Title
Essays on behavioral responses to development interventions

Author
Emerick, Kyle Jared

Publication Date
2014

Peer reviewed|Thesis/dissertation

Essays on behavioral responses to development interventions

by

Kyle Jared Emerick

A dissertation submitted in partial satisfaction of the
requirements for the degree of
Doctor of Philosophy

in

Agricultural and Resource Economics

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:

Professor Elisabeth Sadoulet, Chair
Professor Edward Miguel
Professor Jeremy Magruder

Spring 2014

Essays on behavioral responses to development interventions

Abstract

Essays on behavioral responses to development interventions

by

Kyle Jared Emerick

Doctor of Philosophy in Agricultural and Resource Economics

University of California, Berkeley

Professor Elisabeth Sadoulet, Chair

This dissertation combines three papers which are all empirical analyses of agricultural interventions in developing countries. I focus on how new policies, technologies, and institutions affect the behavior of small-scale farmers in both Mexico and India. The first paper focuses on the certification of agricultural land in Mexico while the second and third papers focus on technology adoption in rural India.

Chapter 1, which is based on joint work with Alain de Janvry, Marco Gonzalez-Navarro, and Elisabeth Sadoulet, shows that removing the link between active land use and ownership through certification leads to a reallocation of labor away from agriculture and towards migration. In particular, we use the rollout of the Mexican land certification program from 1993 to 2006 to show that households obtaining land certificates were subsequently 28% more likely to have a migrant member. This response was differentiated by initial land endowments, land quality, outside wages, and initial land security, as predicted by our model. Effects on land under cultivation were heterogeneous: in high land quality regions land under cultivation increased while in low quality ones it declined.

Chapter 2, which is based on joint work with Alain de Janvry, Elisabeth Sadoulet, and Manzoor Dar, shows evidence that risk is an important factor that constrains the decisions made by small farmers. More specifically, the chapter reports results of a field experiment in Odisha India that quantifies the effects of Swarna-Sub1, a promising new rice seed that effectively reduces risk by sharply reducing the susceptibility of the crop to flood damage. In doing so, the chapter offers novel evidence on the effect of a direct reduction in production risk on economic behavior. Specifically, access to this new technology leads to increases in area cultivated, fertilizer used, and the likelihood of using a more modern planting method. Also, the technology reduces precautionary savings of grain for consumption and increases the use of agricultural credit. An important implication from the chapter is that technological progress that directly eliminates weather-induced production variability offers a promising method of advancing agriculture in areas that are prone to extreme weather.

Chapter 3 builds on the promising results in Chapter 2 by studying diffusion of Swarna-Sub1. I provide an experimental test of whether informal exchange of Swarna-Sub1 between

farmers produces an efficient allocation. I report results on a field experiment, also in Odisha, to compare decentralized trade of Swarna-Sub1 through networks with an approach where demand was revealed via door-to-door sales. While 84% of farmers are expected to gain from Swarna-Sub1, only 7% adopt in networks. Conversely, 40% of farmers adopt when demand is revealed in door-to-door sales. Using variation across the sample in estimated gains in revenue, I show that 63% of the gains from door-to-door sales are lost with decentralized trade through networks. Frictions preventing interactions between farmers from different social groups offer an explanation for the results. Sub-caste and surname association with suppliers are strong predictors of adoption in networks, but have no effect in door-to-door sales. The main implication from the chapter is that relying on exchanges between farmers to disseminate new seed varieties will not produce an allocation where demand is met.

Contents

List of Figures

List of Tables

Acknowledgments

I am very grateful to both Elisabeth Sadoulet and Alain de Janvry for their guidance. I will always be thankful for them supporting me throughout my studies, including in the beginning. They gave me a chance to succeed, which I truly appreciate.

Of course, I am thankful to my wife Em for her constant support. She always let me focus on my studies when needed. In addition, she was always there to hear me complain when I felt the need, which is probably too often. Our son Gavin has also been quite an inspiration by always giving us something to smile about.

Finally, I am thankful for my classmates at Berkeley. In particular, I have benefitted from numerous discussions about research with Marshall Burke, Sam Heft-Neal, and Manuel Barron. I have also enjoyed the numerous discussions about things other than research, most often sports.

