

Subsidies and the African Green Revolution: Direct Effects and Social Network Spillovers of Randomized Input Subsidies in Mozambique

Michael Carter,^{*} Rachid Laajaj,[†] and Dean Yang[‡]

April 17, 2019

Abstract: The Green Revolution bolstered agricultural yields and rural well-being in Asia and Latin America, but bypassed sub-Saharan Africa. We study the first randomized controlled trial of a government-implemented input subsidy program (ISP) in Africa. A *temporary* subsidy for Mozambican maize farmers stimulates Green Revolution technology adoption and leads to increased maize yields. Effects of the subsidy persist in later unsubsidized years. In addition, social networks of subsidized farmers benefit from spillovers, experiencing increases in technology adoption, yields, and beliefs about the returns to the technologies. Spillovers account for the vast majority of subsidy-induced gains. ISPs alleviate informational market failures, stimulating learning about new technologies by subsidy recipients and their social networks.

Acknowledgments: We are indebted to the International Fertilizer Development Corporation for their close collaboration, and in particular we thank Alexander Fernando, Robert Groot and Marcel Vandenberg. Aniceto Matias and Ines Vilela provided outstanding field management. We thank Luc Behaghel, Doug Gollin, Karen Macours, Craig McIntosh, Karthik Muralidharan and seminar participants at the NBER Transforming Rural Africa Workshop in Cambridge, MA, Barcelona GSE Summer Forum 2014, PACDEV 2014, MIEDC 2014, Paris School of Economics, UC Davis, UC San Diego and UC San Francisco for helpful feedback. The BASIS Assets and Market Access Innovation Lab the United States Agency for International Development (grant number EDHA-00-06-0003-00) provided financial support for this study. We obtained IRB approval from the University of Michigan (HUM00033387).

^{*} Department of Agricultural and Resource Economics, University of California, Davis, CA, 95616, USA and Department of Economics, University of Cape Town. ORCID: 0000-0003-0960-9181.

[†] Department of Economics, Universidad de los Andes, Bogotá, Colombia.

[‡] Corresponding author. Email: deanyang@umich.edu. Department of Economics, Gerald R. Ford School of Public Policy, and Institute for Social Research, University of Michigan, Ann Arbor, MI 48109, USA. ORCID: 0000-0002-5165-3816.

1. Introduction

The “Green Revolution” reshaped agriculture across Asia and Latin America during the last four decades of the 20th century, but largely bypassed sub-Saharan Africa. Adoption of high-yielding seeds and chemical fertilizers led to rapid growth of agricultural yields in Asia and Latin America. In contrast, in sub-Saharan Africa, the modest growth in agricultural production that did occur was driven by land expansion, with little technological change, nor yield growth (Evenson and Gollin, 2003). While Green Revolution technology adoption allowed the rest of the developing world to take off economically, sub-Saharan Africa has become the repository of an ever-larger share of the world’s severely poor people. Twenty-five years ago, 17% of the world’s absolutely poor lived in sub-Saharan Africa. Since that time, that figure has risen to 51% (World Bank, 2018).

In response to the Green Revolution that wasn’t, African nations signed the Maputo Declaration in 2003, pledging to invest 10% of their national budgets in agriculture to achieve a 6% rate of annual agricultural growth. The aptly named Alliance for a Green Revolution in Africa (AGRA) was launched in 2006, led by former UN Secretary-General Kofi Annan. To promote the adoption of the Green Revolution technologies that drove the economic take-off of other regions, a number of countries recently introduced (or revived) input subsidy programs (ISPs), which provide Green Revolution technologies (mainly fertilizer and improved seeds) at below-market prices.⁴ These “second-generation” input subsidy programs are intended to be “smart” in the sense that they (i) complement development of private input markets; (ii) target beneficiaries with high potential gains from input adoption; and (iii) are temporary rather than

⁴ Input subsidy programs were widespread in sub-Saharan Africa in prior decades, mainly involving input distribution by state-owned enterprises. These earlier ISPs were discontinued in the context of 1980s and 1990s structural adjustment programs.

permanent. Input subsidy programs are now consuming substantial resources across the continent (Morris *et al.*, 2007). While not all African nations have reached the Maputo Declaration budget targets, ten countries currently implement “second-generation” input subsidy programs, devoting 20-25% of public spending on agriculture (nearly US\$1 billion annually) to such programs. Contrary to the third guiding principle behind “smart” subsidies, these programs appear to be relatively permanent fixtures of implementing governments’ agricultural policies, with no apparent phase-out plans (Jayne *et al.*, 2018).

In the context of this recent wave of African ISPs, fundamental questions remain unresolved. What are the impacts of ISPs on adoption of subsidized technologies, agricultural output, and household well-being? Conclusions of existing observational studies are mixed, and have difficulty establishing causal impacts: no prior research on ISPs has used a randomized control methodology. Other questions have been barely addressed, if at all, in prior studies. If subsidies are only temporary, would impacts persist in later, unsubsidized periods? Do ISP impacts extend to social networks of subsidy recipients? What market failures, if any, are ISPs helping to address?

Answers to these questions are key to credibly estimating the full societal benefits of ISPs, and for optimal policy design. It is important to understand persistence of impacts beyond the subsidized period, to fully assess gains of temporary programs and to understand whether governments can phase out ISPs over time. Relatedly, estimates of spillover impacts to social networks allow a full accounting of societal gains.

In addition, an understanding of market failures helps identify optimal policy responses. For some market failures, subsidies are not the obvious remedy. For example, if farmers cannot finance the technology or bear the additional risk technology adoption implies, then policy

should facilitate markets for financial services (e.g., credit, insurance, or savings) rather than providing subsidies (Karlan *et al.*, 2014). By contrast, informational market failures (say, imperfect information on the returns to the technology) may call for subsidies that overcome individuals' reluctance to learn-by-doing. Because information is non-rival, it may readily spill over to social network connections of subsidy recipients, raising others' adoption and magnifying societal gains from the subsidy (Foster and Rosenzweig, 1995). Information market failures may only justify temporary subsidies that can be removed once induced adopters learn about the technology. On the other hand, if persistent behavioral biases such as present-bias inhibit adoption, permanent interventions may be needed (Duflo *et al.*, 2011).

To contribute to these open questions, we conducted the first and (so far) only randomized controlled trial of a "second-generation" ISP. We study the Farmer Support Program (*Programa de Suporte aos Produtores*), a European Union-funded program for Mozambican farmers managed by the UN's Food and Agriculture Organization (FAO) and implemented by Mozambique's Ministry of Agriculture with technical advice from the International Fertilizer Development Center. The program is representative of ISPs that have recently spread in Sub-Saharan Africa. Our results therefore have direct bearing on understanding the impact of ISPs in the region more generally.

We find that a temporary subsidy for Mozambican maize farmers promotes Green Revolution technology adoption and increases maize yields. While the subsidy was provided for just a single input package in one agricultural season, effects of the subsidy are persistent in later unsubsidized years. Magnitudes of impacts are substantial, but comfortably within the range of potential yield impacts of Green Revolution technologies (as detailed below in Section 5.2). We also observe spillovers from subsidized farmers to their social networks: agricultural contacts of

subsidized farmers also see increases in technology adoption and yields. Spillovers account for the vast majority of subsidy-induced gains. Both subsidized farmers and their social networks report higher beliefs about expected returns to the technologies. We interpret these results as revealing that ISPs help to reduce market failures related to information, stimulating learning about new technologies by subsidy recipients and their social networks.

While our main contribution is providing the first causal estimates based on a randomized controlled trial of new generation of input subsidy programs in Africa, we also contribute a new combination of findings to the literature on technology adoption in developing countries. In this context, ours is the first paper to show that a temporary subsidy for agricultural production technology has lasting impacts on adoption after the subsidy ends. In non-agricultural technology contexts, persistent impacts of a temporary technology subsidy on adoption have been found for health goods (Dupas, 2014) and for labor migration (Bryan *et al.*, 2014). One prior non-experimental study failed to find that ISPs lead to persistent adoption post-subsidy (Ricker-Gilbert and Jayne, 2017). Recent randomized studies of Green Revolution agricultural inputs in Africa do not investigate post-subsidy persistence (Duflo *et al.* 2008, Beaman *et al.* 2013, Abate *et al.* , 2018). Social learning about technologies has been documented previously in prior studies, using both observational approaches (e.g., Bandeira and Rasul 2006, Munshi 2004) and randomized study designs (e.g., Magnan *et al.* 2015, Laajaj and Macours 2016, Beaman and Dillon 2018, Beaman *et al.* 2018 and BenYishay and Mobarak forthcoming), but none of these have been in the context of input subsidy programs.

In addition, prior randomized studies showing social learning have been researcher-implemented interventions, not government-implemented programs. Muralidhran and Niehaus (2017) argue that more randomized evaluations of government-implemented programs are

needed, to enhance external validity of findings. Similar to the findings of Baird et al. (2016) in the context of a deworming program, we find that the benefit-cost ratio of the program increases substantially when spillovers are taken into account, highlighting the importance of incorporating long-term and spillover effects in program evaluation.

This paper is organized as follows. In Section 2, we provide further detail about the subsidy program. In Section 3, we describe the randomized research design and the study sample. Section 4 details the empirical regression specification. We present the empirical results in Section 5. In Section 6, we provide a cost-effectiveness calculation. We provide concluding thoughts in Section 7. An Online Appendix provides additional analyses and robustness tests, to which we refer throughout the main text.

2. Mozambique's Input Subsidy Program

Mozambique's Farmer Support Program provided once-off input subsidies to 25,000 smallholder farmers in five provinces. The program embodied key "smart" features considered ISP best practices: it supported development of input markets by providing the subsidies via vouchers to be redeemed at private agricultural dealers, targeted farmers thought to have high potential gains from the inputs, and provided only temporary subsidies. The technology package we study was designed for maize production.⁵ The subsidy provided a 73% discount on a package of chemical fertilizer (50 kg of urea and kg of NPK 12-24-12) and improved maize seeds (12.5 kg of either a hybrid or an open pollinated improved variety), valid for one use in the 2010-11 agricultural season. Subsidy users had to make the 27% co-pay (863 MZN or about \$US32) when redeeming the voucher at an agricultural dealer.

