propensity scores as weights. Final rows and columns provide treatment effects with p-values in parentheses calculated using randomization inference.

these all point in the 'wrong' direction. Among medium and sparse areas, sales were *less* common in LC1s that received treatment; among untreated LC1s, sales were less common in areas with few neighboring ambassadors. Contrary to expectations there is no (positive) evidence of either a direct or indirect effect.

An examination of the map in Figure 6 confirms this pattern—or lack of pattern. We see from the map that large clusters of buyers can be found in sparse untreated areas as well as dense areas, there also exist dense areas (and sparse areas) with no sales at all. Indeed, as we will show below, the evidence suggests that many ambassadors were entirely unsuccessful at generating any sales.

Table 6 provides results consistent with those in Table 5: this time using survey data we examine whether individuals in treated or proximate areas are more likely to claim that they had heard of the woodstoves (more detailed data than this was not collected). Again we find no evidence that the presence of ambassadors led to a rise in information and indeed here we find, perversely, that the most informed areas are those untreated LC1s in zones with few other treated LC1s.

4 Individual Level Analysis

Whereas the treatment was administered at the level of local government areas, in practice the exposure to ambassadors was not constant across in-

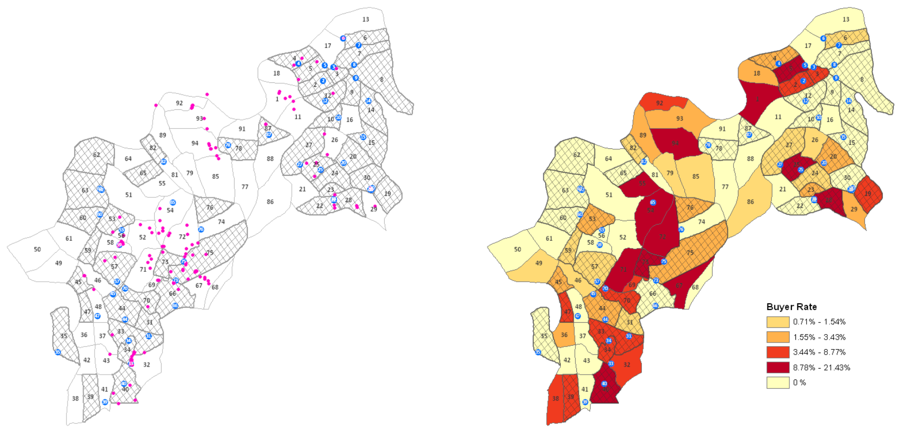

Figure 6: Location of buyers and ambassadors. Left figure shows true locations for all surveyed buyers; right figure shows location information for all buyers accurate only to the LC1 level.

	Dense	Medium	Sparse	All	Dense − Sparse
Direct	0.56	0.41	0.61	0.56	-0.05
	(18)	(11)	(16)	(45)	(0.36)
Not Direct	0.39	0.57	0.83	0.63	-0.44
	(16)	(12)	(17)	(45)	(0.00)
All	0.52	0.40	0.74	0.30	-0.22
	(34)	(23)	(33)	(90)	(0.00)
Direct − Not direct	0.17	-0.16	-0.22	-0.07	0.39
	(0.12)	(0.16)	(0.13)	(0.23)	(0.03)

Table 6: **Knowledge**. Cell entries are the probability that subject has heard of stove. Final rows and columns provide treatment effects with p-values calculated using randomization inference. Responses are based on survey data and adjusted for sampling probabilities and propensity weights.

dividuals but reflected the position of subjects within their communities. Ambassadors, while based on local government areas, were not required to operate in their home areas only. For this reason one can think of the treatment as a more continuous allocation of seeds across a wide space and think of exposure to treatment in a more continuous manner in terms of exposure to these seeds (as well as to other units exposed to these seeds). To what extent any given individual is exposed to ambassadors reflects however both the random selection of ambassadors *and* subject-level characteristics that render one subject more likely than another to be exposed to selected ambassadors.

