As Good as the Networks They Keep?: Expanding Farmer's Social Networks Using Randomized Encouragement in Rural Uganda

Kathryn Vasilaky

November 29, 2010

Abstract

This research isolates the impact of female social networks for subsistence farmers in rural Uganda for a re-emerging cash crop. We devised a social networking intervention (SNI), randomized at the village level, to tease out the pure effects of females' social networking on both females' and males' agricultural outcomes. Difference in difference estimates of the treatment effects show that the expansion of females' social networks significantly increases productivity for farmers producing at the average yield of production, and up to four times the average household's annual yield for cotton. The impact of the SNI exhibits diminishing returns for the highest yielding quantile of producers, suggesting that learning between farmers is most productive for low and mid yielding producers. The intervention has its strongest impact on females' production, but also spills over to males' yields, increasing overall welfare of the village. We also find that these effects are comparable to the effects of a conventional agricultural training program at a fraction of the cost. From a policy perspective, these findings are substantial. In many developing countries, women supply the majority of agricultural labor, exhibit substantially lower yields compared to their male counterparts; however, due to cultural norms, are rarely the recipients of training programs, particularly those that generate their own cash flow. A simple expansion of females' networks to promote new technologies is a not-yet utilized, but clearly effective, tool for economic development.

JEL: O, Q

This work was part of a larger randomized control trial entitled "Gender Dimension of Cotton Productivity in Uganda" led by Laoura Maratou (University of Maryland) and John Baffes (World Bank Prospects Group). Many thanks to Kenneth Leonard as an advisor, mentor, and enormous resource in developing the ideas in this paper; Markus Goldstein and the Economic Research Group at the World Bank, who funded this research; and participants at the Yale Development Seminar who provided very useful comments and suggestions.
Kathryn Vasilaky kvasilaky@arec.umd.edu: Agricultural and Resource Economics, 220B Symons Hall, University of Maryland 20742.

1 Introduction

In the last few decades, the focus of economic growth in developing countries has shifted from country-wide prescriptions to micro-development programs at the local level. Agricultural growth, in particular, is seen as the building block for alleviating hunger and poverty, as agriculture is the primary source of livelihood in the rural developing world. Programs aimed at increasing agricultural productivity are regarded as the most powerful means to reducing poverty as compared to nonagricultural programs (Gates, 2009). An essential stage in any program to increase productivity is the dissemination of new techniques and technologies by agricultural extension agents and trainers. This stage is frequently one of the weakest links in the process. One of the reasons for the lack of clear success in this effort is that trainers' success in reaching and affecting all individuals in a particular location relies on the effectiveness of social networks, which are often uknown to an outsider, and difficult to identify. While extension agents may bring new technologies with each program, what works best in practice in a remote village can widely differ from what is taught or what outside trainers can perceive as being important for local production. It is through individuals' personal ties that external information is disseminated within a remote area, localized, and that usable and believable knowledge is created. Thus, many welfare improving technologies are never adopted because individuals are not connected to effective social networks.

Understanding the impact of social networks on individuals' outcomes is, thus, central to development at the microeconomic level. Identifying these impacts, however, suffers from serious problems and it is difficult to prove that such impacts even exist and to what degree they impede or assist progress. There is no shortage of evidence that individuals with strong links to social networks, large social networks or almost any measure of social connectedness are more likely to adopt and experience better outcomes. However, social connectedness is endogenous and therefore we cannot isolate the impact of social networks on decision making for the reason that dynamic individuals belong to social networks. Unobservable characteristics of an individual, such as networking ability and sociability, which affect an individual's productive outcomes, are correlated with the type of network that an individual forms, confounding the impact of network effects, and biasing the estimated impact of social network measures.

This paper examines a research project that measured the impact of social networks for subsistence farmers in rural Uganda. To deal directly with the identification problem we used a social networking intervention (SNI), randomized at the village level, to tease out the pure effects of females' social networking on both females' and males' agricultural outcomes. The SNI exogenously increased the size of the average woman's social network in treatment villages and left existing networks intact in control villages. We show that the treatment increases productivity for farmers producing at the average yield of production, and up to four times the average household's annual yield for cotton. The intervention has its strongest impact on females' production, but also spills over to males' yields, increasing overall welfare of the village.

By using an intervention to exogenously increase the size of networks, we are avoiding many of the problems faced in the literature on social networks and are able to measure the value, on the margin, of adding to network size for the average female farmer. Thus, we avoid the type of network endogeneity that occurs when measures of the social network are defined using descriptive statistics of the networks' outcomes: the size of the network, the average age and work experience of the network, and the education level of the network. All of these characteristics of an individual's network reflect her ability to connect with such individuals, which would likely be correlated with her productive outcomes. Another common way to measure an individual's network is by summarizing the average outcomes of the individuals in the network: e.g. the number of individuals who decide to adopt a new technology, or the percentage of contacts who choose one input amount over another. These measures suffer from endogeneity issues known as the reflection problem. The reflection problem refers to the idea that an individual's outcome may seem to be affected by his or her network only because her network faces the same unobservable shocks or influences that simultaneously influence the individual, and not because the individual is in fact mimicking her network's actions (Manski, 1993). More complex graph-based measures of networks-including cohesion[1] or reach of the network reach[2]- lead to better understandings of social networks, but do not deal with the endogeneity problem.

This is one of two research studies, to our knowledge, on social networks in the development literature that

[1] Cohesion refers to the minimum number of nodes that would need to be removed to disconnect a group.
[2] Reach refers to the number of nodes within X number of steps from an individual.

uses a randomized encouragement design aimed at exogenously changing the social networks of women. Field et al. (2010) is another current study that encourages new microfinance groups to form in Bangladeshi villages, and varies the meeting frequency of these groups to study the impact of network effects on loan repayment. We are interested in determining whether social networks are a means to improve female's production of a relatively new crop, and to estimate social network (SN) effects without statistical bias. Randomization of a social network intervention (SNI) at the village level allows us to test both these hypotheses. By comparing outcomes of farmers assigned to the SNI to farmers in a control group, over time, we can estimate the impact of expanding a female's network. The estimated network effects will not be diluted by potential spillovers of the SNI, because individuals in the treated and control groups are in separate villages. Furthermore, the SNI was implemented in the presence of a randomly assigned cotton-training program, denoted as TR, which enables us to distinguish between the pure effect of social networks on productive outcomes, and the additive effect of social networks when coupled with a training program.

The decision to structure the SNI around females was inspired by an earlier study in rural Uganda on cotton producers that revealed male-heads' of households yields to be 3-4 times that of female-heads' of households yields (Baffes, 2008a). This is a tremendous welfare loss and reflects the general phenomena in developing countries of females operating far below their full potential, while males continue to receive training (Chambers, 1993). As females supply 70-80 percent of agricultural labor in rural Uganda and are responsible for 80 percent of food crop production (Tanzarn, 2005), this is also a tremendous loss to national welfare. Other studies have looked at possible reasons for these productivity differentials (Quisumbing, 2003; Udry, 1996). They have tested the impact of lower quality inputs, time constraints, disparate production functions, and property rights, where ownership of one's property seemed to be a significant explanation for gender differentials in productivity[3]. No study has yet looked at whether under-utilization of females' social networks could be behind this production schism.

Cotton production is particularly interesting to these purposes because it is being re-introduced in Uganda for the first time since the 1970's. Due to civil war and political unrest cotton production ceased under Idi Amin's regime when the majority of the Indians who managed Uganda's businesses were persecuted and expelled. As a result, at least one generation passed in which no transfer of knowledge ocurred for many of the cash producing crops. Udry's seminal work shows that it is precisely in these circumstances, where new technologies are nascent, that social networks should have their greatest impacts (Conley and Udry, 2010).

[3]Women are unable to allow their land to lie fallow for fear of losing control of their plot (Udry and Goldstein, 2006).

2 Social Networks and Technology Adoption

In development economics, there are two gropus of studies on social networks that focus on estimating the impact of social networks on technology adoption and “learning”, in terms of correct input use and resource allocation. The first group studies the effect of individuals' existing social capital (ego network) on the decision to adopt new technologies (Bandiera and Rasul, 2006; Isham, 2002; Maertens, 2010; Matuschke and Qaim, 2009; Young, 2009). The second group looks at the effect of individuals' social capital on input use, testing whether learning occurs inside the network (Conley and Udry, 2010; Darr and Pretzsch, 2008; Goldstein and Udry, 1999; Kremer et al., 2007; Munshi, 2004).

The above literature employs different methodologies to deal with the endogeneity of social networks. The adoption-network literature attempts to identify network effects by controlling for a gamut of individual level characteristics that may confound individuals' network effects, in the hope that these variables will control for all unobservable characteristics of the individual. Matuschke and Qaim (2009) find that the endogenous group network measures such as the average number of adopters in an individual's network impact an individual's decision to adopt a new crop. Bandiera and Rasul (2006) also use endogenous network measures, i.e. the number of sunflower plant adopters in an individual's network, to predict an individual's decision to adopt. They too find that endogenous network measures significantly influence an individual's outcomes. Specifically, they find that the relationship between the probability that an individual adopts sunflower production and the number of adopters in that individual's network is inverse-U shaped. In other words, the probability of an individual's adopting sunflower production increases with the number of adopters in their networks at a decreasing rate, and eventually declines with the number of adopters. Unlike the latter two studies, Isham (2002) identifies the effects of networks using the exogenous variation in individual's networks caused by ethnic fractionalization and land inequality. He finds that social capital, when instrumented for by tribal affiliations, has significant impacts on adoption decisions. However, there are many other aspects of an individual's production network that are not captured by ethnic affiliation, which a researcher would want to identify.