⁵ Nationally, the program provided 15,000 subsidy vouchers for maize and 10,000 for rice production. Our study occurred in a maize-growing area.

While program design benefited from international technical advice, the Mozambican government had sole responsibility for implementation. The agricultural extension agency identified the beneficiary farmers and distributed the subsidy vouchers. We therefore estimate impacts of an actual government-implemented program, rather than a potentially unrepresentative researcher-implemented intervention (Muralidhran and Niehaus, 2017).

The randomized controlled trial of the input subsidy program was designed by the research team in cooperation with the Ministry of Agriculture and IFDC. Lists of eligible farmers were created by government agricultural extension officers, with input from local leaders and agro-input retailers. Individuals were eligible for a voucher coupon if they met the following program criteria: 1) farming between 0.5 hectare and 5 hectares of maize; 2) being a “progressive farmer,” defined as a producer interested in modernization of their production methods and commercial farming; 3) having access to agricultural extension and to input and output markets; and 4) being able and willing to pay for the remaining 27% of the package cost. Only one person per household was allowed to register. Extension officers informed participants that a lottery would be held and only half of those on the list would win a voucher. Vouchers were then randomly assigned to 50% of the households on the list in each locality. In other words, localities served as treatment stratification cells.

Randomization was conducted by the research team on the computer of one of the PIs, and the list of voucher winners was provided to agricultural extension officers. Extension officers were responsible for voucher distribution to beneficiaries. Voucher distribution occurred at a meeting to which only farmers who won the lottery were invited. Random assignment and distribution of vouchers occurred in September to December 2010. Vouchers were intended to be used for inputs for the 2010-11 season. The annual agricultural season in Mozambique runs from

November (when planting starts) through end of the harvest period the following June. Vouchers expired on January 31, 2011, and this expiration date was strictly enforced. The vouchers were assigned to specific individuals, and names were verified by the input retailer when redeemed.

3. Research Design and Sample

To estimate causal effects, the government collaborated with us to randomly assign subsidies among eligible farmers in 32 villages in Manica province. Within each village, the agricultural extension agency identified farmers eligible for subsidy vouchers (using the same criteria used in the over-arching program). The research team randomly selected half of farmers in each study village to receive subsidy vouchers. These farmers comprise the treatment group, and the remainder comprise the control group.⁶

The sample consists of 514 farmers (247 and 267 in the treatment and control groups, respectively). Agricultural extension officials informed study participants that vouchers would be assigned via random lottery in study villages, announced lottery winners, and distributed vouchers accordingly. We implemented surveys of treatment and control participant households, tracking outcomes in the subsidized 2010-11 season and two annual agricultural seasons afterwards, 2011-12 and 2012-13 (Online Appendix B gives further detail on the sampling and tracking of respondents). We collected data on social network connections, asking each participant to identify others in the village with whom they discuss agriculture (Conley and Udry, 2010). The subsidy was assigned randomly, so controlling for the size of one's social network, the extent to which members of one's social network received the subsidy was also random. We are therefore able to establish the causal impact of the subsidies on a beneficiary's

⁶ Online Appendix A gives further detail on the study context and locality definitions.

own technology adoption over time, and on adoption by social network contacts of beneficiaries. We also examine direct and spillover impacts on agricultural production, consumption and learning about returns to the subsidized inputs.

Table 1 presents key baseline summary statistics. Less than a quarter of households used any chemical fertilizers in the season that preceded the subsidy, and slightly more than a half planted improved maize seeds. Farmers' experience with chemical fertilizer prior to the study appears to be quite limited. We also asked farmers how many years they used fertilizer out of the last 10 years (prior to the beginning of the study). 67% of farmers reported zero years, and 87% reported two years or less (panel C). Based on these reports about prior use of the technologies, there appears to be room for learning about chemical fertilizer, and perhaps less about improved seeds.

Average (median) maize yields are 975 (600) kilograms/hectare, indicating a large yield gap relative to yield expectations from agronomic trials of three to four times that level. Median per-capita consumption (measured using a standard LSMS instrument) is just above the World Bank's standard \$1.95 a day poverty line. Finally, we can see in panel C of the table that our network survey instrument registers substantial variation in the extent to which study respondents were connected to other voucher winners, with 44% registering no connections, and another 23% registering three or more network members who received the voucher.

As is common in studies of real-world programs, we have imperfect compliance with treatment assignment. Only 40.8% of farmers in the treatment group used their vouchers. Most such non-compliance was due to inability to make the input package co-payment (even though claimed ability to pay was a participant selection criterion). Moreover, 12.4% of control group farmers reported using subsidy vouchers for the input package, due to imperfect compliance by

extension agents distributing vouchers (see Online Appendix D). Imperfect compliance with treatment assignment reduces statistical power, but does not threaten internal validity of estimates. The intervention is therefore an “encouragement design” that affects the probability of using a subsidy voucher. The difference in voucher use rates in the treatment and control groups (28.8 percentage points) is statistically significantly different from zero (p-value < 0.001, Appendix Table A3). This treatment-control difference in subsidy voucher use drives all treatment effect estimates.

4. Empirical Specification

Our study focuses on three key outcome variables: use of Green Revolution inputs, learning about returns to those inputs, and living standards. We estimate “intent to treat” (ITT) effects, for the season in which the subsidy was offered, as well as for subsequent unsubsidized seasons (to study post-subsidy persistence of impacts). We also seek to measure spillover effects to social network contacts of treated farmers, over the same time periods. We estimate the following regression equation for outcome variable y_{ict} of household i in locality c in time period t :

$$(1) \quad y_{ict} = \alpha_{Dur}Treat_{ic}Dur_t + \alpha_{Aft}Treat_{ic}Aft_t + \sigma_{Dur}Soc_{ic}Dur_t \\ + \sigma_{Aft}Soc_{ic}Aft_t + \mathbf{X}_{ict}\boldsymbol{\gamma} + \theta_c + \varepsilon_{ict}$$

Outcome variables include measures of input use, yield, consumption and beliefs about the returns of the input package (see Online Appendix E for detailed variable definitions). Time periods are the 2010-11 subsidized season, and two subsequent seasons (2011-12 and 2012-13) when no subsidy was offered. Right-hand-side variables are all indicators (equal to 1 if so; 0 otherwise).

TABLE 1—BASELINE DESCRIPTIVE STATISTICS

A. Percentages for key binary outcomes							
	%						
Used fertilizer on maize	23%						
Used improved maize seeds	54%						

B. Descriptive statistics of continuous variable							
	Mean	Standard Deviation	Percentiles				
			5th	25th	median	75th	95th
Fertilizer used on maize (kg)	23.9	61	0	0	0	0	150
Improved maize seeds used (kg)	19	30	0	0	10	25	75
Maize yield (kg/ha)	975	1,168	104	300	600	1,169	3,305
Expected yield with technology package (kg / ha)	1,944	2,540	221	603	1,163	2,176	6,216
Daily consumption per capita (MZN/day/hh member)	77	51	29	45	63	92	172

C. Frequency of discrete variables							
	Mean	0	1	2	3	4	5 or more
Number of years using fertilizer out of 10 years before the study	1.03	67%	13%	6.6%	3.6%	1.8%	7.5%
Number of social network contacts who are study participants	3.17	31%	16%	12%	7.2%	8.6%	25%
Number of social network contacts in treatment group	1.54	44%	18%	15%	8.2%	5.7%	8.8%

Notes: Data are from the 514 households that were surveyed in all three of the annual surveys. Fertilizer, seed, maize yield, and expected yield are in kilograms. Daily consumption per capita is in Mozambican meticaais (27 MZN \approx 1 USD). Continuous variables are truncated at their 99th percentile prior to calculation of means and standard deviations.

$Treat_{ic}$ indicates treatment group households. Dur_t indicates the observation is in the subsidized (“during”) time period, and Aft_t the subsequent (“after”) time periods. Soc_{ic} indicates the household has above-median (two or more) social network contacts who were randomized into the treatment group (we discuss below the rationale for this specification). Social network contacts are defined as those with whom the participant discussed agriculture in the season prior to the subsidized 2010-11 season.⁷ Households with $Soc_{ic} = 1$ have 3.63 treatment group contacts on average; for households with $Soc_{ic} = 0$, the average is 0.29.

The vector of controls \mathbf{X}_{ict} includes indicators for having one, two, three, four, or five or more social network contacts who are study participants (omitted category zero), to control for effects of social network size. Social network size is not exogenously determined, and is mechanically positively correlated with Soc_{ic} (households with more contacts will also have more treatment group contacts). When controlling for social network size, the regression coefficients on the terms including Soc_{ic} can be interpreted as the causal effect of having two or more social network contacts who were offered the input subsidy. To capture common time effects, \mathbf{X}_{ict} includes an indicator for each time period, and interactions between the time period indicators and the social network size indicators (allowing social network size effects to vary over time). θ_c are locality fixed effects (treatment is randomized within locality). ε_{ict} is a mean-zero error term. We employ robust (heteroscedasticity-consistent) standard errors.

⁷ Following Conley and Udry’s (2010) elicitation of “information links”, study participants were presented with the full list of other study participants in the same village, and asked one by one whether they talked about agriculture with this person in the prior season.

Random assignment led to balance on key time-invariant household characteristics with respect to $Treat_{ic}$ and Soc_{ic} .⁸ Survey attrition was low (8.6% on average across rounds) and uncorrelated with $Treat_{ic}$ and Soc_{ic} .⁹

5. Empirical Results

This section presents intention to treat estimates for equation (1) above. Given the statistical and especially significance of the estimated social network effects, we also examine alternative network specifications and test for alternative mechanisms that might underlie our findings.

5.1 Regression results

Regression results are presented graphically in Figure 1, and regression coefficients are reported in Appendix Table A4. Outcome variables in the regressions are expressed in logarithms (findings are robust to alternate dependent variable specifications, such as indicators, kilograms, or Mozambican meticaais as shown in Online Appendix F and Appendix Table A5). In the left-hand column of Figure 1, coefficients represent ITT effects of subsidy assignment on households in the treatment group, during and after the subsidized season (α_{Dur} and α_{Aft} , respectively). The right-hand column of Fig. 1 presents σ_{Dur} and σ_{Aft} , spillover impacts on study participants “more connected” to the treatment group (with above-median treatment group

⁸ See Table A1 in Online Appendix C.