Just as the likelihood that an LC1 is exposed to treatment from neighboring LC1s is a function of characteristics particular to it, so the likelihood that an individual is exposed to influences from multiple ambassadors may depend on characteristics of the individual—in particular the extent to which they are linked in to social networks.

Table 7 examines the effects of exposure to multiple ambassadors, conditioning on the number of potential ambassadors a subject knows, as measured in the December survey.[2] Conditioning in this way is a form of matching on the true propensity score since conditional upon how many individuals one knows, the number that get treated depend on the randomization scheme—if each potential ambassador gets treated with an equal and independent probability then each unit in a given stratum is equally likely to be exposed to each level of treatment. In our case each ambassador is treated with an equal probability (50%), while assignment to treatment is not uncorrelated, the correlations (as seen in Figure 4) are low and we ignore them here. Figure 11 reports the analogous table for knoweldge of the stoves.

The "trend" estimates in these table are the differences between the estimated probabilities of knowing average numbers of ambassadors in that strata with the probabilities associated with knowing no ambassadors, estimated by logistic regressions. When we estimate the overall "trend", we also include the dummies variables for number of potential ambassadors known. The p-values are calculated using randomization inference.

Again we find no consistent evidence of positive treatment effects, whether direct or indirect. Fur purchases we observe positive effects for some strata and negative for others and in no case are these estimated effects statistically significant. For knowledge all effects are negative and none are significant.

[2]We note that one could imagine that ambassadors become known to individuals precisely because they are ambassadors; we can check this hypothesis however by examining the relative likelihood that an individual knows a treated and an untreated ambassador. The likelihood an individual knowns the ambassador in their area is 56.5% in untreated areas and 57.3% in treated areas.

Number of Potential Ambassadors Subject Knows	Number of known potential ambassadors that are selected								
	0	1	2	3	4	5	6	$\tau_{\max -0}$	Trend
0	0.028	-	-	-	-	-	-	-	-
	(62)	-	-	-	-	-	-	-	-
1	0.030	0.026	-	-	-	-	-	-0.004	-0.003
	(38)	(40)	-	-	-	-	-	(0.345)	(0.46)
2	0.040	0.039	0.063	-	-	-	-	0.023	0.009
	(17)	(40)	(18)	-	-	-	-	(0.201)	(0.28)
3	0.025	0.044	0.044	0.007	-	-	-	-0.018	-0.007
	(8)	(20)	(34)	(8)	-	-	-	(0.33)	(0.35)
4	0	0.013	0.028	0.032	0.058	-	-	0.058	0.015
	(3)	(4)	(9)	(7)	(2)	-	-	(0.144)	(0.44)
5	-	0	0.094	1	-	-	-	1	0.63
	(0)	(1)	(3)	(2)	(0)	(0)	-	(-)	(-)
6	-	-	-	1	-	-	-	-	-
	(0)	(0)	(0)	(2)	(0)	(0)	(0)	(-)	(-)
All	0.028	0.032	0.046	0.034	0.058	-	-	0.029	0.001
	(128)	(105)	(64)	(19)	(2)	(0)	(0)	(0.155)	(0.45)

Table 7: **The effects of individual exposure on propensity to purchase.** Cell entries show estimated purchase probabilities. Final rows and columns provide treatment effects with p-values calculated using randomization inference. Responses are based on survey data and adjusted for sampling probabilities. In the final row the 'trend' estimate includes a set of dummies for the number of potential ambassadors known.

5 Discussion and Conclusion

Our results—that the presence of disseminators is not associated with greater sales or knowledge either locally or in neighboring areas (and may even be associated with perverse outcomes)—is puzzling. It is especially puzzling in a context in which actual levels of sales are high, as are reported knowledge of the stoves and willingness to purchase. Under another design the lack of evidence for ambassador effects might be attributed to spillover effects, but under our present design that argument appears more difficult to make.