The learning-network literature relies on dynamic decision making to capture network effects. Namely, the individual only makes decisions after observing the actions of his network's members. If all actions and decisions are captured sequentially, and we believe that the actions of one individual are caused by observing the outcomes of others', then information x at time t for person j, should identify the decision or outcomes of individual i at time $t+1$. With detailed data on the outcomes and order of outcomes for all individuals in a network, this literature argues that the reflection problem is bypassed. Conley and Udry (2010) and Goldstein and Udry (1999) rely on the dynamic decision making assumption to identify learning from one's network, as well as detailed information on geography, soil, credit and family relationships that should control for confounding productivity factors. They find strong evidence of social learning, where farmers' decisions on inputs are affected by the successful outcomes of their neighbors in previous periods. Munshi (2004) adds to this result by showing that learning occurs in more homogenous populations. Maertens (2010) also uses a similar dynamic decision making methodology as Conley and Udry (2010), but for predicting adoption decisions rather than inputs or outputs. Her research goes further and looks at distinguishing the channels by which individuals decide to adopt: social learning, imitation and social pressures, which are similarly outlined in Young (2009). Leonard (2007) looks at the decision to visit a health care facility with a new clinicians as similar to the decision to adopt a new technology and uses a methodology similar to that of Conely and Udry (2010) by assuming individuals can only learn from the experience of people who visited the facility before they became sick. These methodologies rely on meticulous data collection and the belief that the available observable control variables, such as soil characteristics, are sufficient for dealing with confounding unobservable variables, such as weather and other productivity shocks, which concurrently affect the individual and their network.

Duflo et al. (2006) is one study that uses an experimental design to identify social networks' effects. In Duflo et. al, 2006, farmers are randomly selected from among the parents of children on school lists to participate in fertilizer-use trials. They compare the average outcomes of those individuals who were reported speaking to selected farmers with the average outcomes of individuals who were reported speaking to the control group. Essentially, they are exogenously altering the information present in some randomly selected social networks. According to their randomization, their identification strategy relies on the fact that there are no significant spill-over effects of the information from the networks of trained individuals to the networks of untrained individuals. Namely, they state, "farmers participating in each pilot were randomly selected from the parents of a school list, and that participating in the trials is randomly assigned within a school. Parents from the same schools that were not selected form a control group"((Duflo et al., 2006), pg 7). However, when interventions are likely to have significant externalities, randomization across individuals will not capture the full effect of a program. That is, if the networks of trained and untrained are in close enough proximity to each other, it is very likely that individuals who spoke to trained farmers could have then shared the information with individuals in the network of untrained farmers. As a result, the differences in average outcomes of untrained and trained networks will not be detectable, which is what the authors find, when in fact they may exist.

Our research does not rely on controlling for unobservable household variables or the dynamic learning assumption. And unlike Duflo et al. (2006), our experimental design tests for the actual impact of social networks, whereas the Duflo et al. (2006) experimental design attempts to estimate the effect of a training program at diffusing information across already existing networks, but not the impact of social networks themselves. Similar to Field et al. (2010), we directly perturb the networks of our sample population by randomly pairing individuals within selected villages, an event we term from hereon the social network intervention (SNI). New pairs are encouraged to discuss their problems and solutions in growing cotton, create a mutual long term goal for increasing cotton output, and re-exchange information about growing cotton that they received in focus meetings. The SNI is meant to encourage information flow across new links. In this way, we would like to measure the actual impact of adding a new link to a grower's network.

The next section motivates the sample population selected for this study. Section 3 explains the randomization. Section 4 outlines a simple model to motivate our empirical estimation in Section 5. Section 6 presents the results. Section 7 concludes.

3 Women and Cotton in Uganda

We follow Baffes (2008a) and use female heads of households as our sample population. This avoids revisiting the issue of land ownership as a potential cause for gender specific productivity differentials. We expect that the expansion of social networks for production, particularly for a new cash crop, has a high potential for improving females' outcomes. The reason behind these expected gains is due to females' networks traditionally being less oriented toward production alone than males' (Edmeades et al., 2008; Katungi et al., 2006). This may be because females face a starker tradeoff between economic and non-economic social networks. While males' days are delineated by morning work and afternoon discussion with other males, women's days are often a simultaneous combination of work, child-care, and household responsibilities. A wider range of household responsibilities raises the cost and reduces the availability of acquiring new production techniques (Granovetter, 2005). Responsibilities close to the home also restrict females from participating in geographically dispersed social networks and community projects, and force their relations to be dependent on the collaborative tasks that they perform with other females, i.e. collecting water, fuel, and harvesting crops

(Maluccio et al., 2003). Female-headed households are also more likely to be poorer or more marginalized in their community, particularly those who have been widowed or divorced[4]. Hoang et al. (2006) emphasizes that "development workers' inadequate understanding of local social networks, norms, and power relations may further the interests of better-off farmers and marginalize the poor," who are disproportionately female. Large "structural holes" in females' production networks, therefore, likely exist, and establishing new links with a like grower should create a more complete production network for every farmer in a village by closing some of these gaps. Nascent and weaker links are also more likely to propagate new and novel information along their paths, and their introduction can potentially have the greatest impact (Granovetter, 1974).

Our full sample population is comprised of male and female-headed households that grew cotton in 2008 in rural Uganda. The SNI was directed at female-headed households[5], while cotton training was administered to both groups. Randomization was stratified by female and male-headed household status after having randomly selected cotton growing villages from the complete list of all cotton growing villages in one Eastern district (Bukedea District) and one Northern district (Oyam District) of Uganda. The SNI consisted of an in-depth survey of the grower's social networks, participation in information games[6], in which participants learned some of the information that would later be taught to them if randomly selected for the training treatment, and being paired with a "buddy" in his or her village area (who may or may not also be receiving training) with whom they were encouraged to develop an agricultural link[7].

The pairing occurred by first stratifying the cotton growing participants into 2 to 3 geographic areas of the village[8], and then randomly pairing individuals. We used a random number generator to print out lists of numbers randomly drawn from a uniform distribution, $U[0, x]$, where x represents the number of individuals in the group. For example, if the group was comprised of 14 women, then $x = 14$. We would then pair individual "1" with the first listed number on the list of numbers drawn, with replacement, from U[0,14]. If the first number was "1" then we'd select the next number in the list, perhaps "3". Now "1" and "3" would be paired, "3" would be crossed out, and we would continue down the list in this way until all 14 women were paired. The pairing occurred among all the female-headed households in our sample as well as the additional female

[4]However, divorced, separated or widowed females who are not subsumed back into a male-headed households, but retain some own property rights, are not necessarily the most resource constrained individuals in rural society.

[5]The head of household was defined as the individual who made land, resources and income allocation decisions in the household.

[6]The information games are detailed in (Vasilaky, 2010)

[7]See end extension training sheets in the appendix, from which 10-12 points were taught via the games

[8]This to ensure that females were not separated by large geographic constraints.

cotton growers in the village who participated in the information games. A random re-pairing occurred if the individuals were already neighbors, or if both were to receive training[9] to maximize the effects of networking.

3.1 Randomization

In order to capture the effect of a social network intervention, randomization occurred at the village level as we would expect externalities from both programs, SNI and TR, between the treated and untreated within a village. By randomizing the SNI and TR programs across villages, we are able to measure the effect of the SNI treatment, the TR treatment, and the complimentary effect of both treatments. The following diagram shows our three treatment groups that we will compare to the control groups who received no treatments. The Treatments table below represents the combinations of effects between the two treatments:

Treatments, Sample Size

	TR	No TR	Totals
SNI	96	59	155
No SNI	120	50	170
Totals	216	109	325

The first round of a large-scale household survey was administered to 36 villages in 4 regions of Uganda: North (13 villages), Northeast (13 villages), West (5 villages) and West-Nile (5 villages)[10] from February through May 2009. The household survey consisted of questions on household demographics, input use and outputs for cotton and other crops grown, household control of financial assets including sales from cotton, and a separate survey instrument on farmers' social networks regarding adoption, cultivation and marketing of cotton.

To facilitate farmers learning proper cotton growing techniques, and to estimate the impact of a low-budget agricultural training program, villages were randomly selected for participation in the TR. Nine of the 13

[9] This only occurred in villages that were selected for the TR.

[10] This results in a survey of 500 households. Approximately 175 households in each Northern region and 75 households in each Western region were randomly selected for the survey.

villages in the northern and eastern regions received training for a total of 18 trained villages over the course of the 2009 growing season. In each village, approximately 14 heads of households were randomly selected to be visited by a local agronomist three times a week to undergo five training stages in 2009[11]: pre-planting in March through April; planting in May; pesticides use in July through August; harvesting in October through November; and marketing in December and January. Half of the participant sample is female heads of households. Among the 18 villages randomly selected for agricultural training, another subset of villages was chosen to participate in the SNI. Among the 8 villages not selected for agricultural training, 4 received the SNI and 4 did not.

In the SNI group, each pair received a polaroid of themselves and their team member, chose a team name, identified some cultivation issues and chose a collaborative goal, as well as potential times when they would meet to exchange information. They then presented this to their peers at a group meeting. In this way they were strongly encouraged to build a relationship around what they would learn in the coming year about growing cotton via their new link.

4 Model

We use a conceptual model that is limited to the household's decision in choosing inputs to produce cotton, given their access to new links and training, which are exogenous in the model and in our data due to the design of SNI. We look at the household's maximization of yields. Although the household's objective is to maximize their profits, the TR and SNI were designed to affect output; therefore we don't consider the evolution of cotton prices and households'expectation of those prices, which could negatively affect overall profits, even if yields increased. It should be noted, however, that prices for cotton and inputs are known in advance with some noise[12], and because decision makers are subsistence farmers, they are generally working without wages.

[11] This was part of the larger RCT which implemented a cotton training program under "Gender Dimension of Cotton Productivity in Uganda" led by Laoura Maratou (University of Maryland) and John Baffes (World Bank).

[12] Cotton prices are set by the government's announcement of an indicative price. Although the price is not fixed, it is highly suggestive of what price ginners will pay for cotton at harvest time, and has consistently been a conservative projection of the expected price (Baffes, 2008a).

Household i chooses a vector of inputs, $\bar{x}$, to maximize a production function at time period t:

$$F_{it}(\bar{x}_{it}|SN_{it}, KN_{it}) = b * \bar{x}_{it}^{\eta} SN_{it}^{\gamma} KN_{it}^{\delta} \tag{1}$$

subject to a budget constraint $\bar{p}'\bar{x}_{it} \leq I_{it}$ where $t = 0....T$, SN is a continuous variable representing one aspect of the i's social network and is affected by the exogenous variation from SNI: $SN = f(SNI)$, and KN, knowledge, is affected by the exogenous variation from the TR program $KN = g(TR)$. The b, η, γ, and δ are unknown parameters (for ease of notation we suppress the i subscript unless necessary). We choose to model the problem statically, as the decision to grow cotton is not a dynamic one in terms of inputs, i.e. cottonseed cannot be carried over from one season to the next. Social networks would generally be modeled to evolve over time, and could be endogenized in the model; however, their evolution, particularly for females, is likely determined outside the realm of cotton production networks, and because of our SNI experiment is exogenous in this particular problem.