⁹ Table A2 in Online Appendix C, which reports these results.

contacts), during and after the subsidy, respectively. We will first discuss direct impacts before turning our attention to the spillover impacts.

The direct effect of the subsidies on treatment group members is an increase in technology adoption and maize yields (coefficients in the left-hand side of the figure). Direct effects during the subsidized period (α_{Dur}) are large and positive for adoption of fertilizer and improved seeds, as well as for maize yields. In the after-period (α_{Aft}), treated households show some persistence in use of fertilizer, but not seeds, which can be due to the fact that improved seeds were more widely used and known than fertilizer before the program. Direct impacts on fertilizer use after the subsidy become smaller in magnitude, which is to be expected after the end of the subsidy, but they remain substantial and statistically significant at the 1% level. Direct impacts on yields remain almost as high in the period following the subsidy, which can be due to the farmers re-using the inputs when not subsidized, and also to persistence in the benefits of fertilizer used in the subsidized season though nutrients remaining in soils. Returns to fertilizer can also increase because of learning about how to use fertilizer, or a selection effect, where only farmers who observed high yields purchase the inputs after the subsidy period.

We also show impacts on living standards, measured by per capita daily consumption in the household (Figure 1, fourth row). Consumption is useful to examine as a summary measure of household well-being. It is additionally useful because we do not have a measure of agricultural profits, which would require data on all agricultural inputs used (in particular difficult-to-measure labor inputs). Examining impacts on consumption can therefore indirectly reveal whether agricultural profits rose. Direct impacts on the treatment group are close to zero in the “during” period, but large and positive in the “after” period. Spillover impacts are large

and positive, with magnitudes that are relatively stable across periods. These results provide an indirect indication that unobserved agricultural profits did rise and benefit households.

Finally, to examine learning, we estimate impacts on beliefs about expected yields with the technology package. To do this, for each study participant, after identifying their main maize plot, we asked what production the farmer would expect in this plot if they used the technology package on this parcel in 1) a normal year, 2) a very good year, and 3) a very bad year. We then asked the farmer to say, on average, out of 10 years, how many are very good years, very bad years, and normal years. This set of simple questions allows us to calculate the expected yield when using the technology package. The last row of Figure 1 shows that being a direct recipient of the subsidy significantly increased the yield expected by the farmer when using the technology package. The positive effects on expected returns are stable in the “during” and “after” periods (in terms of magnitudes and statistical significance).

In addition to positive direct effects on treated households, there are substantial spillover effects from treated households to their social network contacts. We find no statistically significant spillover impacts during the subsidy season (σ_{Dur} coefficients). However, in subsequent seasons (as represented by σ_{Aft} coefficients) households who have above-median connections to treated households saw increases in fertilizer use, improved seed use, maize yields, and beliefs about the returns to the technology package. Impacts on all these outcomes in the after period are statistically significantly different from zero at conventional levels.

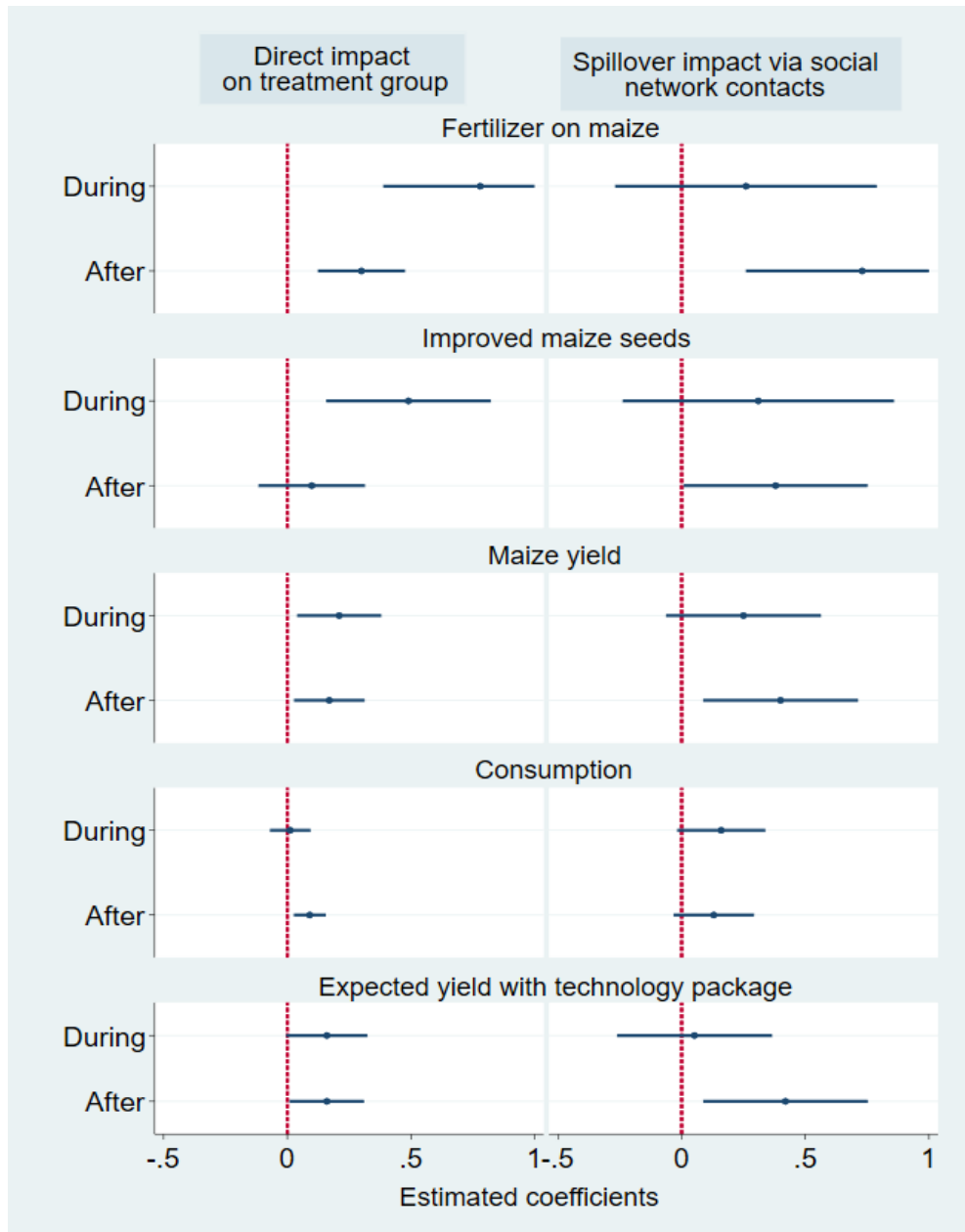


FIGURE 1—DIRECT AND SPILLOVER IMPACTS OF SUBSIDIES FOR GREEN REVOLUTION TECHNOLOGY

Notes: Results from estimation of equation (1). Dependent variable x expressed as $\log(1+x)$ for fertilizer and improved seed outcomes (originally in kilograms, which includes zeros), and $\log(x)$ for other outcomes (data include no zeros). Maize yield originally expressed in kilograms per hectare. Daily consumption per capita originally expressed in Mozambican meticaais. Expected yield with the technology is respondent's estimate of maize output (in kilograms per hectare) on household's main farming plot if using the subsidized Green Revolution technology package. Regression coefficients presented in left-hand column are α_{Dur} , α_{Aft} ; those in right-hand column are σ_{Dur} , and σ_{Aft} . Lines represent 95% confidence intervals. Regression coefficients are presented in Appendix Table A4.

5.2 Magnitudes of Effects

Our estimated intent-to-treat (ITT) impacts are large and economically consequential. Given 28.8% compliance with treatment assignment, our ITT estimate of a 0.21 increase in log yields (a 19% yield increase) for subsidy-recipient households implies an undiluted treatment-on-the-treated (TOT) impact of a 66% yield gain. Recent efforts to estimate the yield gap for eastern and southern African maize farmers find the gap to be between 2.5 and 3.5 tons/hectare (Sadras *et al.*, 2015).¹⁰ Hence farmers in our study area, who produced a bit less than 1 ton per hectare before the program, could have tripled their yields if they closed the gap. From this perspective, our results are well within the bounds of what is believed to be technologically feasible.

We find that in the post-subsidy period, network effects often exceed the direct impact of the subsidy. Since farmers with above median number of neighbors treated have on average 3.34 more farmers in their network who are treated, it is possible that the spillover impact exceeds the direct impact if farmers learn sufficiently from discussing agriculture with their neighbors. Even when summing the estimates of the direct and spillover impacts, we find a yield increment of 143%, which is high, but remains within the possible effects of closing the gap, estimated by Sadras *et al.* (2015).¹¹ Section 6.1 shows that the yield response to input use is very much within expectations from an agronomic perspective.

Finally, given that about 60% of household income comes from maize, the estimated consumption impacts are also in line with what is possible and what would be expected from Asian experience with Green Revolution technologies (Otsuka and Larson, 2016).

¹⁰ The yield gap is the difference between the yields farmers obtain and what is technologically possible using improved seeds and fertilizers given farmers' soils and the weather conditions they face.

¹¹ An increase of 0.51 in the log is equivalent to a 41% increase as the ITT, and a 143% increase in the average treatment effect.

5.3 Alternate specification of network effects

While the presence of social network effects is consistent with what we might expect if subsidies help resolve underlying information failures, this section provides additional analyses of the role of the social network. It explores alternative specifications and tests for mechanisms beyond information spillovers that might drive the findings.

5.3.1 Effects of number of subsidy beneficiaries in social network

Spillover effects are specified in our main regressions (Figure 1 and Table A4) as the effects of having above-median (two or more) social network members in the treatment group. This provides a reasonable approximation of the spillover effects observed when estimating a more flexible specification. Table 2 estimates spillover effects in such a specification, with five separate indicators for the number of one's social network contacts in the treatment group (indicators for one, two, three, four, and "five or more").

The general pattern in Table 2 is that the estimated coefficients on the social network variables tend to be positive and significant in the "after" period, but mostly not in the "during" period. In the after period, coefficient magnitudes rise as one moves from one social network contact to two social network contacts in the treatment group, with the effect remaining roughly stable thereafter. These patterns roughly approximate a step-function at two or more social network contacts in the treatment group. Figure 2 displays the spillover effect coefficients for the after period (σ_{Aft}), using the same flexible specification, for fertilizer use and maize yields.