In this concluding section we examine a number of possibilities related to this finding, including the possibility that we have mischaracterized ambassador networks and motivations, that heterogeneous effects are being masked in our analysis, that we have insufficient power to detect effects or that spillover effects may be too strong, and finally that the market has not yet cleared and that the early effects of the dissemination are unrepresentative

Number of Potential Ambassadors Subject Knows		Number of known potential ambassadors that are selected								
		0	1	2	3	4	5	6	$\tau_{\max -0}$	Trend
	0	0.45	-	-	-	-	-	-	-	-
		(61)	-	-	-	-	-	-	-	-
	1	0.61	0.40	-	-	-	-	-	-0.21	-0.11
		(35)	(39)	-	-	-	-	-	(0.10)	(0.46)
	2	0.72	0.69	0.48	-	-	-	-	-0.24	-0.09
		(17)	(38)	(15)	-	-	-	-	(0.24)	(0.28)
	3	1	0.78	0.60	0.52	-	-	-	-0.48	-0.19
		(7)	(20)	(33)	(6)	-	-	-	(0.11)	(0.35)
	4	1	1	0.77	0.55	1	-	-	0	-0.12
		(3)	(4)	(8)	(8)	(2)	-	-	(1)	(0.44)
	5	-	1	0.09	1	-	-	-	-	-0.91
		(0)	(1)	(3)	(2)	(0)	(0)	-	(-)	(-)
	6	-	-	-	1	-	-	-	-	-
		(0)	(0)	(0)	(2)	(0)	(0)	(0)	(-)	(-)
	All	0.58	0.61	0.59	0.53	1	-	-	0.42	-0.12
		(128)	(105)	(64)	(19)	(2)	(0)	(0)	(0.15)	(0.05)

Table 8: **The effects of individual exposure on knowledge.** Cell entries show basic knowledge of the stoves from responses to household surveys. Final rows and columns provide treatment effects with p-values calculated using randomization inference. Responses are based on survey data and adjusted for sampling probabilities. In the final row the 'trend' estimate includes a set of dummies for the number of potential ambassadors known.

of the ultimate effects.

5.1 Are our assumptions wrong?

A number of assumptions underpinned our hypothesis of ambassador effectiveness. Two key ones were that (a) ambassadors are locally linked, as are the people living in their communities (b) ambassadors will see value in the stoves and work to encourage not discourage purchases. If instead ambassadors operate primarily outside their regions, or if they actively discourage stove use, then our predictions would clearly be inappropriate. We believe data from this research provides grounds for continuing to expect that our entering assumptions are valid. Table 2 provides consistent evidence that individuals are likely to have very locally based social networks and this is as true for ambassadors as for others. Moreover, our data suggests that while individuals often know ambassadors from neighboring areas, they are significantly more likely to know the ambassadors from their own area than they are proximate ambassadors. For argument (b) we note that the design built in a measure of ambassador interest, by allowing ambassadors to purchase stoves for their own use (at a subsidized price). We take the fact that all ambassadors took up this offer, together with more qualitative assessments, as evidence that ambassadors were not hostile to the technology.

5.2 Heterogeneous Populations and Heterogeneous Effects?

A second possibility is that the overall mixed-to-negative effects are masking different patterns taking place among different subpopulations. Plausibly the marginal effects of local information providers is much stronger for less connected people–indeed that is a motivation for using this type of strategy in the first place. To assess the claim, we examined heterogeneous effects across populations that differ in their social connectedness. Connection is measured by whether they are the Chair of the LC1, whether they are in the LC1 committee or whether they have ever been, how many MVP staff that they know, number of groups that they are involved in, and number of group meetings that they attended for the past two weeks. Every variable is coded so that higher values mean stronger connections. A k-mean algorithm is implemented to separate the households into two groups: "Highly Connected" and "Poorly Connected", which maximizes the between-group variances and minimizes the in-group variances.

The evidence from Table 9 suggests that the aggregate effects are not simply masking a positive effect—the 'wrong' results are found in both subgroups.