The sign and magnituted of η, γ, and δ is representative of the returns to output from any one of these inputs, which is an empirical question to be answered with the data. SN can be thought of as the sum of weighted links: $SN_{it} = \Sigma_{i \neq j} \delta^{n_{ijt}} sn_{ijt}$, as in Jackson and Wolinsky (Jackson and Wolinsky), where n_{ij} is the number of links for the shortest path between i and j ($n_{ij} = \infty$ if there is no path between i and j), sn_{ij} is the value of one link between i and j, and $0 < \delta < 1$ indicates that the value of a link is proportional to the distance between i and j.

The optimal, non-corner solution, will yield the function $x_t^*(SN_t, KN_t, I_t, \bar{p})$, and the optimized production function[13],

$$F^*(\bar{x}_t{}^*(SN_t, KN_t, I_t, \bar{p})|SN_{it}, KN_{it}) = b\bar{x}_t^*(SN_t, KN_t)^{\eta} SN_t^{\gamma} TR_t^{\delta} \tag{2}$$

If $F(.) = -e^{-(.)}$[14], and substituting in SNI and TR for SN and KN respectively, then taking logs gives us an estimatable function[15]:

[13]For now we exclude income, I, and prices $\bar{p}$, from our optimal solution, since our focus is on the effects of SNI, and SNI relative Tr, on individuals' outcomes. Including them as controls would reduce some of the variance in the error term, but estimates will remain unbiased based on our identification strategy.

[14]This does not impose any strict assumptions on the utility function when F(.) is exponential. For instance the measure of Absolute Risk Aversion with respect to x is not constant as it would be with only one input, is $\frac{\delta U^2/\delta X}{\delta U/\delta X} = \eta\gamma\text{SN}\delta\text{KN}$. Similarly, the relative risk aversion is $\eta X \gamma \text{SN} \delta \text{KN}$

[15]Taking the log of yields will also be useful empirically, as a number yields are close to zero, and a log transformation re-weights the distribution towards the lower tail.

$$\log y_t^*(SN_t, TR_t) = \beta + \eta x_t^*(SN_t, TR_t) + \gamma SN_t + \delta TR_t \tag{3}$$

We are interested in the difference in outcomes between the control group vs. the treated groups as a result of a change in SNI, where SNI and TR equal one if an individual received a new link or training, and zero otherwise. This is captured by the above equation in first differences for those who did and did not receive the SNI, controlling for the TR treatment:

$$\log y_t - \log y_{t-1} = \gamma(SN_t - SN_{t-1}) + \eta(X_t - X_{t-1}) + \delta(TR_t - TR_{t-1}) \tag{4}$$

This can also be written using a dummy variable for time[16]:

$$\log y_t = \alpha + \beta t + \rho SNI_t + \mu TR_t + \nu SNIxTR_t + \eta SNxTRxt + \gamma SN_t xt + \delta TR_t xt \tag{5}$$

As the model cannot capture all determinants of yields, we observe y with some error u, such that Equation 2 becomes:

$$\log y_t = \alpha + \beta t + \rho SNI_t + \mu TR_t + \nu SNIxTR + \eta SNxTRxt + \gamma SNIxt + \delta TRxt + u_t \tag{6}$$

Using our data on outcomes and treatments we can estimate the above equation. Assuming that the u_{it} are iid distributed disturbances with some known distribution that are uncorrelated with the regressors, or $E[SNI_t u_t | Z_t] = 0$ where $Z_t = [SNI_t, TR_t, SNIxTR_t, SNxTRxt, TR_t xt]$, the estimated effect of the SNI, $\hat{\gamma}$, will be unbiased.

The estimation of η in Equation 3 is equivalent to a triple difference across both treatments and time, and γ captures the double difference across time and SNI. The estimated $\hat{\gamma}$ captures the average treatment effect (ATE) of the SNI that is: $\gamma = [E(y|SNI = 1, t = 1, TR = 0) - E(y|SNI = 1, t = 0, TR = 0)] - [E(y|SNI = 0, t = 1, TR = 0) - E(y|SNI = 0, t = 0, TR = 0)]$, and $\hat{\delta}$ captures the simultaneous average effect of SNI and

[16] Where holding TR constant at zero, we can see that the two specifications yield equivalent results:$[(y|SNI = 1, t = 1, TR = 0) - (y|SNI = 1, t = 0, TR = 0)] - [(y|SNI = 0, t = 1, TR = 0) - (y|SNI = 0, t = 0, TR = 0)] = [(\alpha + \rho) - (\alpha + \rho + \beta + \gamma)] - [\alpha - (\alpha + \beta)] = \gamma$, just as it would in the first difference equation if TR is held constat at 0.

TR[17], on yields. The ATE is equivalent to $E[y|SNI = 1, t = 1, TR = 0] - E[y|SNI = 0, t = 1, TR = 0]$, or the average treatment effect on the treated, where $t = 1$,and TR is held constant at zero, if we believe that there would have been no difference in yields between our treatment and control groups in the absence of the SNI and TR, i.e. $[E(y|SNI = 0, t = 1, TR = 0) = E(y|SNI = 0, t = 0, TR = 0)]$. This is a fair assumption to make, given that our program was randomly assigned. However, because we were fortunate enough to follow our control and treatment groups over time, we can control for such trends, where β in Equation 3 captures $E[y|SNI = 0, t = 0, TR = 0] - E[y|SNI = 0, t = 1, TR = 0] = \alpha - (\alpha + \beta)$, or the differential trend in yields over time in the absence of the interventions.

Because the SNI is an encouragement design, our estimates reveal the intent to treat (ITT), or the intent to change individuals' networks. That is, everyone who participates in the SNI meeting is regarded as having participated in the SNI, even if they did not follow any of our suggestions over the course of the year. Therefore, our estimates are only a lower bound for the possible effect of SNI[18], further strengthening our results.

The above outline frames a number of testable hypotheses:

(1) $\frac{\delta y_t}{\delta SN_t} = \frac{\gamma}{SN_t} F^*() > 0$, or the marginal impact of social networks is positive.

(2) $\frac{\delta^2 y_t}{\delta^2 SN_t} = \gamma(\gamma - 1)\frac{\delta^2 F^*()}{\delta^2 SNI} \leq 0$, or decreasing returns to scale in SNI.

(3) $\frac{\delta y_t}{\delta sn_t} = \gamma \delta^{n_{ij}}$ implies that the value of an additional link to person j is decreasing with the distance from j.

5 Empirical Estimation

5.1 Output

A summary of the data are shown in Table 1. The data indicate that the interventions were evenly allocated across control and treatment groups, with slightly under half the total number of villages receiving the SNI, and slightly over half receiving TR. The average Ugandan cotton farmer in our sample produces between 100

[17] $\delta = [E(y|SNI = 1, t = 1, TR = 0) - E(y|SNI = 1, t = 0, TR = 0)] - [E(y|SNI = 0, t = 1, TR = 0) - E(y|SNI = 0, t = 0, TR = 0)] - [E(y|SNI = 1, t = 1, TR = 1) - E(y|SNI = 1, t = 0, TR = 1)] - [E(y|SNI = 0, t = 1, TR = 1) - E(y|SNI = 0, t = 0, TR = 1)]$

[18] Provided SNI>0.

and 200 kilograms per year. This concurs with previous studies on cotton production in Uganda, which find that the average subsistence farmer produces about 100 kilograms of cotton lint per annum, while an average US cotton farm yields about 500 kilograms per acre Baffes (2008a). To situate this in tangible terms, one third of a kilogram of seed cotton[19] amounts to approximately one half to one full cotton t-shirt, where one cotton t-shirt requires 0.30 kilograms of seed cotton, or about a quarter of a kilogram of cotton lint, which equates to 90 US cents for the Ugandan farmer. Standard deviations for the yield of cotton (kilograms per acre) and level of cotton (total kilograms produced) are particularly high. This is due to the stark drop-off in production from 2009 to 2010, as well as to yields being right skewed, as seen in Figure 1. The average farmer produces less than 500 kilograms per season, which is well below the maximum producer in 2009 of 2,000 kilograms, resulting in a high variance in yields.

The number of acres used to grow cotton ranges between one half to one acre on average. Land is generally not a scarce resource, though having the means to clear and prepare the land is. Therefore measures of yield will reveal this constraint, while the total kilograms of cotton harvested will not. We also summarize yield per seed, denoted as "ypseed", since yield alone may not reflect accurate planting technique and input use. One farmer may have access to more seed and can replant in areas where no germination takes place, while another succeeds with the first round of seeds because of good technique. Yield per seed was 52 kilograms in 2009 and fell down to 37 kilograms in 2010. It should be noted, however, that seed is freely or nearly freely provided by cotton ginners, so that yield alone may be the most appropriate outcome measure. The drop in yields, acreage used for cotton and yield per seed, is the result of delayed rains in Northern and Eastern Uganda during the course of the intervention. Our interest remains in measuring the impact of the SNI in two ways. First, we measure the impact that the SNI had on increasing the probability that a household maintained cotton as a cash crop despite the drought. Second, we estimate the impact that the SNI had on output, and intermediate input decisions for farmers, while controlling for the impact of the TR intervention.

We first look at the impact of the SNI on farmers' decision to grow cotton in the presence of the training intervention, clustering all standard errors at the village level to account for within village correlations between households' error terms on outcomes. Table 2 estimates the effect of the SNI and TR on remaining a cotton grower between 2009 and 2010, despite the adverse weather shocks mentioned earlier. We use a Probit model to predict the probability that a grower continues to grow cotton. Column 1 indicates that the presence of the SNI in a village positively and significantly impacted a farmer's decision to continue to

[19]Seed cotton refers to the harvested cotton lint and seed, where the seeds have not been filtered from the lint. Cotton seed refers to the actual seeds that cotton produces.

grow cotton, where the outcome variable is zero if the individual ceased to grow cotton in 2010, and equals one if they planted cotton. The marginal effect of expanding a farmer's network by one link increases the probability of remaining a cotton grower by 18 percent. Whereas, introducing training to a farmer increases the probability of remaining a cotton grower by only 11 percent and is insignificant.