TABLE 2—REGRESSIONS WITH MORE FLEXIBLE SPECIFICATIONS OF SPILLOVER EFFECT

		Fertilizer on maize	Improved maize seeds	Maize yield	Daily consumption per capita	Expected yield with technology package
<i>Direct impacts on treatment group members</i>	During	0.78*** [0.21]	0.51*** [0.15]	0.20** [0.088]	0.019 [0.043]	0.15* [0.086]
	After	0.31*** [0.092]	0.11 [0.11]	0.17** [0.072]	0.096*** [0.033]	0.16** [0.079]
<i>Spillover impacts DURING subsidy period of having x contacts in treatment group</i>	1 contact	-0.53* [0.31]	-0.19 [0.29]	0.14 [0.13]	0.052 [0.080]	0.059 [0.17]
	2 contacts	-0.025 [0.39]	0.12 [0.35]	0.33 [0.22]	0.21** [0.096]	0.12 [0.20]
	3 contacts	-0.57 [0.61]	0.20 [0.45]	0.44 [0.27]	0.084 [0.13]	0.056 [0.22]
	4 contacts	-0.079 [0.71]	0.68 [0.43]	0.28 [0.23]	0.11 [0.17]	-0.25 [0.28]
	5 contacts	-0.057 [0.54]	1.01 [0.52]	0.030 [0.28]	0.29 [0.15]	-0.18 [0.29]
<i>Spillover impacts AFTER subsidy period of having x contacts in treatment group</i>	1 contact	0.33* [0.18]	-0.29 [0.26]	0.18 [0.15]	0.069 [0.058]	-0.055 [0.19]
	2 contacts	0.95*** [0.25]	0.14 [0.27]	0.53*** [0.20]	0.18* [0.092]	0.44** [0.22]
	3 contacts	0.98*** [0.32]	0.21 [0.34]	0.48** [0.20]	0.16* [0.085]	0.15 [0.22]
	4 contacts	0.94** [0.46]	0.53 [0.37]	0.60*** [0.23]	0.27** [0.13]	0.24 [0.24]
	5 contacts	1.17*** [0.35]	0.66* [0.35]	0.39 [0.24]	0.31*** [0.12]	0.11 [0.28]
Observations		1,428	1,404	1,346	1,393	1,273

Notes: Data are from 2011, 2012, and 2013 follow-up surveys. Dependent variables are as in Figure 1. Regressions are based on modified version of equation 1 in main text, but with five separate indicators for the number of one’s social network contacts in the treatment group (indicators for one, two, three, four, and “five or more”) instead of a single indicator for above median (two or more) social network contacts in the treatment group. Robust standard errors in brackets.

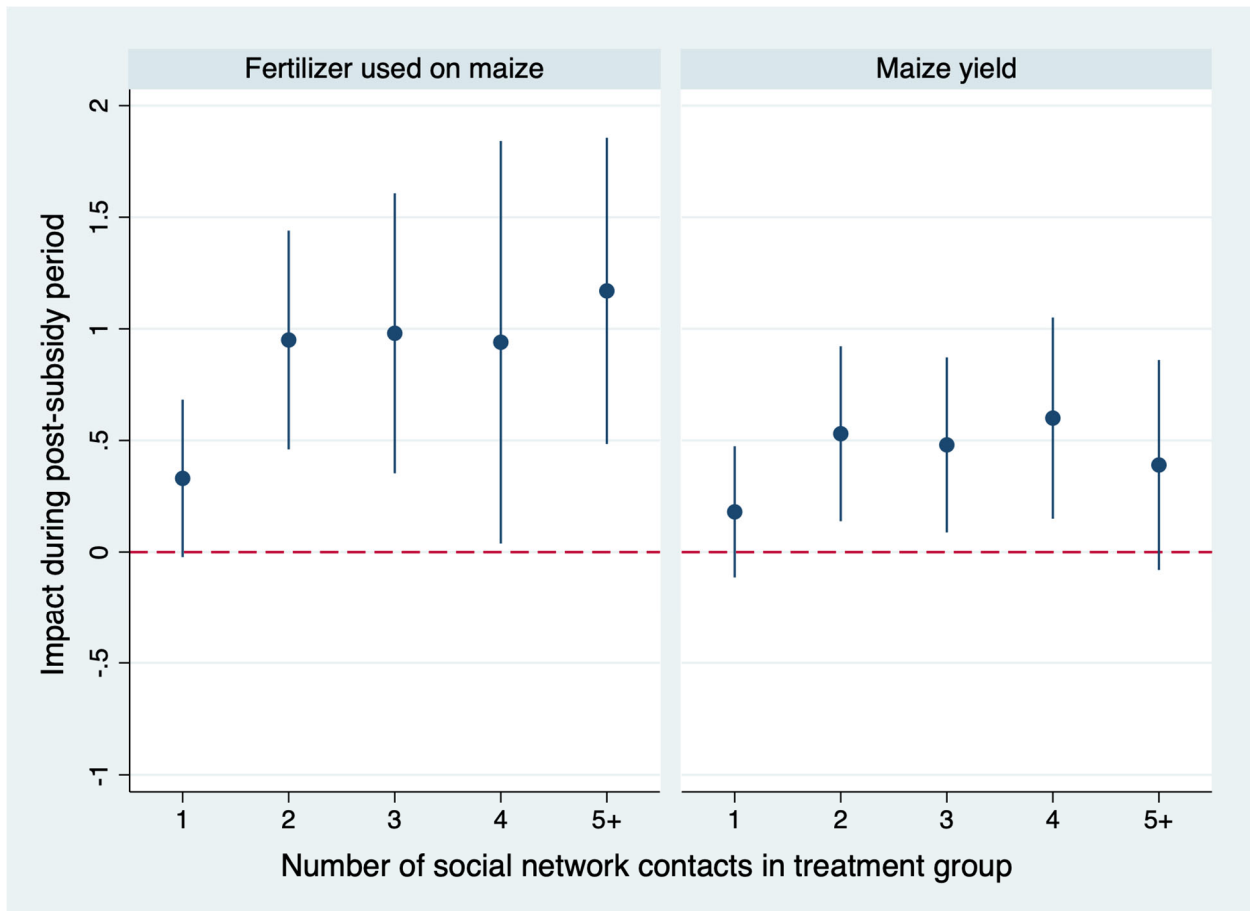


FIGURE 2—SPILLOVER EFFECTS BY NUMBER OF SOCIAL NETWORK CONTACTS IN TREATMENT GROUP

Notes: The specification is the same as Table 2. See Table 2 for regression coefficients and other details.

5.3.2 Separate social network effects for subsidy recipients and non-recipients

Spillovers are often thought of as impacts on those who did not receive the treatment themselves (the control group, subsidy non-recipients). But spillovers can affect treatment group members (subsidy recipients) as well (Baird *et al.*, 2014), and so the spillover effect coefficients in our analysis (σ_{Dur} and σ_{Aft}) incorporate spillovers to both treatment and control group members. In Table 3, we estimate these spillover effects to farmers in the control group and in the treatment group separately. We allow these spillover effects to differ by treatment group by

modifying equation (1) so that $Soc_{ic}Dur_t$ and $Soc_{ic}Aft_t$ are each interacted with the indicator for the treatment group ($Treat_{ic}$), and separately with an indicator for the control group ($Cont_{ic}$).

TABLE 3—DIRECT AND SPILLOVER EFFECTS OF INPUT SUBSIDIES, WITH SPILLOVER EFFECTS ESTIMATED SEPARATELY FOR TREATMENT AND CONTROL GROUP MEMBERS

		Fertilizer on maize	Improved maize seeds	Maize yield	Daily consumption per capita	Expected yield with technology package
<i>Direct Impacts</i>	During (σ_{Dur})	0.83*** [0.24]	0.37* [0.21]	0.11 [0.11]	0.0028 [0.057]	0.11 [0.10]
	After (σ_{Aft})	0.31* [0.16]	0.055 [0.15]	0.12 [0.082]	0.073* [0.039]	0.18** [0.086]
<i>Spillover Impacts on Control Group</i>	During (σ_{Dur}^{Cont})	0.33 [0.31]	0.16 [0.30]	0.12 [0.17]	0.14 [0.10]	-0.012 [0.18]
	After (σ_{Aft}^{Cont})	0.75*** [0.24]	0.33 [0.25]	0.33* [0.17]	0.11 [0.076]	0.44** [0.17]
<i>Spillover Impacts on Treatment Group</i>	During (σ_{Dur}^{Treat})	0.19 [0.35]	0.47 [0.35]	0.37** [0.17]	0.17 [0.11]	0.11 [0.17]
	After (σ_{Aft}^{Treat})	0.72** [0.31]	0.44* [0.24]	0.47*** [0.18]	0.16 [0.099]	0.40** [0.19]
Observations		1,428	1,404	1,346	1,393	1,273

Notes: Variable definitions are as in Figure 1. Regressions are a modified version of in equation 1 in main text, with $Soc_{ic}Dur_t$ and $Soc_{ic}Aft_t$ each interacted with the indicator for the treatment group ($Treat_{ic}$), and separately with an indicator for the control group ($Cont_{ic}$). Robust standard errors in brackets.

Spillover effects for treatment and control group members are quite similar, as it turns out, with some nuanced differences. The main difference is that for the treatment group only, spillovers lead to higher maize yield in the subsidized (“during”) period, not only in the post-

subsidy “after” period, suggesting that treatment group members may have helped each other learn to use the novel technologies more productively in the initial subsidized year.

5.4 Learning versus alternate mechanisms

Our results are consistent with the benefits of the input subsidy programs being driven, at least in part, by learning about the Green Revolution technologies. First, we directly observe that study participants report higher expected returns to the technology package when they are treated or have more than two treated social network contacts. Second, the increase in technology adoption, yield and consumption persists in periods after the end of the subsidy. Third, the greater effect on fertilizer than on improved seeds is consistent with the fact that fertilizer was less used and known than improved seeds prior to the program. Finally, the fact that the spillover effects mostly occur with a lag (only appearing in the “after” period) is reasonable, as farmers may wait to fully observe outcomes of neighbors’ experimentation before experimenting themselves (Foster and Rosenzweig, 1995). Altogether, these findings strongly suggest that the subsidy alleviates information imperfections related to subsidized technologies.