In this case however it is also possible to conceive of heterogeneity at the ambassador level. Strictly such heterogeneity would suggest heterogeneity in treatment not necessarily heterogeneity of treatment effects for any one treatment. In effect perhaps our ambassadors constitute a mix of people some of whom are effective and some of whom are not. In this case we might think of an 'intention' to provide information hubs in communities which sometimes is and sometimes is not effective. Under this interpretation the estimates we provide are intention to treat effects and are not the same as the effect of introducing *effective* ambassadors.[3] Using our vendor data we can link

[3]We note that such distinctions should be made with some caution so that one does

"Highly Connected" Subjects	Sparse	Medium	Dense	All	Dense − Sparse
Indirect	0.023	0.052	0.051	0.042	-0.028
	(29)	(19)	(41)	(89)	(0.21)
Direct	0.016	0.101	0.086	0.060	-0.069
	(21)	(42)	(52)	(115)	(0.10)
All	0.019	0.062	0.074	0.024	-0.055
	(50)	(61)	(93)	(204)	(0.03)
Direct − Indirect	0.007	-0.049	-0.034	-0.018	0.042
	(0.30)	(0.11)	(0.24)	(0.22)	(0.18)
"Poorly Connected" Subjects					
	Dense	Medium	Sparse	All	Dense − Sparse
Direct	0.014	0.015	0.025	0.021	-0.012
	(21)	(23)	(24)	(68)	(0.26)
Indirect	0.005	0.070	0.041	0.031	-0.037
	(10)	(14)	(22)	(46)	(0.01)
All	0.010	0.030	0.036	0.013	-0.026
	(31)	(37)	(46)	(114)	(0.02)
Direct − Indirect	0.009	-0.055	-0.016	-0.010	0.025
	(0.33)	(0.00)	(0.21)	(0.12)	(0.24)

Table 9: **Heterogeneous Effects**: Cell entries are the probability of purchase, broken down by strongly connected and weakly connected subjects. Final rows and columns provide treatment effects with p-values calculated using randomization inference. LC1 level responses are based on survey data and adjusted for sampling probabilities. Propensity weights are used to estimate treatment effects for each group.

individual sales to individual ambassadors to determine which ambassadors were effective, which purchasers were influenced directly by ambassadors and which were influenced indirectly or through third sources (such as vendor promotion).

The results (presented in Figure 7) suggest strong heterogeneity across ambassadors (although we note that from this data alone one cannot tell whether cluster patterns reflect ambassador activism or local market conditions). They confirm that to a large extent ambassadors operate locally. But they also suggest that the clustering of sales do not always reflect ambassador influences (there are a set of areas with clusters of purchases in sparse areas which are unrelated to ambassador activity, and in many cases there are clus-

not define a treatment in terms of its effectiveness.