Table 2, Column 2, estimates an Ordered Probit model, where the decision to not plant is 0, the decision to plant but then realize no yields is assigned a 1, and the decision to plant and realize positive yields is assigned a 2. Our estimates reveal the significance of the SNI and TR in effecting the outcome variable. The ordering of these outcomes reflects the somewhat subjective perspective that planting and realizing positive yields is the best possible outcome. Of course, from a general equilibrium perspective, we cannot say that growing cotton is necessarily an optimal component a househould's production basket.

A hurdle model might also be appropriate, where the decision to plant is modeled as a Logit or Probit, and conditional on a non-zero yield the distribution is modeled as a Poisson. However, this model would not capture the difference between a zero yield due to no attempt to plant cotton vs. a zero yield where the farmer made an attempt but yielded zero, two susbtantially different decisions and outcomes. Table 2, Columns 3 and 4 estimate a hurdle Logit poission model. Column 3 shows that the SNI had both a significant impact on individuals' decision to continue growing cotton between 2009 and 2010, but also had a significant impact on the potential output that they realized. Even more surprising, is that the the SNI had a stronger and more significant positive impact on growing behavior than the TR.

Table 3 estimates Equation 3, the triple difference in difference (η coefficient), and difference in difference across the TR and SNI variables (γ and δ coefficients repsectively) on log of yields in Columns 3 and 4. We also run our estimations with yields in levels as we are interested in the interpretation of the programs' effects on yields, not log yields, in Columns 1 and 2. We are interested in the coefficients on $SNIxt$, γ, and $SNIxTRxt$, η, that is, the pure impact of the SNI intervention over time on outcomes, and the additive impact of the SNI relative to the TR treatment over time. At the same time, we also check that the estimated coefficients on SNI, and TR are insignificant. SNI, and TR are dummies for having been selected for the SNI and TR treatments. They capture whether selected households are significantly different in their yields from households who were not. Similarly, the t variable measures whether there is a significant time trend in yields, which we expect to be negative given the drop in yields between 2009 and 2010. The first four columns of Table 2 show our initial estimates in yields and log of yields. Selection into the programs was random as indicated by the insignificant effect of TR and SNI. The negative and significant coefficient on t reveals the downward trend in yields that is exhibited in the summary statistics of Table 1. The estimated impacts of $SNIxt$, $\hat{\gamma}$, and of $SNIxTRxt$, $\hat{\eta}$ are insignificant. However, both estimates are significant under the log yields specification in Columns 3 and 4. The additive effect of SNI on TR program is insignificant everywhere.

As Table 1 indicates, yields are overdispersed, where the variance in yields exceeds its mean. As Figure 1 shows, the average producer, before and after the treatments, is clustered below a 500 kilogram yield per year, so that the deviation from the mean yield is quite high for those few producers in the right tail of the distribution. Hence, the above result that the SNI treatment had an insignificant impact is not surprising if the upper portion of the yield distribution could gain nothing from the program. We would not expect a significant impact from social networks for the highest producers, who are already far above the mean yield, given that their knowledge base is likely saturated. It is farmers with production yields in the low to mid quantiles that we would expect to benefit the most from new networks. We did not exclude these farmers from the study, however, because they may play a critical role in information dissemination. We now look at the average impact of the program for those producers located around the mass of the yield distribution in Figure 1. These are individuals who yielded 500 kilograms per acre or less in 2009. Those who yielded greater than 500 kilograms per acre in 2009 are removed from the sample, which constitutes 15 percent of our original sample [20].

Columns 5 and 7 of Table 3 estimate Equation 3, conditional on having grown 500 kilograms of cotton per acre or less in 2009:

$$E(logy_t|y_t < 500) = E(\alpha + \beta t + \rho SNI_t + \mu TR_t + \nu SNIxTR + \eta SNxTRxt + \gamma SNIxt + \delta TRxt + u_t|y_t < 500) \tag{7}$$

They show that the SNI treatment has a positive and significant impact on yields and log yields respectively. Dropping down to households who harvested less than 400 kilograms of cotton in 2009, reveals an even greater impact of SNI, as shown in Columns 6 and 8. This result is also of economic significance, as the average cotton yield in rural Uganda is 100-200 kgs per year, and the significance of our effects extend to households who began with yields of up to 400-500 kilograms. That is, even households who are well above the mean yield, benefit from the SNI. In fact, the impact of the SNI program for these producers ranges from 66 to 74 additional kilograms of cotton per acre, which is 50 percent increase from the average farmer's

[20] A household is dropped from the sample in both years if its yield in 2009 was less than 500 to maintain a balanced panel

cotton yield between 2009 and 2010[21].

To gain a more complete picture of the SNI's impact on output for every quantile of producers by output, we plot the marginal impact of SNI on yields, conditional on households' yields being less than X kilograms per acre in 2009, where $X \in (0, 2000)$, and the estimates' corresponding t-statistics[22]. Figures 2 and 3 are smoothed plots of the estimates of γ, the impact of SNI, from Equation 3 for yields$< X$ for the total sample of households, and for female-headed households alone. These graphs confirm hypothesis (2), namely that the marginal effect of SNI is decreasing for higher yielding farmers. This also shows that the impact of the SNI is greatest for female-headed households producing up to 400 kilograms per year. Females producing between 0 and 400 kilograms per year experience an increase in yields of up to 70 kilograms per acre for the additional link that is added to their social network, as seen by the peak in the distribution in Figure 3. Additionally, this effect spills over to male-headed households in the lowest quantile of producers, i.e. those yielding up to 200 kilograms per year, as seen in Figure 4. The effects for males do not reach statistical significance at the 10 percent level, but are nevertheless non-negligible. This confirms hypothesis (3), namely that the value to a male farmer i of an additional link to person j is decreasing with the distance to j. Granted, we do not directly test for this relationship by mapping the networks within villages, and then estimating the average effect of the SNI program along specific network paths. However, villages are quite small, and the individuals in our sample are very likely connected to one another within a few degrees of separation. Therefore, males in the sample who did not participate in the SNI, are in some way connected to the females who did. Given that the females did expand their networks, the males likely did so as well, and this appears to have a muted effect on males' yields.

5.1.1 GLM Estimation

The above estimation assumes that $E[ln(y)|X] = Xb$, which shifts the distribution of yields below zero when a small constant, c, is added to zero valued yields. It may be more appropriate to assume: $ln(E[y|X]) = Xb$, which can be estimated by a generalized linear model with a log link. That is, the mean of the datum is linked to its predictors by a logarithmic function. The benefit of this specification is that the conditional mean should be positive, but the realized outcome can be zero (Nichols, 2010), something that occurs frequently in the labor literature with income and wage data, and with developing country data where yields and income

[21] The average yield across both years is 140 kilograms/acre. A 70 kilogram increase in output would result in 50 percent increase in yields for the average farmer.

[22] Using the triple difference estimate from a linear regression with yields in levels:

exhibit a mass near zero. We need only to specify a distribution for $(y_i|Z_i)$, so that the $E[y|X]$ is defined. Our results are robust to several distributional specifications (Gamma, Poisson, Gaussian), but modeling the conditional yields as a Poisson fit the data best. If $(yield_i|Z_i) \sim P(\mu_i)$, then the mean of the distribution is defined as $\mu_i = exp(\alpha + \beta t + \rho SNI_t + \mu TR_t + \nu SNIxTR_t + \eta SNxTrxt + \gamma SN_t xt + \delta TR_t xt)$. Table 4 estimates Equation 3 using a GLM log link and Poisson distribution. The significance of SNI's effect still holds for producers producing 400 kilograms or less, as seen in Column 2. The estimated marginal effects (not listed) of the SNI from the GLM estimation are a 30 percent increase in yields for the average farmer, a 36 percent increase for women, and a 19 percent increase for men.

5.1.2 SNI and Training as Substitutes

We have found that the complementary impact of SNI on TR is insignificant, that is, the estimate of η is insignificant in all of our specifications. This may because the TR program induces its own social networking effect such that SNI does not bring any additional gain to individuals who received TR. Therefore, each intervention seems to effect individuals' outcomes independently. We therefore, look at how the impact of the SNI (with and without TR) compares to the impact of TR (with and without SNI).

Table 4, Column 1 and 2 allow us to compare the effect of the SNI for those who received training compared to those who did not receive training in Columns 1 and 2. Column 2 compared to Column 1 shows that the SNI had its greatest economic and statistical impact for individuals who did not receive TR, where SNI increased yields by 74 kilograms per acre for those without TR[23], vs. 26 kilograms per acre for those who did receive TR, which was insignificant. In Column 4, we estimate the converse of Column 2; we estimate the impact of TR for individuals who did and did not receive the SNI. The results show that the effect of the TR where there was no SNI administered increased yields by 82 kilograms per acre and was significant, but increased yields by 34 kilograms per acre where SNI did occur and was insignificant. If we compare the effects of SNI vs. TR, we see that the two programs are of comprapable efficacy for increasing cotton yields. Therefore, the two programs appear to be feasible substitutes at increasing productivity in villages. The results hold true for the GLM specification as well, as shown in Columns 4-6.

Whether social capital behaves as a substitute or complement to standard training programs, may depend on the program type itself. (**?**) suggests that when programs are delivering private goods with large information

[23]Which we can also see in Table 3 Column 6

spillovers then the influence of social capital on information sharing is high. The highest returns to investments in social capital, however, are when "the economic good that a development project is designed to deliver is characterized by high levels of nonexclusiveness or non-rivalry." Of course, most training programs aim to deliver new knowledge, where knowledge is the quintessential public good. In that sense, we believe that the marginal investment in a social assessment will be relatively small compared to the potential benefits of the investment, regardless of whether the agricultural training itself is meant to deliver a private good.