That said, social network spillovers are also consistent with mechanisms other than learning. A first possibility is that farmers simply kept some fertilizer for the following season, shared it with neighbors, or sold it to neighbors. We ask about fertilizer saving or sharing in our surveys, and find that it is quite rare: immediately following the subsidized 2010-11 season, the vast majority of respondents (88.8%) reported they had already used all the inputs for agriculture, 2.8% reported that they had not used it, and only 1.4% reported that they sold the inputs.¹² Even though it was an option, exactly zero farmers reported that they had given away

¹² Also, 1.4% declared that they used the inputs in some other way and 5.6% did not respond to this question

any of the inputs. Based on these data, there appears to be little scope for farmers to have shared their inputs with others. This is consistent with our qualitative observations during our presence in the field. Sharing may have also been limited by the fact that the vouchers could only be redeemed by the intended beneficiary named on the voucher certificate. In addition, we also estimate whether the likelihood of using the voucher for one's own agriculture was affected by the indicator for having two or more social network members in the treatment group. If sharing was happening, one would expect that having more neighbors treated should reduce one's own use of the voucher, but the effect is small in magnitude and not significant (Appendix Table A6).

Another possible channel that can generate the spillovers is resource transfers from treated farmers to their social network contacts (Maertens and Barrett, 2013). However, we find that the treatment, and social network connections to the treatment group, are not significantly related with the likelihood of providing assistance to others (Appendix Table A7 and Online Appendix G). Resource transfers are therefore unlikely to explain the large spillovers that we observe.

6. Cost-Effectiveness

How cost-effective was the subsidy program, and what fraction of the benefits occurs in subsequent (post-subsidy) periods and via spillovers? We calculate benefit-cost ratios of the subsidy program, in total and then separately for direct subsidy beneficiaries and their social network contacts. We also distinguish between the subsidized and post-subsidy periods. In this calculation, benefits are taken as the increase in maize output net of increases in the costs of

fertilizer and improved seeds, while costs include the cost of the subsidies to the government, as well as all logistical costs (calculated from detailed implementation budgets).¹³

The top panel of Table 5 displays the decomposition of benefits. Most studies, without a post-subsidy-period follow-up and a specific design to capture spillovers, would focus on estimating benefits accruing to direct subsidy beneficiaries during the subsidized period; we find that this only accounts for 10% of total benefits. But even when only accounting for this small minority of total benefits, the benefit-cost ratio would be 2.0. The remaining 90% of benefits accrues via spillovers from subsidized farmers to their social network contacts, as well as in post-subsidy periods. 69% of benefits occur through spillovers. 74% of benefits occur in the years after the subsidy ended. Accounting for both spillovers and post-subsidy effects leads to a roughly ten-fold increase in the benefit-cost ratio, from 2.0 to 20.5.

7. Conclusion

In sum, we find that temporary input subsidies can cost-effectively promote learning about Green Revolution technologies, adoption of those technologies, and improvements in agricultural output and living standards among both subsidy beneficiaries as well as their social network contacts. Viewed through the lens of economic theory, input subsidies address two kinds of market failures. First, they alleviate *imperfect information*, stimulating learning about the true productive returns to the technology among farmers who were previously underestimating those returns. Second, they mitigate the underprovision of goods that generate *positive externalities*. Subsidies induce experimentation with the technologies, and information spills over from subsidy beneficiaries to their social network contacts, who benefit from the

¹³ Online Appendix H details calculation of the benefits and costs.

information as well. When goods generate positive information or knowledge externalities, individuals have incentives to *free-ride*, avoiding costly experimentation so as to learn from others' experimentation instead. Subsidies induce some who would have engaged in free-riding to experiment themselves, moving society closer to socially optimal levels of experimentation. When information constraints are important, well-designed public policy that successfully encourages experimentation can generate the sort of spillover-driven highly favorable benefit cost ratio that we estimate for this program. In short, there is a strong economic case for temporary input subsidies, understood as a once-off inducement to experiment and learn.

TABLE 5—INPUT SUBSIDY PROGRAM BENEFIT-COST ESTIMATES

A. Shares of benefits			
	Subsidized year	Two years following the subsidy	All years
Direct effect	10%	21%	31%
Spillover effect (through social network contacts)	17%	53%	69%
Direct and indirect effects	26%	74%	100%
B. Benefit-Cost Ratios			
	Subsidized year	Two years following the subsidy	All years
Direct effect	2.0	4.3	6.3
Indirect effect (through social network contacts)	3.4	10.8	14.2
Direct and indirect effects	5.4	15.1	20.5

Notes: Benefits are increases in value of additional maize yields, minus costs of additional fertilizer and improved seeds used for maize. Direct effects accrue from being randomly assigned to treatment group (being eligible for subsidy voucher oneself). Indirect (spillover) benefits accrue from having above-median (two or more) social network contacts randomly assigned to treatment group. Costs include the value of input subsidies and subsidy program management and distribution costs.

As with all empirical work, subsequent studies should determine the generalizability of these results. It is important to note that prior investments in crop research led to the existence of improved seeds and fertilizers suitable for the conditions in our study area. In addition, a prior donor-funded effort led to availability of the technologies through a network of private agro-input dealers (Nagarajan 2015). Future research may find that effects of subsidies are attenuated in areas where available technologies are less suitable, or that are less accessible to input markets. Other dimensions along which subsidy effects may be heterogeneous include prior experience with modern inputs, geographical and climate conditions, crop types, and formal financial development. Policymakers should be cautious about expanding ISPs before future studies can measure direct impacts, post-subsidy persistence, and social network spillovers under different conditions, as guidance for locally-specific benefit-cost analyses.

Pending further studies to establish external validity, our findings have direct policy implications. In contexts with strong post-subsidy adoption persistence and social network learning spillovers, subsidy programs can achieve substantial gains even if scaled back, compared to current subsidy policies implemented by governments in Africa. Input subsidy programs need not be permanent nor universal to benefit farmers and their social networks in substantial ways. Temporary, targeted subsidies can make major progress in bringing the gains of new technologies to populations previously bypassed by the Green Revolution.

References

- Abate, G., T. Bernard, A. de Brauw, and N. Minot.** 2018. “The impact of the use of new technologies on farmers’ wheat yield in Ethiopia: evidence from an RCT,” *Ag. Econ.* 49:409-412.
- Baird, S., A. Bohren, C. McIntosh, and B. Ozler.** 2014. “Designing experiments to measure spillover effects,” *World Bank Policy Research Working Paper* 6824.
- Baird, S., Hicks, J.H., Kremer, M. and Miguel, E., 2016.** “Worms at work: Long-run impacts of a child health investment”, *The quarterly journal of economics*, 131(4), pp.1637-1680.
- Bandeira, O., and I. Rasul,** Social networks and technology adoption in northern Mozambique, *Econ. J.* 116(514):869-902 (2006).
- Beaman, L., A. BenYishay, J. Magruder, and A. Mobarak.** 2018. “Can network-theory-based targeting increase technology adoption?” *NBER Working Paper* 24912.
- Beaman, L., and A. Dillon.** 2018. “Diffusion of agricultural information within social networks: Evidence on gender inequalities from Mali,” *J. Dev. Econ.* 133:147-161.
- Beaman, L., D. Karlan, B. Thuysbaert, and C. Udry.** 2013. “Profitability of fertilizer: experimental evidence from female rice farmers in Mali,” *AEA P&P* 103(3):381-86.
- BenYishay, A., and A. Mobarak.** Forthcoming. “Social learning and incentives for experimentation and communication,” *Rev. Econ. Stud.*
- Bryan, G. S. Chowdhury, and A. Mobarak.** 2014. “Under-investment in a profitable technology: the case of seasonal migration in Bangladesh,” *Econometrica* 82(5):1671-1748.
- CIMMYT** 1988 “From Agronomic Data to Farmer Recommendations: An Economics Training Manual” Completely revised edition. Mexico, D.F.
- Conley, T. and C. Udry.** 2010. “Learning about a new technology: pineapple in Ghana,” *Am. Econ. Rev.* 100(1):35-69.
- Duflo, E., M. Kremer, and J. Robinson.** 2008. “How high are rates of return to fertilizer? Evidence from field experiments in Kenya,” *AEA P&P* 98(2):482-88.
- Duflo, E., M. Kremer, and J. Robinson.** 2011. “Nudging farmers to use fertilizer: theory and experimental evidence from Kenya,” *Am. Econ. Rev.* 101(6):2350-90.
- Dupas, P.** 2014. “Short-run subsidies and long-run adoption of new health products: Evidence from a field experiment.” *Econometrica* 82(1):197-228.
- Evenson, R. and D. Gollin.** 2003. “Assessing the Impact of the Green Revolution, 1960 to 2000,” *Science* 300(5620):758-762.
- Foster, A., and M. Rosenzweig.** 1995. “Learning by doing and learning from others: human capital and technical change in agriculture,” *J. Polit. Econ.* 103(6):1176-1209.
- Jayne, T., N. Mason, W. Burke, and J. Ariga.** 2018. “Taking stock of Africa’s second-generation agricultural input subsidy programs,” *Food Policy* 75:1-14.