Chapter 1

Land certification and migration in Mexico

1.1 Introduction

Well-defined and secure property rights over land have long been recognized as essential for economic development (Demsetz, 1967; North and Thomas, 1973; De Soto, 1989). There are however different ways in which these rights can be established. Contrary to the norm in developed countries in which rights are established with land tiles, in many developing countries rights are established by contingent use. In the latter case, security of access requires evidence of active use (production); i.e., leaving land idle implies a risk of reallocation without compensation. This can be inefficient as it imposes restrictions on the amount of labor used on the land by requiring that it be kept in production at an accepted standard of use, ignoring the return to labor in other activities. With a focus on improving the security of access to land and stimulating investment, land certification and titling programs have been proposed (De Soto, 2000), resulting in the implementation of large-scale certification programs sponsored by national governments and international development agencies (Heath, 1990). While the focus has been on land productivity, little attention has been given to the potentially large effects on the spatial reallocation of labor. The importance of this effect becomes clear once one considers that in developing countries value added per worker is on average four times higher in the non-agricultural sector than in agriculture (Gollin et al., 2012). For the specific case of Mexico, in the early 1990s agriculture represented only 3.8% of GDP while 34.4% of the population lived in rural areas.

In reviewing the literature, Galiani and Schargrodsky (2011) find that the benefits from well-defined and secure property rights over land can materialize through four channels: enhanced investment incentives (Alchian and Demsetz, 1973; Lin, 1992), facilitation of land trades (Besley, 1995; Deininger, 2003), increased use of land as collateral to access credit (Feder, Onchan, and Chalamwong, 1988; De Soto, 2000), and improved intra-household labor allocations (Field, 2007). There is no clear distinction, however, as to whether rights are established by use or by certification/titling, for as long as they are well defined and secure. Yet, the difference on labor and land use can be very important: use-based rights can restrain migration out of agriculture and keep inferior land in production (Feder and Feeny, 1991).

The classic economic argument regarding the impact of weak property rights on migration is based on treating insecurity as a tax on output. Improving property rights is then predicted to increase the marginal products of agricultural land and labor, decreasing incentives to migrate. In this paper, we argue that a pre-title regime where use-based property rights require presence of the owner on the land and his active use of the land, creates a distortion working in the opposite direction, inefficiently tying labor to the land.[1] We use a simple household model to show that implementation of a land certification program delinking land rights from land use can lead to increased outmigration. In the model, the inefficient labor tying result rests on two main conditions: a preexisting suboptimal farm size and the land use requirement.

We test the model's predictions using data from Mexico's large-scale land certification program (Programa de Certificación de Derechos Ejidales y Titulación de Solares, or Procede). The program was rolled out nationwide from 1993 to 2006 to issue certificates of ownership over ejido land. Ejidos are agrarian communities that were created over the 1914 to 1992 period as part of an ambitious land reform program in which community members (ejidatarios) were granted use and residual claimant rights over individual agricultural plots. Security of access for individuals was closely linked to usage. Any land that was left fallow for more than two years could be granted to another beneficiary. Procede revoked this pattern of property rights. It gave ejidatarios land certificates specifying the name of the owner of each agricultural plot alongside with a GIS-based map of the plot. Similar documents were provided for residential plots, while a certificate was issued to each ejidatario giving ownership of a share of common use lands. Procede was massive in scale, providing certificates to over 3.6 million families by the end of the program. We use this large-scale land certification experiment to assess the migration and land reallocation impacts of redefining property rights from use-based to title-based.

Because the program provided certificates to the entire community simultaneously, selection concerns are minimized.[2] We use a fixed-effects econometric specification that compares

[1]There are many examples of use-based property rights with implications on the efficiency of land use. In Brazil, cultivation of more than 50% of the potentially productive area in large farms is required by the constitution of 1988 as a "social obligation" of land ownership, with the right to expropriate at the demand of occupants if deemed under-used (Navarro, 2009). By contrast, occupants making active use of the land cannot be removed as long as they are growing crops. In China, under the household responsibility system introduced in 1978, land belongs to the community and individual farmers have usufruct rights that can be subject to expropriation. Households engaging in off-farm employment are more likely to see part or all of their land reallocated (Rozelle and Li, 1998). In Ghana, Goldstein and Udry (2008) find that individuals with more secure property rights due to their political position can reduce land use, leaving it idle over longer fallow periods to restore soil fertility.

[2]Typically, distribution of land titles is demand driven. See for example, Alston, Libecap, and Schneider (1996).

changes in migration between households in early certified and later certified ejidos.[3] We establish the migration result using three independent datasets with the following results. First, using panel data on rural households, we find that households in certified ejidos were subsequently 28% more likely to have a migrant household member. Second, using locality level data from two successive population censuses, we find that certification led to a 4% reduction in population. Third, we use a nationwide ejido census to confirm that certification led to more young people leaving the ejido for work reasons. Our estimates imply that about 70,000 people–or some 20% of the total number of migrants from these communities–can be attributed to the certification program.