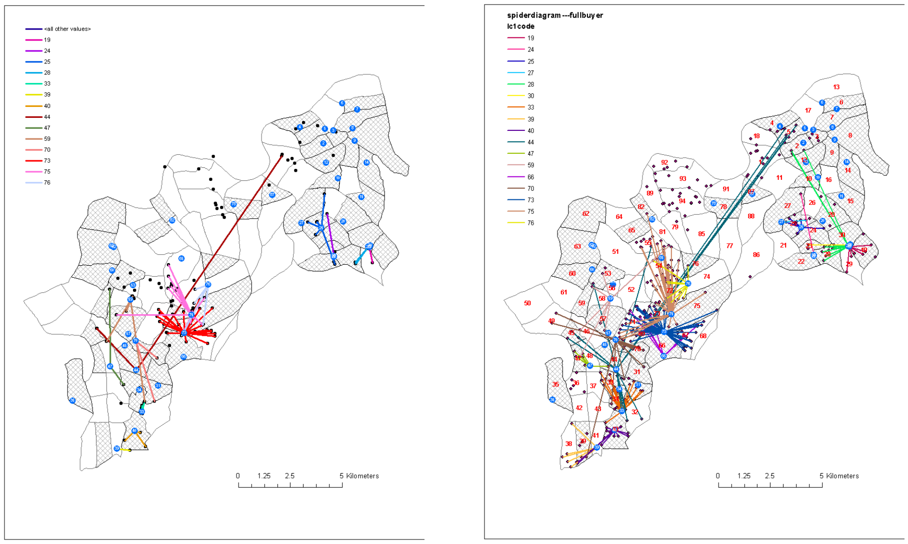

Figure 7: Links between buyers (black points) and ambassadors (blue points). Left figure shows true locations for all surveyed buyers; right figure shows location information for all buyers accurate only to the LC1 level. Black points without connections are buyers that did not provide ambassador reference; blue points without connectors are ambassadors with no referrals.

ters of ambassadors that have not promoted any purchases) and that activist ambassadors are not producing strong local ripple effects—isolated sales do not concentrate in the vicinity of referred sales. Overall these results suggest ambassador heterogeneity and are consistent with the conclusion that while some ambassadors are effective, a large majority are not.

5.3 Are ambassadors *too* effective? Spillovers and power

Another possibility is that there really are positive effects but our power is too weak or spillover effects are too strong for us to detect them. Our previous examination of the actual pattern of sales and referrals suggest that the strong spillover concern is not likely to be the determining factor here; nevertheless we now address this concern together with a related power concern more formally. The power concern is the concern that no evidence of an effect is not the same as evidence of no effect. For this study, power analyses were conducted ex ante, nevertheless one might wonder that the relatively small number of LC1 units and the clustered nature of allocation to treatment make it hard to measure true positive effect. In this case this concern might be coupled with the concern that non-results in the presence of spillovers face a particular interpretational challenge. In particular we face the difficulty of distinguishing between the possibility that outcomes are *identical no matter how treatments are allocated* (assuming some treatment is allocated to at least some units) and outcomes are the same as they would be if no treatments were allocated, with these two possibilities corresponding to what Rosenbaum refers to as "no primary effect" and "no effect" (Rosenbaum, 2007). The challenge is similar to the difficulty of determining a constant of integration: even if we can estimate the marginal effect of a high and low level of exposure with great accuracy, we cannot estimate the total effect unless we have information on outcomes when there is zero exposure. For our purposes the difference is important as it corresponds to the difference between a conclusion that the introduction of ambassadors have no effect at all and the conclusion that ambassadors are uniformly effective in reaching the entire population no matter how they are distributed.

To address these twin concerns we ask the following question: for different possible true direct and indirect effects, what would be the probability that we would observe estimated effects as small or smaller than we do? Let d_{ij} denote the distance between household i and ambassador j (measured from centroid to centroid of the relevant LC1s). Normalize distances such that 1 is the greatest distance that there can be between a household and an ambassador and 0 is the smallest distance. Under this normalization the average distance between a random household and a random ambassador is 0.35, the distance between an individual and an ambassador in their LC1 is 0. Now assume that, absent influences from other ambassadors, the probability that an individual i purchases a stove, due to the influence of a given ambassador j, is given by $\phi_{ij} = \frac{a}{(1+bd_{ij})}$, where a indicates the effectiveness of ambassadors, b is a measure of parochialism — greater values of b imply shorter reach of ambassador influence; zero spillovers occur when $d = \infty$ in which case $\phi_{ij} = a$ if there is direct treatment and 0 otherwise. The upper panel of Figure 8 demonstrates the resulting influence of a given ambassador for $d_{ij} = 0.35$.