What is significant about this finding is that whereas a training program requires the coordination of several agricultural extension agents[24], repeated travel to remote villages along unpaved roads, as well as coordination with the recipients of the training, the SNI is a one time pairing of individuals and dissemination of information. A training program such as TR would cost between 300-600 dollars per village per year. Considering that Uganda has over 95 districts, each with around 10 subcounties, and 5-10 villages per subcountry. With over five thousand villages, the cost of a training program could range from one to three million, depending on the number of trainers and their expertise. The SNI would amount to a one time travel cost and the time of one individual to organize the SNI. At the national level the SNI would cost in the order of one hundred to five hundred thousand.

5.2 Inputs

There are two channels through which the SNI could impact yields: (1) it could change the input decisions for cotton production and (2) it may change the techniques used by farmers (timing, weeding, thinning, and harvesting) to produce cotton. The differential impact of the SNI on outcomes between males and females may be a result of a change, or lack thereof, in either intermediate step. Our first check should naturally be to detect whether the SNI impacted the use of inputs in producing cotton.

We again look at the impact of the SNI on input use across the entire sample, and for males and females separately, using yields and log yields in a triple difference in Tables 6 and 7. These estimations suggest that there is a shift in the number of acres used for cultivating cotton, and a less consistent shift in the amount of seed used as a result of the SNI. These results are strongest for the log-transformed data, and remain significant for the female sample, but not male.

[24] Agricultural education, extension and training programs ensure that information on new technologies, plant varieties and cultural practices reaches farmers and those who need them most.

That said, another channel through which the SNI may have impacted output is through the improvement in planting techniques themselves. We will look at this in future research, where we can compare the knowledge correctly learned about planting cotton before and after the training and SN interventions through our tests on best practices in cotton production.

6 Conclusion

This is the first experimental design in the development literature to identify the causal impacts of social networks on productive agricultural outcomes. Previous research has not been able to claim a causal effect of social networks, while other literature estimated the effect of a training program at propagating new information across existing networks. One other study has implemented a comparable study to our own, but varies the frequency of meeting between financial network members to measure its effect on loan repayment. To circumvent these shortcomings, we exogenously perturbed networks, focusing only on females, whose output lags behind men's and whose potential to improve yields via social networks appeared to be large. Our estimates are robust to several specifications, including a general linearized model which approximates a linearized production function for yields of cotton.

We find that social networks have a causal impact on production. We estimated the SNI's own impact, and additive impact on the TR program using linear regression, log linear, and generalized linear model with a log link for mean yields. All of our results indicate that the SNI had a significant impact on yields for individuals who produced less than 500 kilograms per acre in 2009, where the average Ugandan farmer produces between 100 and 200 kilograms per acre per year. In particular, the difference in difference estimates of SNI on yields show that that an additional link to a female's social network increases yields by about 70 kilograms per acre, and this effect declines for the highest yielding farmers. Much of this impact is driven by an increase in females' yields in villages where there was no TR. This is a substantial finding, given that females comprise 80 percent of the agricultural labor force in Uganda, yet rarely receive direct agricultural training.

In addition to contributing to the development literature on social networks, the SNI also serves as a potential development program in and of itself. Given that the the average cotton producer yields 100-200 kilograms of seed cotton per year, the SNI increases a female's household productivity by 50-60 percent in our data, which is nearly as much as the training program does alone, but at a fraction of the cost. As (??, Ish) stresses, it is important that such findings on social capital be given a context for policy practitioners. "A mix a pinch of trust with a dash of social cohesion; then let simmer for six or seven centuries" is not a strategy for development ?? (Ish). As a practical implementation of our findings, we advocate that for optimal effects, programs similar to the actual social network intervention outlined here be used, rather than the more traditional idea of advocating groups, such as farmers' groups or female groups. We find, from the experimental extension of this work (Vasilaky (2010)), and through our qualitative studies, that promoting groups strengthens already existing social structures in a village, while developing new and random links helps to propagate new information from peripheral individuals, whose voice might otherwise be subsumed by a well regulated social structure. Furthermore, we find that competitive incentives rather than group incentives are better at propagating information exchange. The social network intervention outlined and tested here was successfully designed to achieve that.

Agricultural development is a foundational step to progressing past rural poverty, yet one of the most difficult to implement. Agricultural programs require time, inputs, and transmission of new knowledge. The SNI acted as a program which increased farmers' output without the continual intervention of outside agents, which the TR required. In addition to potentially being a substitute for training programs, the SNI also circumvents the potential biases that exist in developing countries towards male focused training. As such, our findings are relevant to the developing country context where males disproportionately receive more training programs as compared to females, and where there is limited information exchange across genders. This is true not only in East Africa, but in rural areas of Latin America, India, and the Caribbean. Social networks can be instrumental at increasing productivity and can substitute for traditional training programs at a fraction of the cost.

References

Social Capital and Economic Development: Well-being in Developing Countries. Edward Elgar Publications.

(2000). Committee on the convention on the elimination of all forms of discrimination against women: Report of states parties: Uganda (july 3, 2000), cedaw/c/uga/3 at 57. Technical report, Sponsored by the UN.

(2005). Building uganda's global competitiveness in agribusiness, 2005-2010: Uganda's cotton competitive plan. *Working Draft. Kampala: SCOPE.*.

Acemoglu, D. (2007). *Introduction to modern growth, manuscript*. Cambridge, MA: The MIT Press.

Acemoglu, D., M. Dahleh, I. Lobel, and A. Ozdagla (2008). Bayesian learning in social networks. *Working Paper*.

Agree, E. M., A. E. Valente, and W. Thomas (2005). Intergenerational transfers of tesources between older persons and extended kin in taiwan and the philippines. *Population Studies 59*(2), 181–195.

Baffes, J. (2008a). The 'full potential' of uganda's cotton industry. *Development Prospects Group, The World Bank. September 2008*.

Baffes, J. (2008b). The gender dimension of cotton productivity in uganda. *Accepted project proposal from the Gender Action Plan, World Bank March 2008*.

Bandiera, O. and I. Rasul (2006). Social networks and technology adoption in northern mozambique. *The Economic Journal 116*(514), 869–902.

Behrman, J. R., H.-P. Kohler, and S. C. Watkins (2001). The density of social networks and fertility decisions: evidence from south nyanza district, kenya. *Demography 38*(1), 43–58.

Behrman, J. R., H.-P. Kohler, and S. C. Watkins (2002). Social networks and changes in contraceptive use over time: evidence from longitudinal study in rural kenya. *Demography 39*(4).

Besteman, C. (1995). Agricultural extension. In E. H. Helen Kreider Henderson (Ed.), *Gender and agricultural development: Surveying the field*, pp. 105–37. Arizona: University of Arizona Press.

Bramoulle, Y., H. Djebbari, and B. Fortin (2). Identification of peer effects through social networks. *IZA DP No. 2652*.

Cameroon, C. and P. Trivedi (2005). *Microeconometrics: Methods and Applications*. Cambridge, MA: Cambridge.

Caughman, S. and M. N. Thiam (1989). The markala, mali cooperative: A new approach to traditional economic roles. In A. Leonard (Ed.), *Seeds*, pp. 105–37. CUNY, NY: Feminist Press.

CGIAR (2001). Livelihood zones of Uganda. Technical report, Foodnet at Consultative Group on International Agricultural Research (CGIAR).

Chambers, R. (1993). London: Intermediate Technology Publications.

Conley, T. and C. Udry (2001). Social learning through networks: The adoption of new agricultural technologies in ghana. *American Journal of Agricultural Economics 83*, 668–732.

Conley, T. G. and C. R. Udry (2010). Learning about a new technology: Pineapple in ghana. *American Economics Review 100*(1), 35–69.

Cornelius, C. E. (2001, Februrary). Cotton subsector development project. World Bank Policy Research Working Paper 4122, World Bank.

Darr, D. and J. Pretzsch (2008). The spread of innovations within formal and informal farmers groups-evidence from rural communities of semi-arid eastern africa. *Institute of International Forestry and Forest Products, Germany, Working Paper*.

Davis, B., G. Stecklov, and P. Winters (2005). Domestic and international migration from rural mexico: Disaggregating the effects of network structure and composition. *Population Studies 59*(3), 265–282.

Duflo, E. (2007). Using randomization in devleopment economics research: A toolkit. *Working Paper No. 6059, CEPR, London*.

Duflo, E., M. Kremer, and J. Robinson (2006). Understanding technology adoption: Fertilizer in western kenya, evidence from field experiments. *Manuscript*.

Duflo, E. and E. Saez (2003). The role of information and social interactions in reteriment plan decisions: Evidence from a randomized experiment. *Quarterly Journal of Economics 118*(3), 1815–842.

Edmeades, S., E. Katungi, and M. Smale (2008). Gender, social capital and information exchange in rural uganda. *Journal of International Development 20*(1), 35–52.

Ellis, A., C. Manuel, and M. Blackden (2006). *Gender and economic growth in Uganda: Unleashing the power of women.* Washington DC: The World Bank.

Field, E., B. Feigenberg, and R. Pande (2010). Building social capital through microfinance. *University of Maryland Seminar, November 9, 2010*, http://econweb.umd.edu/ davis/eventpapers/FieldBuilding.pdf.

Foster, A. and M. Rosenzweig (1995). Learning by doing and learning from others: Human capital and technical change in agriculture. *Journal of Political Economy 103*(6), 1176–1209.

Gates (2009). Seattle, WA: Bill and Melinda Gates Foundation.

Goldstein, M. (2002). Chop time, no friends: Intrahousehold and individual insurance mechanisms in southern ghana. *Manuscript. Long, UK: London School of Economics*.

Goldstein, M. and C. Udry (1999). Agricultural innovation and resource management in ghana. *Final Report to IFPRI under MP17*.

Goldstein, M. and C. Udry (2004). Gender, power and agricultural investment in ghana. *Manuscript*.

Goldstein, M. and C. Udry (2008). The profits of power: Land rights and agricultural investment in ghana. *Journal of Political Economy 116*(6).

Granovetter, M. (1974). The strength of weak ties. *American Journal Sociology 78*(1), 1360–1380.

Granovetter, M. (2005). The impact of social structure on economic outcomes. *Journal of Economic Perspectives 19*(1), 33–50.

Helleringer, S. and H.-P. Kohler (2005). Social networks, perceptions of risk, and changing attitudes towards hiv/aids: New evidence from a longitudinal study using fixed-effects analysis. *Population Studies 59*(3), 265–282.