- Karlan, D., R. Osei, I. Osei-Akoto, and C. Udry.** 2014. “Agricultural decisions after relaxing credit and risk constraints,” *Quar. J. Econ.* 129(2):597–652.
- Laajaj, R., and K. Macours.** 2016. “Learning-by-doing and learning-from-others” in *Learning for Adopting: Technology Adoption in Developing Country Agriculture*, A. de Janvry, E. Sadoulet, K. Macours, Eds. (Ferdinand), pp. 97-102.
- Maertens, A., and C. Barrett.** 2013. “Measuring social networks’ effects on agricultural technology adoption,” *Amer. J. Agr. Econ.* 95(2):353-359.
- Magnan, N., D. Spielman, T. Lybbert, and K. Gulati.** 2015. “Leveling with friends: social networks and Indian farmers’ demand for a technology with heterogeneous benefits,” *J. Dev. Econ.* 116:223-251 (2015).
- Morris, M., V. Kelly, R. Kopicki, and D. Byerlee.** 2007. *Fertilizer Use in African Agriculture* (World Bank Publications, 2007).
- Munshi, K.** 2004. “Social learning in a heterogeneous population: technology diffusion in the Indian green revolution,” *J. Dev. Econ.* 73(1):185–215.
- Muralidharan, K. and Niehaus, P.,** 2017. “Experimentation at scale”. *Journal of Economic Perspectives*, 31(4), pp.103-24.
- Nagarajan, L.** 2015. “Impact assessment of the effectiveness of agro-dealer development activities conducted by USAID-AIMS project in Mozambique” (International Fertilizer Development Center).
- Otsuka, K. and D. Larson** (eds.). 2016. *In Pursuit of an African Green Revolution: Views from Rice and Maize Farmers’ Fields* (Springer Press, New York).
- Ricker-Gilbert, J., and T. Jayne.** 2017. “Estimating the enduring effects of fertiliser subsidies on commercial fertiliser demand and maize production: panel data evidence from Malawi,” *J Ag Econ.* 68(1), 70-97.
- Sadras, V.O., K.G.G. Cassman, P. Grassini, A.J. Hall, W.G.M. Bastiaanssen, A.G. Laborte, A.E. Milne, G. Sileshi , and P. Steduto.** 2015. “Yield Gap Analysis of Field Crops: Methods and Case Studies” (FAO Water Reports No. 41, Rome, Italy.)

Online Appendix: Not for Print Publication

Subsidies and the African Green Revolution: Direct Effects and Social Network Spillovers of Randomized Input Subsidies in Mozambique

Michael Carter, Rachid Laajaj and Dean Yang

Appendix A. Background on Study Context and Locality Groupings

The study that is the subject of this paper is nested within a larger research program on the interaction between input subsidy programs and formal savings programs. Localities in Manica province were selected to be part of the larger-scale research program on the basis of inclusion in the provincial input subsidy program as well as access to a banking program run by Banco Oportunidade de Mocambique (BOM, the implementation partner for the savings component of the research program). To be accessible to the BOM savings program, a village had to be within a reasonable distance to one of BOM's branches (including places visited weekly by a truck-mounted bank branch). These restrictions led to the inclusion of 94 localities in the research program, across the districts of Barue, Manica, and Sussundenga. Each of the selected 94 localities was then randomly assigned to either a “no savings” condition or to one of two savings treatment conditions (“basic savings” and “matched savings”), each with 1/3 probability. As we analyze in other work, the addition of the savings intervention creates a complex set of interactions with the input subsidy intervention and we here focus on the pure cases of subsidy versus no subsidy in order to cleanly identify the impact of temporary input subsidies. The analysis of this paper thus focuses on the 32 localities randomly selected to be in the “no savings” condition, which did not experience any savings treatment.

The localities we use were defined by us for the purpose of this project, and do not completely coincide with official administrative areas. We sought to create “natural” groupings

of households that had a high level of connection to one another and limited connection to others. In most cases our localities are equivalent to villages, but in some cases we grouped adjacent villages together into one locality, or divided large villages into multiple localities.

Appendix B. Sampling and Survey Data

Our sample consists of households of individuals who were included in the Sep-Dec 2010 voucher randomization (both voucher winners and losers), and who we were able to locate and survey in April 2011. Key research design decisions could only be made once the government had reached certain points in its implementation of the 2010 voucher subsidy program. In particular, the government's creation of the list of potential study participants in the study localities (among whom the voucher randomization took place) did not occur until very close to the actual voucher randomization and distribution. It was therefore not feasible to conduct a baseline survey prior to the voucher randomization. Instead, we sought to locate individuals on the voucher randomization list (both winners and losers) some months later, in April 2011, and at that point requested their consent to participate in the study.

Individuals who consented to participate in the study were then surveyed; we refer to this as the April 2011 "interim survey". 704 individuals were included in the list for randomization of subsidy vouchers in 2010. Of these, 514 (73.0 %) were located, consented, and surveyed. One worry this approach raises is possible selection bias, if subsidy voucher treatment status affected the individual's likelihood of inclusion in the study sample. The first point to note is that we find no statistically significant difference in success rates in the April 2011 interim survey by subsidy treatment status: interim survey success rates for treatment and control group members were 71.0% and 75.0%, respectively (p-value of the difference is 0.246). A second key point is that, in

the vast majority of cases the reason an individual could not be surveyed in April 2011 was poor quality of the farmer lists prepared by extension workers at the time the vouchers were randomized. Extension agents had a relatively short period of time to make the lists of farmers to include in the voucher randomization, and hence sometimes used pre-existing lists of farmers, without carefully checking whether farmers were still present in the local area. The vast majority of individuals in the initial list of 704 who were not successfully surveyed in April 2011 had died or moved away prior to voucher randomization, or were otherwise not known to local guides who assisted in finding farmers for the April 2011 survey. Only 2.0% of the initial sample of 704 actually refused to be surveyed.

The final sample consists of 514 study participants and their households in the 32 study localities. The data used in the analyses come from household survey data we collected over the course of the study. Surveys of study participants were conducted in person at their homes. We fielded follow-up surveys in August 2011, August 2012, and July-August 2013. These follow-up surveys were timed to occur after the May-June annual harvest period, to capture input use, production, and other outcomes in the agricultural season leading up to that harvest. These surveys provide our data on key outcomes examined in this paper.

Appendix C. Tests for Experimental Balance and Differential Attrition

In this section we document balance on key time-invariant household characteristics with respect to the key randomly-generated independent variables in our analyses: the indicator for assignment to the treatment group ($Treat_{ic}$) and the indicator for having above-median social network contacts in the treatment group (Soc_{ic}). The balance test consists of estimating the following regression equation for household characteristic y_{ic} of household i in locality c :

$$(A1) \quad y_{ic} = \alpha Treat_{ic} + \sigma Soc_{ic} + \mathbf{X}_{ic}\gamma + \theta_c + \varepsilon_{ic}$$

This regression is a simplified version of equation (1) in the main text. The dependent variable is a time-invariant household characteristic, so there is only one observation per household; time subscripts t are therefore dispensed with. Now \mathbf{X}_{ic} are a set of dummies for the number of persons in the social network of household i .

As discussed above, due to uncertainties in the timing of voucher distribution and delays in the creation of the list of study participants, it was not feasible to conduct a baseline survey prior to the subsidy voucher lottery. Instead, we implemented a survey after the distribution of vouchers (in April 2011), which included questions on variables that are not expected to be manipulable in response to treatment. This balance test focuses on four key household characteristics that can plausibly be considered non-manipulable: education, gender, age, and literacy of household head. Education and age are measured in years. Gender is an indicator for the head being male, and literacy is an indicator for being literate.

As in equation (1), $Treat_{ic}$ indicates treatment group households, and Soc_{ic} indicates the household has above-median (two or more) social network contacts who were randomized into the treatment group. The vector of controls \mathbf{X}_{ic} includes indicators for having one, two, three, four, or five or more social network contacts who are study participants (omitted category zero). As discussed in the main text, social network size is not exogenously determined and so must be controlled for in the regression for Soc_{ic} to be considered exogenous. θ_c are locality fixed effects (treatment is randomized within locality). ε_{ict} is a mean-zero error term. We report robust (heteroscedasticity-consistent) standard errors.

Results are presented in Table A1. None of the coefficients on either $Treat_{ic}$ or Soc_{ic} are large nor are they statistically significantly different from zero. Conditional on the \mathbf{X}_{ic} ,

randomization of treatment status appears to have led to balance on key household characteristics with respect to the key right-hand-side variables of interest, $Treat_{ic}$ and Soc_{ic} .

TABLE A1—BALANCE RESULTING FROM TREATMENT ASSIGNMENT.

	Education of household head (years)	Household head male (%)	Household head age (years)	Household head lit (%)
$Treat_{ic}$	-0.059 [0.26]	0.015 [0.028]	0.26 [1.02]	-0.044 [0.035]
Soc_{ic}	-0.14 [0.58]	0.12 [0.072]	0.89 [2.54]	0.020 [0.085]
Observations	475	504	491	500

Notes: Level of observation is the household. Data are from April 2011 interim survey. Summary statistics of dependent variables: education of household head, mean 4.75 (standard deviation 3.19); 84.9% of household heads male; household head age, mean 46.14 (standard deviation 13.95); 77.8% of household heads literate. Regression is as in equation S1, including locality fixed effects and dummies for one, two, three, four, or “five or more” social network contacts who are study participants. Robust standard errors in brackets.

In addition to testing for baseline balance, it is important to consider attrition from the study sample, and whether such attrition could lead to biased treatment effect estimates. We attempted to survey everyone in the initial April 2011 sample at each subsequent survey round (in other words, attrition was not cumulative), so all attrition rates reported are vis-à-vis that initial sample. Across the three rounds, attrition rates range from 7.3% to 9.9%. In Table A2, we examine whether attrition is related to treatment assignment. The regressions are specified as in Equation 1 in the main text, with three observations per household (in 2011, 2012, and 2013). The dependent variable is an indicator for a household having attrited from a given round of the survey. None of the four coefficients are large in magnitude or statistically significantly different

from zero at conventional levels. Attrition bias is therefore not likely to be a concern in this study.

TABLE A2—TEST FOR ATTRITION RELATED TO TREATMENT

		Dependent variable: Attrition indicator
<i>Direct Impacts</i>	During (σ_{Dur})	-0.015 [0.028]
	After (σ_{Aft})	0.025 [0.023]
<i>Spillover Impacts</i>	During (σ_{Dur})	-0.013 [0.058]
	After (σ_{Aft})	0.013 [0.027]
Observations		1,524

Notes: Level of observation is the household. Data are from 2011, 2012, and 2013 follow-up surveys. Dependent variable is an indicator for a household being missing from the sample in a given round. Attrition rates in the three survey rounds are 8.6%, 9.9% and 7.3% in 2011, 2012, and 2013, respectively. Regression is as in equation (1) in main text. Robust standard errors in brackets.