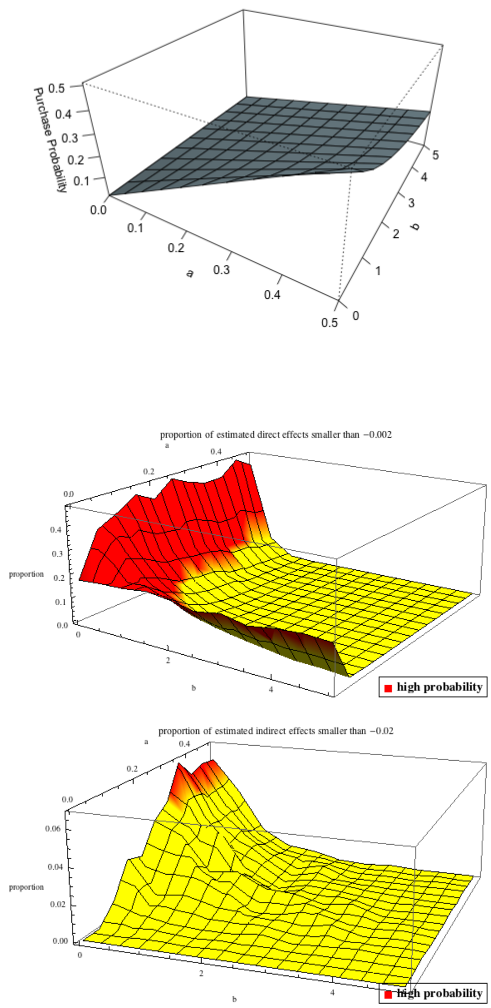

Figure 8: Power and Spillovers. The upper panel shows the relation between effectiveness parameter a and parochialism parameter b on simulated probabilities of influence. The lower figures show the probability that we would find results as small or smaller for different hypothesized values of a and b. Red areas indicate parameter combinations that yield probabilities of observing results similar to ours greater than 10%.

Given this data structure we then simulate, conditional on our randomization, distributions of buyer/seller profiles, under the assumption that each individual will make independent purchasing decisions with probability $p_i = max_j\{\phi_{ij}\}$. Thus here, contrary to our previous analyses, we do not keep potential outcomes fixed but instead consider stochastic outcomes as function of a fixed allocation of treatment.[4] The lower panels show the probability of observing results as low as we do or lower for different values of a and b. The numbers in the tables are the proportions of estimated direct and indirected effects smaller than -0.002 (for direct effects) or lower than -0.02 (for indirect effects)under different parameter settings (see Table 5).

The results are reassuring, suggesting that it is unlikely that we would observe the patterns we do unless true effectiveness were low or spillover effects were very great. For effective ambassadors (a large) we would require a very high level of spillover in order to observe the results we do with moderate probability (for example $a = 0.5$ would require $b \leq .5$ to produce a 10% probability of producing a result so small). Conversely with weak spillovers ambassadors would have to be very ineffective for us to observe the results we see (with $b = 5$ we would require $a \leq 0.05$ to produce a 10% probability of producing a result so small). While it is possible that spillovers operate through channels that we have not yet fully captured, these results suggest that spillover and power problems do not account for our negative findings.

5.4 Is it just too soon to say?

The final possibility we consider is that it is simply too soon to determine what the true effects of the diffusion strategy are. In this analysis we are analyzing results from constrained data—stoves were sold out in November 2009 with all indications that demand is not satisfied. There are reasons to expect that the sample of sales from the constrained market may not be representative of true demand; this could arise for example if a segment of the market, such as those close to vendors, or the highly connected, are able to learn and act more quickly than potential buyers that learn from more decentralized information sources. Initial data suggests slow local diffusion, nevertheless with new stoves arriving on the market in March 2010 we suspend our ultimate conclusions until the market has had time to clear.

[4]One could instead undertake a similar exercise setting a distribution of potential outcomes across areas and examining estimates from different realizations of the lottery.