Hoang, L. A., J.-C. Castella1, and P. Novosad (2006). Social networks and information access: Implications for agricultural extension in a rice farming community in northern vietnam. *Journal Agriculture and Human Values 23*(4), 1572–8366.

Hoddinott, J., S. Dercon, and P. Krishnan (2005). Networks and informal mutual support in 15 ethiopian villages. *Unpublished manuscript*.

Isham, J. (2002). The effects of social capital on technology adoption: Evidence from rural tanzania. *Journal of African Economies 11*(1), 39–60.

Jackson, M. O. and A. Wolinsky. A strategic model of social and economic networks.

Janice Jiggins, P. M. and M. Masona (1995). Breaking new ground: Reaching out to women farmers in western zambia. In A. Leonard (Ed.), *Seeds 2*, pp. 105–37. CUNY, NY: Feminist Press.

John Maluccio, L. H. and J. May (1999). Social capital and income generation in south africa, 1993b98. *Food Consumption and Nutrition Division, International Food Policy Research Institute,FCND DISCUSSION PAPER NO. 71*.

J.R.Bibangambah (1996). *Marketing of Smallholder Crops in Uganda*. Kampala, UG: Fountain Publishers.

Katungi, E., S. Edmeades, and M. Smale (2006). Gender, social capital and information exchange in rural uganda. *Collective Aciton and Property Rights Working Paper No. 5 December 5, 2006 International Food Policy Research Institute*.

Kilpatrick, S. (2007). Building social capital in groups: Facilitating skill development for natural resource management. *Journal Agriculture and Human Values 17*(3), 248.

Kremer, E. M. M. (2003). Worms: Identifying impacts on education and health in the presence of treatment externalities. *Econometrica 72*(1), 159–217.

Kremer, M., E. Miguel, and R. Thornton (2007). Incentives to learn. *Manuscript*.

Leonard, K. (2007). Learning in health care: Evidence of learning about clinician quality in tanzania. *Economic Development and Cultural Change 55*, 531–555.

Maertens, A. (2010). Who cares what others think (or do)? social learning, social pressures and imitation in cotton farming in india. *University of Pittsburgh, Working Paper*.

Maluccio, J., L. H. L, and J. May (2003). Social capital and gender in south africa. In A. Quisumbing (Ed.), *In Household Decisions, Gender and Development, A Synthesis of Recent Research*. Washington DC: International Food Policy Research Institute.

Manski, C. (1993). Identification of endogenous social effects: The reflection problem. *The Review of Economic Studies 60*(1), 531–542.

Matuschke, I. (2008). Evaluating the impact of social networks in rural innovation systems. *IFPRI Discussion Paper 00816*.

Matuschke, I. and M. Qaim (2009). The impact of social networks on hybrid seed adoption in india. *Agricultural Economics 40*, 493–505.

Mouw, T. (2003). Social capital and finding a job: do contacts matter. *American Sociological Review 68*, 868–870.

Munshi, K. (2004). Social learning in a heterogenous population: Technology diffusion in the indian green revolution,. *Journal of Development Economics 73*, 185–213.

Mwaka, V. M. (1993). Agricultural production and women's time budgets in uganda. In J. H. Momsen and V. Kinnaird (Eds.), *Different Places, Different Voices: Gender and Development in Africa, Asia and Latin America*, pp. 46–52. London, UK: Routledge.

Nichols, A. (2010). Regression for nonnegative skewed dependent variables. Technical Report 2, BOS10 Stata Conference.

Quisumbing, A. (2003). Social capital, legal institutions, and property rights: Overview. In A. Quisumbing (Ed.), *Household Decisions, Gender and Development*, pp. 139–144. Washington DC: International Food and Policy Research Institute.

Rangel, M. and D. Thomas (2003). Out of west africa: Evidence on the efficient allocation of resources within farm households. *Harris School Working Paper Series 05.15*.

Romer, P. (1986). Increasing returns and long-run growth. *Journal of Political Economy 94*(5), 1002–1037.

Romer, P. (1990). Endogenous technological change,. *Journal of Political Economy 98*, S71–S101.

Stevenson, L. and A. St-Onge (2005). Support for growth-oriented women entrepreneurs in uganda. *International Labour Office, Geneva*.

Sylvester, C. (2000). *Producing Women and Progress in Zimbabwe*. Portsmouth, NJ: Heinemann.

Tanzarn, N. (2005). Revisiting the past, reflections on the future: Gender in science, technology and agricultural development. In N. Tanzarn (Ed.), *Gender in Agriculture and Technology*, pp. xvii–xxxii. Kampala, Uganda: Women and Gender Studies.

Udry, C. (1996). *Working Paper No. 6059, CEPR, London*.

Udry, C. (1996-1998). Ghana survey drafts. http://www.econ.yale.edu/ cru2/Surveys/.

Udry, C. and M. Goldstein (2006). Addressing unequal economic opportunities, a case study of land tenure in ghana. *World Bank, Development Outreach*.

USAID (2008). Cotton value chain case study for northern uganda guided case studies in value chain development for conflict-affected environments. Technical Report 91.

Vasilaky, K. N. (2010). Incentives for information exchange:getting women to share in rural uganda. *Univeristy of Maryland, Work in Progress*.

Wooldridge, J. M. (2008). *Econometric analysis of cross Section and panel data, 2nd Edition.* Cambridge, MA: The MIT Press.

You, L. and J. Chamberlin (2004). Spatial analysis of sustainable livelihood enterprises of uganda cotton production. *EPTD Discussion Paper, 121. International Food Policy Research Institute. Washington D.C.*.

Young, P. (2009). Innovation diffusion in heterogeneous populations-contagion, social influence, and social learning. *American Economic Review 99*, 1899–1924.

Table 1: Means over time

Means in 2009 & 2010			
	2009	2010	Total
SNI	0.475	0.475	0.478
	(0.500)	(0.500)	(0.500)
TR	0.658	0.658	0.660
	(0.474)	(0.475)	(0.474)
sex	1.488	1.480	1.484
	(0.501)	(0.500)	(0.500)
kgs cotton	140.8	79.54	109.9
	(201.5)	(129.2)	(171.6)
[1em] acres	0.983	0.586	0.783
	(0.701)	(0.593)	(0.678)
yield (kgs/acre)	182.0	139.5	160.6
	(208.7)	(234.9)	(223.1)
kgs seed	4.976	3.232	4.097
	(3.799)	(3.000)	(3.527)
yperseed	52.83	36.96	44.83
	(78.32)	(62.70)	(71.27)

Mean of each variable with standard deviation in parentheses.

Table 2: Probit, Oprobit, HPlogit

EQUATION	VARIABLES	(1) 0=Dropped logit & poisson	(2) 0=Dropped 1=Attempted 1=Attempted&0 2=>0	(3)	(4)
planter	SNI		0.699***	0.922**	
			(3.115)	(2.208)	
	TRAINING		0.428**	0.578	
			(2.286)	(1.614)	
	TrxSNI		0.0159	0.212	
			(0.0558)	(0.409)	
SINGLE	SNI	0.565**		0.922**	
		(2.272)		(2.208)	
	TRAINING	0.334		0.578	
		(1.579)		(1.614)	
	TrxSNI	0.0657		0.212	
		(0.207)		(0.409)	
	TrxSNI	(0.554)		0.212	-0.184
					(-1.078)
	SNI				0.335**
					(2.022)
	TRAINING				0.241
					(1.536)
	Constant				-0.0161
					(-0.106)
	Observations	325	325	325	325

Robust z-statistics in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 3: Triple and Double Differences

Double and Triple Differences

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
							DD	DD
					DD	DD	LnYield	LnYield
VARIABLES	TRIPLE D: Yield	DD	TRIPLE D:Ln Yield	DD	Yield<500	Yield<400	Yield<500	Yield<400
t	-97.84***	-103.9***	-3.127***	-2.965***	-87.63***	-76.74***	-3.061***	-3.176***
	(-4.275)	(-4.951)	(-8.404)	(-7.631)	(-3.506)	(-3.685)	(-7.751)	(-7.013)
SNI	58.85	48.78	-0.0983	0.123	-11.32	-8.464	-0.428	-0.419
	(1.569)	(1.480)	(-0.251)	(0.577)	(-0.460)	(-0.359)	(-1.164)	(-1.034)
TR	26.98	19.30	0.124	0.292	-3.574	8.357	0.0257	0.0950
	(0.745)	(0.651)	(0.368)	(1.228)	(-0.147)	(0.432)	(0.0809)	(0.311)
TrxSNI	-15.22		0.334		42.23	31.53	0.681	0.638
	(-0.256)		(0.725)		(1.282)	(1.050)	(1.621)	(1.420)
SNIxt	1.332	12.70	1.593**	1.291**	66.68*	74.69***	1.795**	2.118***
	(0.0457)	(0.461)	(2.699)	(2.502)	(1.828)	(2.899)	(2.592)	(3.428)
TRxt	75.83*	84.51***	1.160	0.929*	100.6**	82.73*	1.165	1.234
	(1.791)	(3.483)	(1.645)	(1.979)	(2.163)	(1.881)	(1.659)	(1.667)
TRxSNIxt	17.20		-0.457		-38.33	-48.27	-0.725	-0.989
	(0.352)		(-0.491)		(-0.632)	(-0.909)	(-0.723)	(-1.025)
Constant	140.2***	145.6***	4.534***	4.415***	130.8***	113.2***	4.496***	4.398***
	(7.076)	(6.206)	(17.07)	(19.39)	(6.921)	(8.307)	(16.98)	(17.41)
Observations	646	646	646	646	596	574	596	574
R-squared	0.047	0.047	0.238	0.237	0.047	0.045	0.232	0.245

Robust t-statistics in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Figure 1:

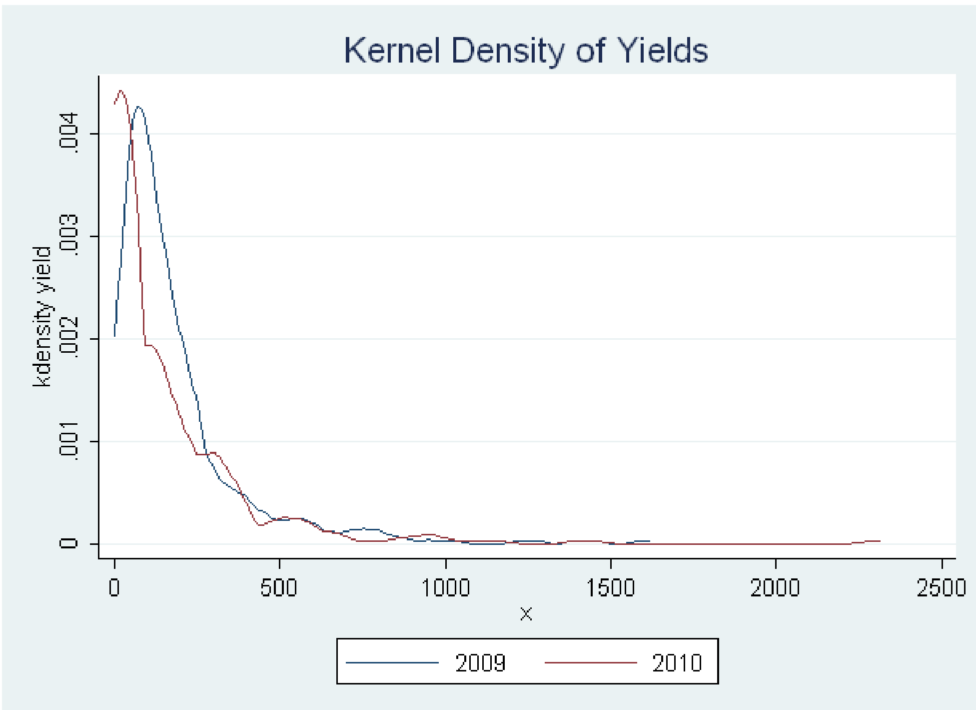

Kernel Density of Yields
kdensity yield
0
.001
.002
.003
.004
0
500
1000
1500
2000
2500
x
2009
2010

Table 4: Double Difference, GLM

EQUATION	VARIABLES	(1) DD	(2) DD yield1<400 GLM	(3) DD F yield1<400 GLM	(4) DD M yield1<400 GLM
yield1	t	-0.944***	-1.109***	-1.482***	-1.012***
		(-3.714)	(-2.989)	(-2.760)	(-2.706)
	SNI	0.268	-0.155	-0.497*	0.0489
		(1.484)	(-0.767)	(-1.916)	(0.226)
	TR	0.107	-0.0277	-0.0124	-0.0213
		(0.659)	(-0.149)	(-0.0680)	(-0.102)
	TrxSNI		0.348	0.592*	0.226
			(1.389)	(1.774)	(0.846)
	SNIxt	0.173	1.001**	1.468**	0.862*
		(0.682)	(2.351)	(2.126)	(1.910)
	TRxt	0.774***	1.206***	1.020*	1.554***
		(3.502)	(2.621)	(1.669)	(3.279)
	TRxSNIxt		-0.830	-0.694	-1.168**
			(-1.585)	(-0.884)	(-2.096)
	Constant	4.993***	4.874***	4.847***	4.884***
		(34.63)	(33.93)	(49.53)	(28.92)
	Observations	646	592	288	304

Robust z-statistics in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 5: Double Difference, GLM

EQUATION	VARIABLES	(1) Linear ALL yield1<400 Trained	(2) Linear ALL yield1<400 Untrained	(3) Linear ALL yield1<400 SNI	(4) Linear ALL yield1<400 No SNI	(5) GLM ALL yield1<400 Trained	(6) GLM ALL yield1<400 Untrained	(7) GLM ALL yield1<400 SNI	(8) GLM ALL yield1<400 No SNI
Linear	t	5.996	-76.74***	-2.049	-76.74***				
		(0.153)	(-3.556)	(-0.132)	(-3.623)				
	SNI	23.06	-8.464						
		(1.227)	(-0.347)						
	SNIxt	26.42	74.69**						
		(0.564)	(2.797)						
	TRAINING			39.89	8.357				
				(1.702)	(0.424)				
	TRxt			34.47	82.73*				
				(1.135)	(1.850)				
		(8.764)	(8.016)	(5.341)	(8.168)				
	Constant	121.6***	113.2***	104.8***	113.2***				
GLM	t					0.0481	-1.132**	-0.0198	-1.132**
						(0.159)	(-2.478)	(-0.129)	(-2.517)
	SNI					0.174	-0.0777		
						(1.212)	(-0.344)		
	SNIxt					0.154	1.112**		
						(0.466)	(2.305)		
	TRAINING							0.323	0.0712
								(1.566)	(0.426)
	TRxt							0.222	1.180**
								(1.092)	(2.172)
	Constant					4.801***	4.729***	4.652***	4.729***
						(42.24)	(38.22)	(24.99)	(38.82)
	Observations	386	188	265	309	386	188	265	309
	R-squared	0.014	0.073	0.041	0.033				

Robust t-statistics in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 6:Triple Diff of Seeds(kgs

and Acreage Used for Cotton

	(1)	(2)	(3)	(4)	(5)	(6)
VARIABLES	Seed (kgs)	Acreage	Seed:Females	Acreage:Females	Seed:Males	Acreage:Males
t	-2.577***	-0.691***	-2.717***	-0.425***	-2.571***	-0.800***
	(-5.890)	(-6.707)	(-5.434)	(-3.665)	(-3.955)	(-5.198)
SNI	0.385	-0.308	0.676	0.124	0.409	-0.498*
	(0.392)	(-1.473)	(0.771)	(1.049)	(0.296)	(-1.819)
TR	0.198	-0.155	1.128	0.269**	-0.397	-0.383
	(0.280)	(-0.827)	(1.393)	(2.762)	(-0.471)	(-1.515)
TrxSNI	-1.556	0.190	-1.873	-0.292*	-1.447	0.459
	(-1.296)	(0.772)	(-1.701)	(-1.779)	(-0.870)	(1.370)
SNIxt	0.492	0.382**	1.823	0.108	-0.378	0.497**
	(0.385)	(2.143)	(1.332)	(0.523)	(-0.212)	(2.267)
TRxt	0.494	0.313**	0.109	-0.0143	1.241	0.514**
	(0.680)	(2.149)	(0.129)	(-0.0841)	(1.143)	(2.741)
TRxSNIxt	0.894	-0.281	0.126	0.0348	0.957	-0.466
	(0.590)	(-1.222)	(0.0770)	(0.142)	(0.451)	(-1.557)
sex	-1.402***	-0.329***				
	(-5.570)	(-5.619)				
Constant	7.205***	1.651***	3.679***	0.627***	6.083***	1.465***
	(13.72)	(8.294)	(5.877)	(9.675)	(12.94)	(6.439)
Observations	645	650	312	315	333	335
R-squared	0.114	0.150	0.154	0.146	0.066	0.099

Robust t-statistics in parentheses

Table 7: Triple Diff of Ln(Seeds
(kgs
and Ln(Acreage
Used for Cotton

	(1)	(2)	(3)	(4)	(5)	(6)
VARIABLES	LnSeed (kgs)	LnSeed (kgs)	LnSeed:Fem	LnAcreage:Fem	LnSeed:M	LnAcreage:M
t	-2.867***	-2.181***	-3.715***	-2.469***	-2.580***	-2.099***
	(-12.21)	(-17.88)	(-10.66)	(-7.414)	(-6.615)	(-9.859)
SNI	0.0184	-0.277	0.171	0.0832	-0.113	-0.457*
	(0.101)	(-1.275)	(0.971)	(0.316)	(-0.585)	(-2.052)
TR	0.107	-0.0287	0.217	0.393**	-0.141	-0.391*
	(0.801)	(-0.161)	(1.358)	(2.130)	(-1.046)	(-2.008)
TrxSNI	-0.405	0.00593	-0.672**	-0.415	-0.0670	0.336
	(-1.574)	(0.0199)	(-2.135)	(-1.141)	(-0.250)	(1.012)
SNIxt	1.260*	1.139**	2.077*	1.413*	0.990	1.067**
	(1.744)	(2.274)	(1.996)	(1.900)	(1.329)	(2.158)
TRxt	0.751	0.717	1.110	0.688	1.166	1.093**
	(1.165)	(1.593)	(1.454)	(1.140)	(1.546)	(2.318)
TRxSNIxt	0.124	-0.312	-0.203	-0.475	-0.312	-0.488
	(0.125)	(-0.436)	(-0.155)	(-0.502)	(-0.294)	(-0.654)
sex	-0.582***	-0.603***				
	(-4.169)	(-4.985)				
Constant	2.290***	0.725***	1.198***	-0.725***	1.680***	0.216
	(12.65)	(3.571)	(10.73)	(-5.355)	(26.58)	(1.568)
Observations	645	650	312	315	333	335
R-squared	0.211	0.203	0.256	0.226	0.167	0.158

Robust t-statistics in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Figure 2:

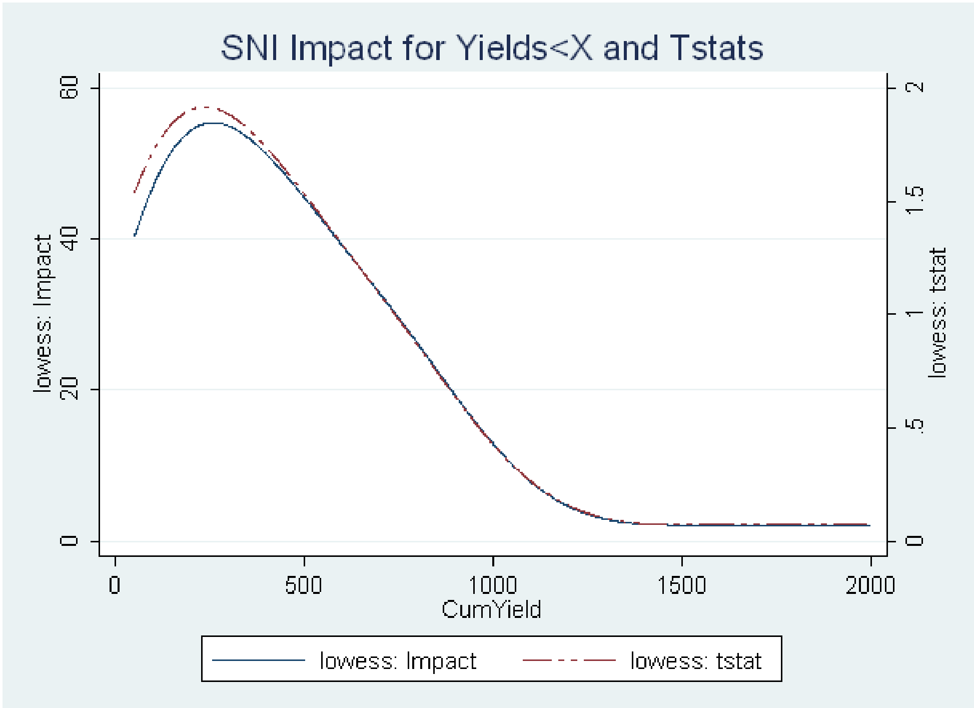


Figure 3:

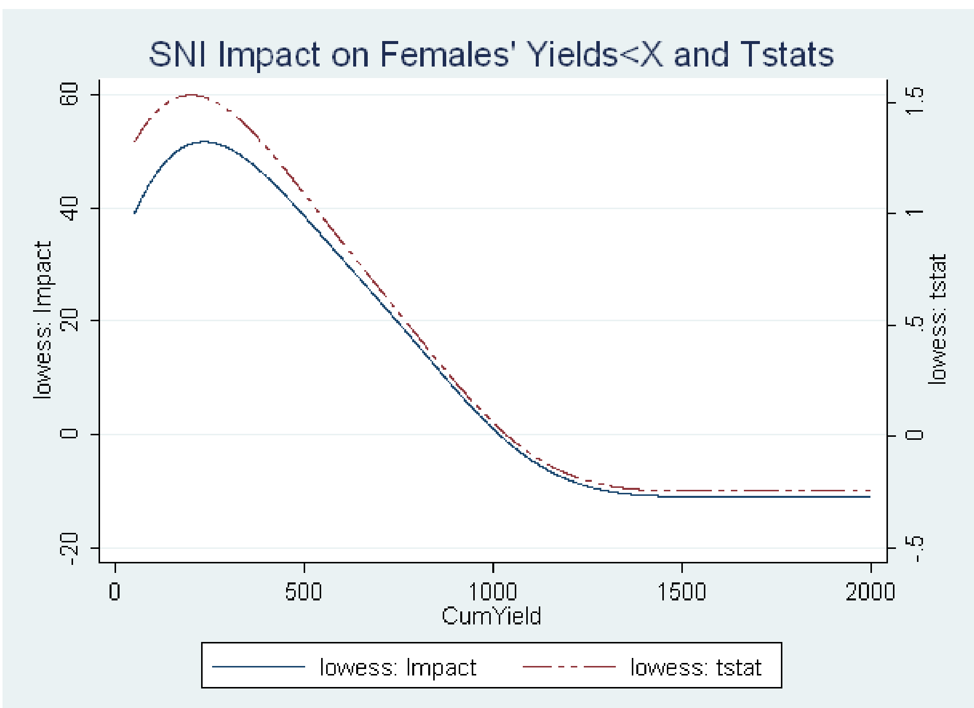


Figure 4:

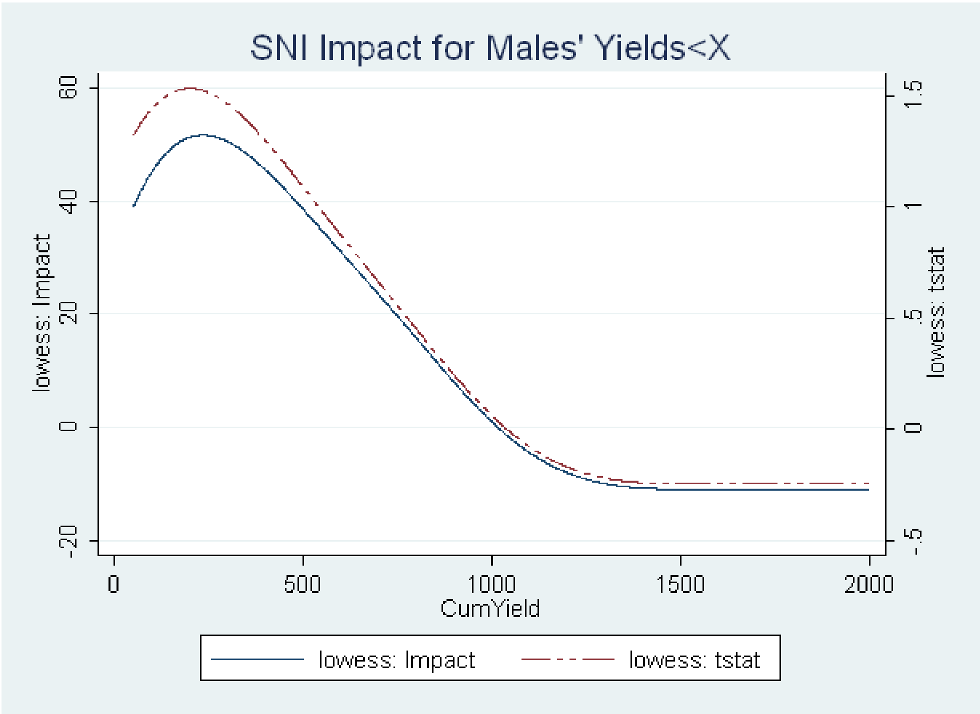


Figure 5:

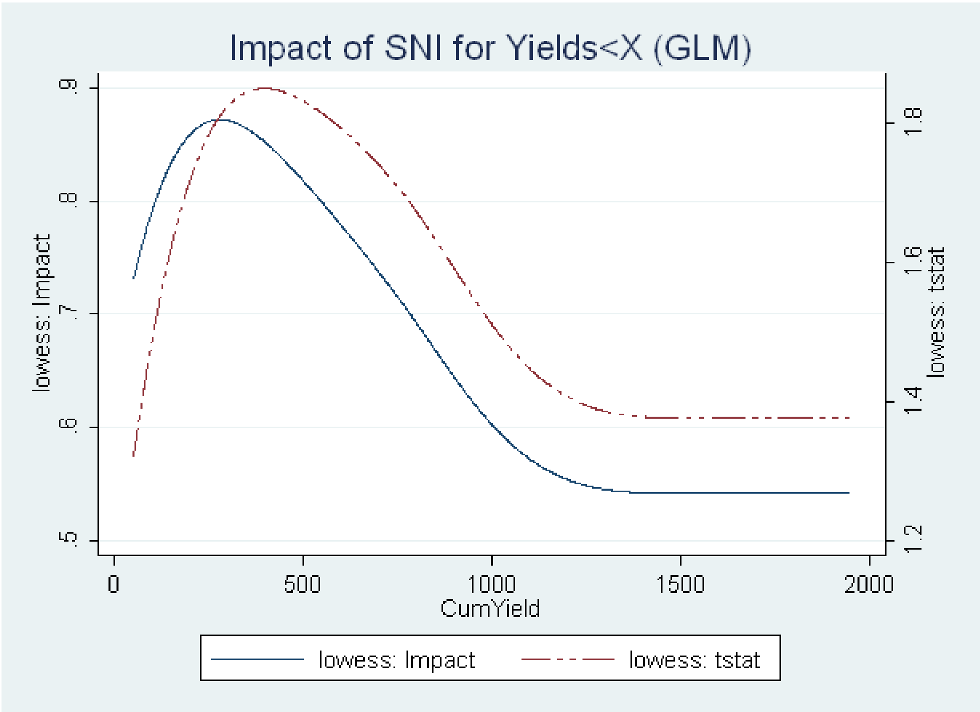
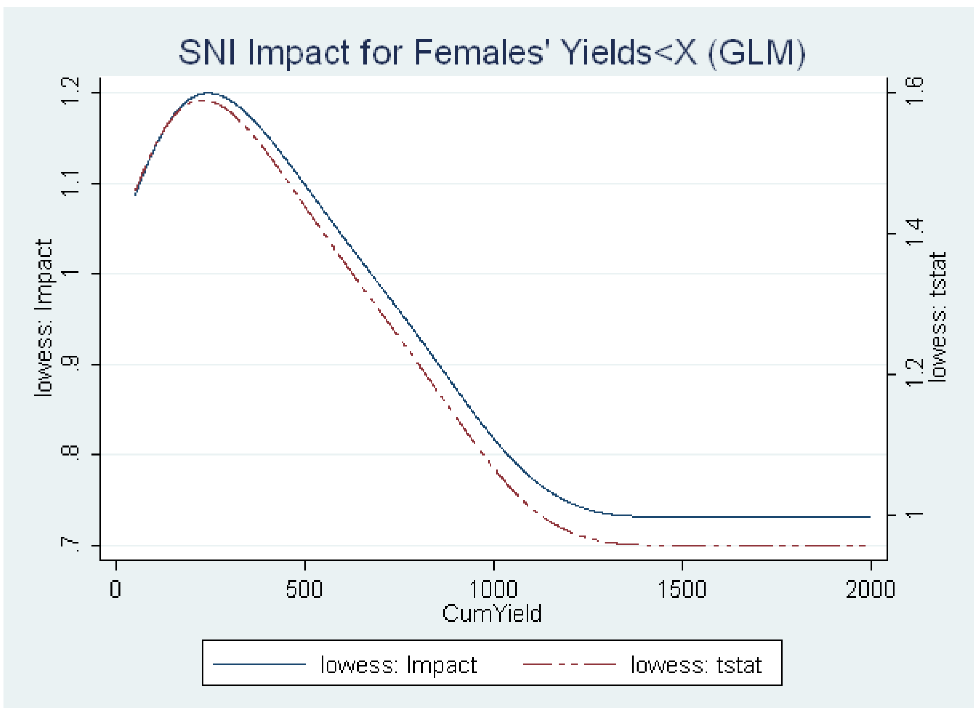


Figure 6:



SNI Impact for Females' Yields<X (GLM)
lowess: Impact
lowess: tstat
CumYield
.7
.8
.9
1
1.1
1.2
1
1.2
1.4
1.6
0
500
1000
1500
2000
lowess: Impact
lowess: tstat