Appendix D. Imperfect Compliance with Treatment Assignment

We have imperfect compliance with treatment assignment. First, only 40.8% of farmers in the treatment group redeemed and used their vouchers. Most such non-compliance stemmed from an inability to make the input package co-payment (even though claimed ability to pay was a selection criterion). Second, 12.4% of control group farmers reported using subsidy vouchers for the input package. Ground-level agricultural extension agents were instructed by their MinAg superiors to distribute vouchers to study participants in accordance with their randomly-determined treatment status. Control group receipt of vouchers was likely due to a mismatch in

incentives between extension agents and their MinAg superiors: extension agents had quotas of vouchers to distribute, and vouchers in study localities that were unused by treatment-group farmers were supposed to have been distributed in other, non-study localities. Some extension agents apparently chose not to bear the travel and effort costs of redistributing unused vouchers in other localities, instead distributing them to some control farmers in study localities. The difference in voucher use rates in the treatment and control groups (28.8%) is statistically significantly different from zero (Table A3). Our experiment therefore constitutes an “encouragement design”.

TABLE A3—COMPLIANCE WITH TREATMENT

	Redeemed voucher (indicator)
$Treat_{ic}$	0.288*** [0.0443]
Observations	511
Avg. in Control group	0.124
Robust standard errors in brackets	
*** p<0.01, ** p<0.05, * p<0.1	

Notes: Level of observation is the household. Data are from April 2011 interim survey. Dependent variable is an indicator for whether the household received and redeemed the voucher. Regression is as in equation A1 above. Robust standard errors in brackets.

Appendix E. Definitions of Outcome Variables

The outcome variables of interest are use of fertilizer for maize, use of improved maize seeds, maize yield, expected maize yield with the technology package, and per capita consumption in the household. We describe these variables in detail here, and then turn to the definitions of the social network contacts variables.

We focus on outcomes in log transformation, to deal with extreme values of outcome variables. The log transformation of maize yield, and expected yield with the technology, which contain no zeros, is straightforward. For other variables (fertilizer and seeds) that contain zeros we add one before taking the log. In the robustness checks described in Appendix F below we show that the results are qualitatively similar when using alternative measures such as the variables in levels, or simple dummies for nonzero fertilizer and improved seed use.

Fertilizer use, improved seeds use, and yields are obtained from a section of the survey that first asks the respondent to list all plots where maize is produced, before asking further questions plot by plot. The survey asked what quantity of planting fertilizer was used, and what quantity of top dressing fertilizer was used. The fertilizer used by the household on maize in that season is obtained by summing planting and top dressing fertilizer across all plots. In a similar way, we asked for each plot what quantity of seed was used, and what type of seed was used. Possible types of seeds are local, OPV (open pollinated variety) and hybrid. A list of all the common names of OPV and hybrid seeds were provided to help identify the type of seeds with the respondent. We summed the quantities of OPV and hybrid seeds across all plots to obtain the household's use of improved maize seeds during the season.

For each plot, we asked the respondent about the area cultivated. We also asked about harvested maize production. These two questions allow the use of multiple units, to allow the respondent to use the unit that he/she is most comfortable with. We then used conversion factors to convert all areas into hectares, and all production into kilograms. Then we summed the production across all plots, summed the area across all crops, and divided the total production by the total area to obtain the average maize yield of the household.

Expected yield with the technology package is calculated as follows. We take into account that farmers perceive production to be conditional on weather conditions. We therefore aimed to estimate the distribution of potential yield that farmers have in mind across possible weather realizations. We first have the respondent specify the main parcel cultivated by the household. We then ask what production the farmer would expect if he/she uses improved seeds and fertilizer in this parcel in 1) a normal year, 2) a very good year, and 3) a very bad year. We then asked the farmer to say, on average, out of 10 years, how many are very good years, how many are very bad years, and how many are normal years. We then multiply the expected production under each condition by the probability that this condition occurs according to the farmer's perception, to calculate expected production. We then divide the expected production by the area of the plot to obtain the expected yield with the technology package.

The calculation of consumption built on the following three survey sections:

- Food expenditures over the 7 days prior to the survey (37 items)
- Regular expenditures over the 30 days prior to the survey (13 items)
- Major expenditures over the 365 days prior to the survey (14 items)

Each type of expenditure was divided by the recall period to obtain expenditures per day. The sum of expenditures per day was divided by the adult equivalent size of the household to obtain our measure of consumption per day and adult equivalent household member. Adult equivalents were obtained using weights that are a function of age and gender of each household member.

Social network connections to treatment group members are collected and defined as follows. We have data on social network links prior to treatment, based on elicitation of "information links" (Conley and Udry, 2010). In the April 2011 interim survey, study

participants were presented with the full list of other study participants in the same village, and asked one by one whether they talked about agriculture with this person in the season prior to the survey (2009-10), and if so whether they did so “a bit”, “moderately”, or “a lot”. For each study participant, others whom they indicated as having talked to about agriculture “moderately” or “a lot” are considered the participant’s social network contacts. We do not require that the social network links be reciprocal. (In other words, Person A can be Person B’s social network contact, as reported by Person B, even if Person B is Person A’s social network contact, as reported by Person A.) Because we are interested in understanding spillovers of our randomized treatment within the social network, this elicitation only captures social network links among study participants in the village, not the full set of social network links (which would include study non-participants). Each respondent was asked about their links with 11.5 other study participants on average.

Appendix F. Results and Robustness tests using alternate dependent variable specifications

Table A4 presents the main results, represented graphically in Figure 1 of the main text. In our main regression results (Table A4 and Figure 1 in the main text), all outcome variables are expressed in natural logarithms, so that all coefficients could be interpreted in similar fashion, and to reduce the influence of outliers. In this section we present results to confirm robustness of our findings to alternate dependent variable specifications.

TABLE A4—DIRECT AND SPILLOVER EFFECTS OF INPUT SUBSIDIES.

		Fertilizer on maize	Improved maize seeds	Maize yield	Daily consumption per capita	Expected yield with technology package
<i>Direct Impacts</i>	During (α_{Dur})	0.78*** [0.20]	0.49*** [0.17]	0.012 [0.042]	0.21** [0.087]	0.16* [0.084]
	After (α_{Aft})	0.30*** [0.090]	0.099 [0.11]	0.091*** [0.033]	0.17** [0.073]	0.16** [0.077]
<i>Spillover Impacts</i>	During (σ_{Dur})	0.26 [0.27]	0.31 [0.28]	0.16* [0.091]	0.25 [0.16]	0.052 [0.16]
	After (σ_{Aft})	0.73*** [0.24]	0.38** [0.19]	0.13 [0.083]	0.40** [0.16]	0.42** [0.17]
Observations		1,428	1,404	1,393	1,346	1,273

Notes: Dependent variable x expressed as $\log(1+x)$ for fertilizer and improved seed outcomes (originally in kilograms, which includes zeros), and $\log(x)$ for other outcomes (data include no zeros). Maize yield originally expressed in kilograms per hectare. Daily consumption per capita originally expressed in Mozambican meticaais. Expected yield with the technology is respondent's estimate of maize output (in kilograms per hectare) on household's main farming plot if using the subsidized Green Revolution technology package. Regression specification and control variables are as in equation 1 in text. Heteroskedasticity-consistent (robust) standard errors in brackets.

In Table A5, we present coefficients from estimation of equation 1 of the main text for the same outcome variables, but with different specifications of the dependent variables. Fertilizer and seed outcomes are expressed as indicators for non-zero use and in kilograms. Maize yield and expected returns are expressed in kilograms per hectare. Daily consumption per capita is expressed in Mozambican meticaais (MZN). Because of the existence of large outliers, outcome variables expressed in kilograms or MZN are truncated at the 99th percentile of that variable's distribution in the full sample. Coefficient estimates are qualitatively similar in terms of relative magnitudes and statistical significance levels when compared to the results in Figure 1 in the main text and Table A4.

TABLE A5—REGRESSIONS WITH ALTERNATE SPECIFICATIONS OF DEPENDENT VARIABLES

		Fertilizer on maize		Improved maize seeds		Maize yield	Daily consumption per capita	Expected yield with technology package
		dummy	level (kg)	dummy	level (kg)	level (kg/ha)	level (MZN)	level (kg/ha)
<i>Direct Impacts</i>	During (σ_{Dur})	0.16*** [0.045]	17.0*** [5.38]	0.16*** [0.041]	4.16 [3.18]	177** [87.8]	1.61 [3.46]	181 [184]
	After (σ_{Aft})	0.066*** [0.021]	6.26* [3.37]	0.041 [0.031]	-0.16 [1.70]	210** [89.4]	10.2*** [3.35]	308 [196]
<i>Spillover Impacts</i>	During (σ_{Dur})	0.047 [0.062]	9.23 [6.44]	0.12 [0.081]	2.52 [5.47]	364 [153]	10.6 [9.35]	32.0 [273]
	After (σ_{Aft})	0.18*** [0.053]	18.9** [8.89]	0.11* [0.056]	5.02* [3.01]	679*** [247]	14.9* [8.03]	1,285** [630]
Nb of observations		1,428	1,428	1,404	1,404	1,346	1,393	1,393

Notes: Level of observation is the household. Data are from 2011, 2012, and 2013 follow-up surveys. Regressions are as in equation 1 in main text. Dependent variables are as in Tab. 2 in main text, but with alternate specifications. Fertilizer and seed outcomes are expressed as indicator variables (dummies) for non-zero use and in kilograms. Maize yield and expected returns are expressed in kilograms per hectare. Consumption is expressed in Mozambican meticaís (MZN). Outcome variables expressed in kilograms or MZN are truncated at the 99th percentile of that variable’s distribution in the full sample. Robust standard errors in brackets.

Appendix G. Alternate mechanisms for social network effects

We interpret the social network spillovers as reflecting informational spillovers from treatment group members to their social network contacts. It is important to consider whether other mechanisms may be behind these social network spillovers. The key alternate mechanisms are input sharing (treated farmers sharing fertilizer and improved seeds with their social network contacts) and resource transfers (treated farmers making monetary or goods transfers to their social network contacts). In this section we provide additional evidence and regression results testing whether these alternate channels are likely to be operative.

As explained in the main text, quantitative and qualitative evidence all point to very limited direct sharing of the inputs provided by the package. Among farmers who redeemed their voucher, 88.8%, reported they had already used the inputs for agriculture, 2.8% had not used it at the time of the survey, 1.4% sold the inputs, 1.4% declared that they used the inputs in some other way and 5.6% did not respond to this question. The requirement that vouchers be redeemed only by the original recipients themselves, with checking of names upon redemption, may have contributed to making sharing more difficult.