6 Appendix

6.1 The study site

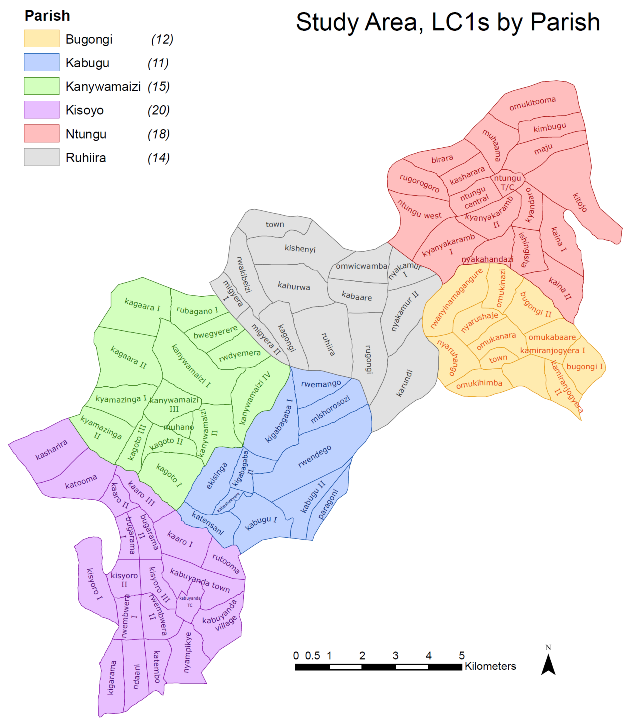

Figure 9: Project Site

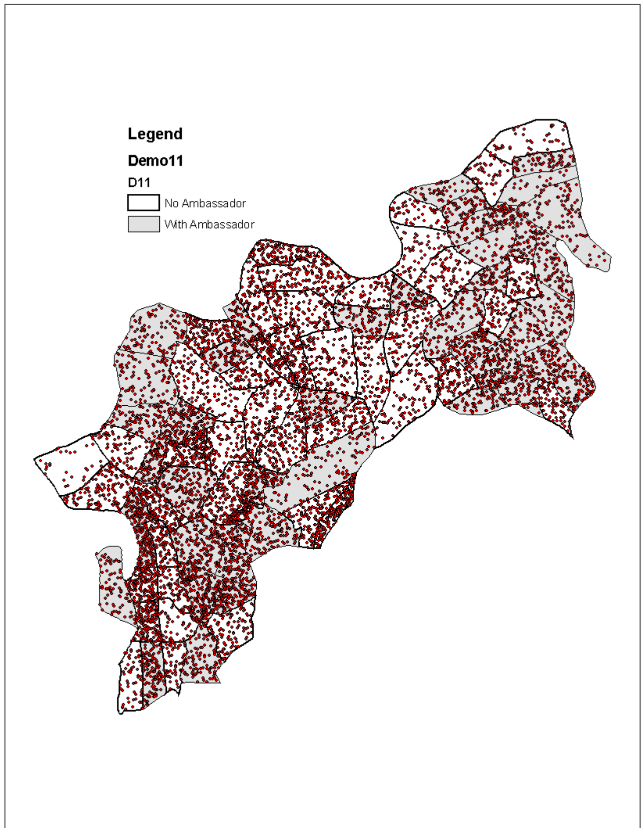

Figure 10: Population Distribution: Each point represents a single household.

6.2 Estimated Effects without Propensity Score Weighting

	Dense	Medium	Sparse	All	Dense − Sparse
Direct	0.027	0.035	0.029	0.030	-0.002
	(18)	(12)	(16)	(45)	(0.21)
Indirect	0.015	0.043	0.041	0.032	-0.027
	(16)	(11)	(17)	(45)	(0.07)
All	0.021	0.039	0.035	0.031	-0.014
	(34)	(33)	(33)	(90)	(0.03)
Direct − Indirect	0.013	-0.008	-0.012	-0.002	0.025
	(0.28)	(0.12)	(0.40)	(0.16)	(0.37)

Table 10: **Sales**. Cell entries are the probability of purchase in a given LC1. Final rows and columns provide treatment effects with p-values calculated using randomization inference.

	Dense	Medium	Sparse	All	Dense − Sparse
Direct	0.60	0.41	0.63	0.56	-0.03
	(18)	(11)	(16)	(45)	(0.48)
Not Direct	0.41	0.50	0.83	0.62	-0.42
	(16)	(12)	(17)	(45)	(0.08)
All	0.53	0.46	0.73	0.59	-0.20
	(34)	(33)	(33)	(90)	(0.06)
Direct − Not direct	0.20	-0.09	-0.20	-0.06	0.39
	(0.22)	(0.48)	(0.31)	(0.28)	(0.22)

Table 11: **Knowledge**. Cell entries are the probability that subject has heard of stove. Final rows and columns provide treatment effects with p-values calculated using randomization inference. Responses are based on survey data and adjusted for sampling probabilities.

6.3 Who are the buyers?

Observation estimates of correlates of sales.

		Bivariate	Multivariate
Household Level Features	Female Headed Households	0.007	0.021
		(0.64)	(1.26)
	Number of Adults Eating	0.003	0.004
		(0.98)	(1.01)
	Education of Head of Household	0.003	0.003
		(2.98)**	(1.76)
	Education of Spouse	0.002	-0.001
		(1.85)	(0.68)
	Main staple is Matoke	-0.001	0.002
		(0.12)	(0.12)
	Household Owns radio	0.007	-0.01
		(0.82)	(0.85)
	Household Owns Cellphone	0.033	0.021
		(4.03)**	(1.74)
Social Connectivity	Individual connected to 'groups'	0.004	0.002
		(0.96)	(0.36)
	Spouse is active in (women's) credit groups	-0.005	-0.005
		(0.43)	(0.36)
	Know local potential ambassador	0.011	0.01
		(1.57)	(1.16)
	Know LC1 Council Members	0.02	0.012
		(2.06)*	(0.88)
LC1 Features	This LC1 has a vendor	0.048	0.042
		(3.37)**	(2.42)*
	LC1 located near MV offices	-0.013	-0.041
		(1.07)	(2.22)*
	LC1 Population	0.04	0.03
		(1.87)	(0.09)
Treatment	Direct exposure	-0.008	-0.017
		(1.12)	(1.65)
	Indirect exposure	-0.014	-0.018
		(3.53)**	(3.26)**
	Constant		0.017
			(0.58)
Observations			311
R-squared			0.03

Table 12: Correlates of Sales. Columns give marginal effects from bivariate and multivariate linear probability models for the probability of purchase. *t*-statistics in parentheses. Regressions weighted by sampling probabilities and clustered by strata.

References

Fisher, R. A. (1935): The Design of Experiments. Oliver Boyd, Edinburgh.

Hudgens, M. G., and M. E. Halloran (2008): "Toward Causal Inference With Interference," Journal of the American Statistical Association, 103, 832–842.

Humphreys, M. (2010): "Bounds on least squares estimates of causal effects in the presence of heterogeneous assignment probabilities," working paper, Columbia University.

Ichino, N., and M. Schuendeln (2009): "Deterring or Displacing Electoral Irregularities? Spillover Eects of Observers in a Randomized Field Experiment in Ghana," Working paper, Harvard University.

Isham, J. (2002): "The Effect of Social Capital on Fertiliser Adoption: Evidence from Rural Tanzania," Journal of African Economies, 11(1), 39–60.

Lunceford, J. K., and M. Davidian (2004): "Stratification and weighting via the propensity score in estimation of causal treatment effects: A comparative study," in Statistics in Medicine, pp. 2937–2960.

McConnell, M., B. Sinclair, and D. P. Green (2010): "Detecting Social Networks: Design and Analysis of Multilevel Experiments," working paper, Yale University.

Miguel, E., and M. Kremer (2004): "Worms: identifying impacts on education and health in the presence of treatment externalities," Econometrica, 72(1), 159–217.

Oster, E., and R. Thornton (2009): "Determinants of Technology Adoption: Private Value and Peer Effects in Menstrual Cup Take-Up," Working Paper, Chicago University.

Rosenbaum, P. R. (2007): "Interference Between Units in Randomized Experiments," Journal of the American Statistical Association, 102, 191–200.