We can also examine whether a treated farmer's social connectedness with other treatment group members affects whether they used their subsidized inputs. If treated farmers shared their subsidized inputs with social network contacts, then having more social network contacts in the treatment group should reduce one's sharing, raising one's likelihood of reporting one had already used the inputs for the household's agriculture in the 2010-11 season. We run a regression analogous to the specification of equation A1 in Appendix C above (there is only one observation per household, for the 2010-11 season). The dependent variable is an indicator for having "already used" one's inputs for agriculture. The sample is also restricted to the treatment group only (because control group farmers were not asked this question), so the indicator for treatment drops out of the regression. In this regression, the coefficient of interest is the coefficient on Soc_{ic} . As shown in table A6, this coefficient is small in magnitude, and not statistically significantly different from zero. There is no indication that social connections to treatment group members affects whether one had already used the inputs for agriculture.

TABLE A6—SPILLOVER EFFECTS ON THE USE OF VOUCHER

VARIABLES	Already used voucher (dummy)
Soc_{ic}	0.035 [0.12]
Nb of observations	245

Notes: Level of observation is the household. Data are from April 2011 interim survey. Dependent variable is an indicator for whether household had already used the voucher in one of its plots at the time of the survey. Regression is as in equation A1 above. Robust standard errors in brackets.

The other possible channel of social network effects is resource sharing from treated farmers to their social network contacts, thus allowing them to alleviate their credit constraint and invest in the inputs. We examine this using survey data on each study participant’s sharing of resources with other households. Dependent variables are (1) an indicator for having given any money or other assistance to other households in the last 12 months (sample mean 0.39), and (2) the number of times one has given any money or other assistance to other households in the last 12 months (sample mean 1.65). These questions were asked only in the latter two follow-up surveys (2011-12 and 2012-13, the “after” period), so each household contributes up to two observations to the regression sample. We conduct regression analyses using the specification of equation A1 in Appendix Section C above (both observations are in the “after” periods so it is not relevant to estimate effects that vary over time).

Regression results are presented in columns 1 and 3 of Table A7. If resource sharing from treated households is a channel for the social network effect, we would expect positive coefficients on $Treat_{ic}$ (treatment group members becoming more likely to provide assistance to

others), and negative coefficients on Soc_{ic} (there is less need to provide assistance when others in the social network also receive the subsidy). None of these coefficients are large in magnitude or statistically significantly different from zero.

The resource sharing hypothesis is not perfectly tested in equation (A1), because the effect of Soc_{ic} on assistance provided to other households should be heterogeneous depending on one's own treatment status. If resource sharing from treated households is a channel for the social network effect, we would expect negative coefficients on Soc_{ic} only for households in the treatment group. We test this in columns 2 and 4 of Table A7, in which we add two interaction terms with Soc_{ic} , one interaction with a dummy for being in the treatment group and another interaction with a dummy for being in the control group. There is no large or statistically significant coefficient between the $Soc_{ic} * Treat_{ic}$ interaction term (nor on the interaction term with the control group indicator). We thus find no indication from these analyses that either input sharing or resource transfers constitute an alternate channel for the social network effects.

TABLE A7—IMPACTS ON ASSISTANCE PROVIDED TO OTHER HOUSEHOLDS

Variables	Indicator: any assistance given to other households		Number of times assisted other households	
Direct beneficiary	0.032 [0.032]	-0.0019 [0.042]	0.11 [0.11]	0.024 [0.12]
Spillover impacts	0.024 [0.062]		-0.064 [0.23]	
Spillover impacts on treatment group		-0.021 [0.066]		-0.18 [0.24]
Spillover impacts on control group		0.069 [0.074]		0.049 [0.27]
Observations	1,022	1,022	1,022	1,022

Notes: Level of observation is the household. Data are from 2012 and 2013 follow-up surveys. Dependent variables are indicator for providing money or other assistance to other households in the last 12 months (columns 1 and 2), and number of times providing such assistance in last 12 months (columns 3 and 4). Regressions in columns 1 and 3 use specification of Appendix Equation A1. Regression in columns 2 and 4 interact Soc_{ic} separately with indicators for households being in the treatment group and in the control group. Robust standard errors in brackets.

Appendix H. Details of the Calculation of the Benefit-Cost Ratio of the Program

Here we describe how we calculate benefit cost ratio (the ratio of the benefits to the costs) of the subsidy program. We first describe how we calculated the benefits, then how we calculated the costs. Both the benefits and costs are estimated in terms of intent-to-treat effects (ITT).

Benefits

We quantify benefits in terms of maize output, net of the increased costs of the use of the subsidized inputs (fertilizer and improved seeds). This is consistent with calculation of “value cost ratios” (VCRs), typical approach to quantifying farm input profitability (FAO 1975). This approach will somewhat overstate gains, to the extent that farmers induced by the subsidy to

increase their input use also bear costs to increase the use of other inputs (in particular, labor), for which we have no data.

We break up benefits into the following four categories:

A: the direct effect of the subsidy during the subsidy period

B: the direct effect of the subsidy after the subsidy period

C: the indirect effect of the subsidy during the subsidy period (indirect refers to the effect from having at least two social network contacts in the treatment group)

D: the indirect effect of the subsidy after the subsidy period

We estimate the value of increased maize production in each of the four categories of effects. To calculate these benefits, we take the following steps:

- We estimate impacts on maize yield, fertilizer use, and use of improved seeds by estimating regressions where dependent variables are expressed in levels (Tab. A5).
- We transform effects on yields into monetary values by multiplying the coefficients by the average area per household (3.35 ha) to obtain the increase in maize production for an average farmer, and then we multiply the result further by the median market price of one kg of maize (5 MZN, obtained from our survey data on maize sales) to obtain the value of the additional maize production.
- We calculate the value of the additional fertilizer and improved seeds used (retail market prices MZN 24 per kg of fertilizer, and MZN 12 per kg of improved seeds) and deduct this value from the additional value of maize production, to obtain net benefits. To be conservative in our benefit-cost ratios, we take into account any positive effect on input use, even if the coefficient in the regression in Tab. A5 is not statistically significantly different from zero.

- To calculate net benefits during the subsidy (“during”) year for direct beneficiaries (category A), we use the value of the farmer co-pay on the package as the cost to farmers, rather than using the total value of additional inputs used. We take into account that treatment group members were 28.4 percentage points more likely to use the voucher, so to estimate the ITT effect on the cost, we multiply the household co-pay by 0.284. The full (unsubsidized) cost of inputs is used when calculating net benefits due to spillover effects in the subsidized (“during”) period (category B), and for both direct and spillover effects in the “after” period (categories C and D).
- For spillover effects, we are interested in the ITT effect per subsidy voucher offered. To calculate these effects (categories C and D) we multiply the spillover coefficients of Tab. A5 by 0.78, which is the proportion of households whose number of social network contact with treatment group members is above the median (0.376), divided by the proportion of households in the treatment group (0.48).
- To account for time discounting, we take the present discounted value of benefits occurring in the “after” period, using a conservative (high) 10% annual discount rate.
- Estimated benefits in the various categories (including column and row totals) resulting from these calculations are in Table A8.

Costs

The cost of the voucher has two main components: the cost of the input subsidy, and logistical costs of identifying beneficiaries, coordinating the different actors, and distributing the vouchers. The latter includes all types of costs from staff and material to the overhead that pays for the corresponding share of the facilities of the implementing institutions. To estimate intent-to-treat costs, we consider the logistical costs for every household who won the voucher lottery,

whether the household used the voucher or not. The subsidy costs themselves are only considered to be disbursed if the household received and used the voucher. We also take into account the subsidy costs arising from the fact that some individuals in the control group managed to obtain and use the voucher.

To calculate costs we take the following steps:

- Logistical costs are calculated from detailed budgets provided by implementing organizations. We first consider costs incurred at the national level by Mozambique's Ministry of Agriculture (MinAg). These costs are for activities such as training and coordination of the program, supervision, and printing of the vouchers. We divide these national costs by 25,000, the number of beneficiaries of the overall national program, to obtain a per-voucher cost. We then consider costs of the provincial authorities in Manica province (including the extension agents who selected farmers and distributed the vouchers), where the research study occurred. These costs are then divided by 5,000, the total number of vouchers in Manica (only a subset of which were for our study participants). We add up national and provincial costs per unit to obtain the logistical cost per voucher intended to be distributed, MZN 694.
- The cost of the input subsidy per voucher used is MZN 2,300 (73% of the total input package value of MZN 3,163).
- To calculate the intent-to-treat costs, we calculate how much a person selected to receive the voucher costs compared to a person not selected to receive the voucher. This cost is equal to the logistics cost (which is spent whether or not the targeted person decides to use the voucher), plus 0.284 times the cost of the input subsidy. The number 0.284

corresponds to the difference in the probability of voucher use in the treatment group (0.408) and control group (0.124).

- Total cost per voucher intended to be distributed (intent-to-treat voucher cost) is MZN 1,347.

Benefit cost ratios

Finally the benefit cost ratios presented in Table 5 in the main text are obtained by dividing the benefits over the costs calculated above. We separately consider the four categories of benefits to illustrate how this ratio varies depending on whether or not post-subsidy period effect is included and whether or not the indirect effect is included.

TABLE A8—BENEFITS DUE TO INPUT SUBSIDY PROGRAM.

	During	After	Total
Direct Benefits	2,718	5,755	8,473
Indirect (Spillover) Benefits	4,578	14,542	19,120
Direct and Indirect Benefits	7,296	20,297	27,593

Notes: All benefits are impacts on maize output minus associated increases in fertilizer and improved seed costs, all denominated in Mozambican meticals (MZN; exchange rate MZN 27 per US\$ at time of study). Direct benefits accrue from being randomly assigned to treatment group (being eligible for subsidy voucher oneself). Indirect (spillover) benefits accrue from having above-median (two or more) social network contacts randomly assigned to treatment group. Benefits are in subsidized “during” period (2010-11 season) or post-subsidy “after” periods (2011-12 and 2012-13 seasons). Benefits in after periods discounted back to during period using 10% annual discount rate. See Appendix I (above) for details.

Appendix References

Food and Agriculture Organization, *Planning and Organization of Fertilizer Use Development in Africa* (FAO, Rome, Italy, 1975).