With this main result established, we proceed to test other predictions of the model. First, we document heterogeneity in migration responses, with larger effects for households with ex-ante weaker property rights (associated with border conflicts and gender of the household head) and with more attractive off-farm wage opportunities. Second, we document that migration effects are smaller where land is more productive, consistent with labor tying being more onerous in less productive land. Third, we find evidence of sorting at the community level regarding who migrates based on differential land endowments. Farmers with more land were less likely to migrate than smaller landholders as a result of the program. The model predicts this differential effect, as the use restriction in the previous property rights regime was more binding for farmers with smaller landholdings. Finally, the model suggests that the difference in migration responses between large and small landholders should be sharper in areas with higher land productivity. We find clear evidence of this in the data. The overall effect of certification for land-rich households in high productivity areas is not statistically different from zero. In contrast, in low land productivity regions the migration effect is statistically significant for large and small landholders and of about the same magnitude.

The focus of the empirical analysis in the first part of the paper is on the labor reallocation effects. The second part of the paper explores the effects of certification on farm consolidation and land use. By allowing consolidation of farm units, the certification program could help resolve the suboptimal farm size problem. Of course, frictions in the land market in spite of certification can also lead to less cultivation if migrants decide to keep the land fallow - but preserve ownership due to its option value or as a retirement activity. We test for this effect using a Herfindahl land concentration index, but cannot reject that there was no consolidation over a four year period, although the coefficient is positive and the magnitude economically significant. Land concentration effects may of course take a longer time to emerge and we only have data on this outcome in a four year window.

The second question regarding land use we focus on is whether the certification program actually led to reductions in cultivated area. Less labor inputs are naturally expected to decrease total output. However, there are two countervailing forces that make this an empirical question. The first is land consolidation in a context of increasing returns to land, while the second is the enhanced investment effect traditionally argued for in the property rights

[3]The robustness checks section provides evidence for the parallel trend assumption necessary for identification.

literature. Investments that are complementary to agricultural land could help expand cultivated area after the program. We use three rounds of satellite land use data to determine that, on average, farmland in ejidos did not decrease after introduction of the program in spite of large population losses. We also find that the impact of certification on land area under cultivation depends on land quality: ejidos in high land productivity areas saw an increase in farmland after the certification program was introduced compared to those in low productivity areas where there was a slight reduction.

An alternative explanation for the increased migration result is that the certification program attracted funds from outside the community through land transactions which helped finance migration by relaxing liquidity constraints.[4] We test and reject that this alternative mechanism is explaining the increased migration after certification. We assess the role of credit constraints by comparing the effect of the certification program between randomly assigned Progresa (a conditional cash transfer program) and non-Progresa localities. Because the former experienced substantial exogenous cash inflows *before* certification, thereby mitigating liquidity constraints, the migration response should be smaller in Progresa localities once certification occurred. We do not find evidence of this in the data.[5]

Our paper relates to a new literature on the effects of property rights on migration in rural areas. In the context of China, a recent working paper by Giles and Mu (2011) shows that tenure insecurity caused by periodic land reallocations, based in part on household land use, has caused farmers to reduce outmigration. Work by Chernina et al. (2013) studies rural to rural migration in the Russian Empire during the early 1900s and argues that increased land liquidity was an important component of the Stolypin titling reform. The authors use a difference-in-differences strategy to show that migration increased significantly after the titling reforms. It is however difficult to attribute the effects of the Stolypin reforms to land liquidity since the reforms occurred concurrently with a large number of government programs designed to incentivize migration to rural Siberia, including giving away land at destination and paying for transportation costs.[6] In a recent paper, Valsecchi (2012) studies the effect of Procede on international migration to the U.S. using a triple differences estimator. However, an important complication arises from his use of posesionarios/avecindados[7] as a non-eligible household control group, since these were often formally recognized as ejidatarios during administration of the program or hired as laborers following the opening of the labor market. Because the program had indirect effects on non-eligible households, this

[4]Angelucci (2012b) shows that conditional cash transfer programs alleviate credit constraints and allow for migration of household members.

[5]Previous research has failed to document a credit access effect from banks using land as collateral after titling (Galiani and Schargrodsky, 2010; Field and Torero, 2006). The Mexican certification program was explicitly designed to limit mortgages (hence the term certification, not title) so we ignore this alternative in the paper. Early evidence on Procede also failed to find any credit access effects (Deininger and Bresciani, 2001).

[6]The authors also show that the main effects of the reform on migration persisted conditional on land sales, suggesting that other mechanisms are potentially contributing to the results.

[7]Avecindados are families living in the ejido without formal access to ejido land. Posesionarios are fringe members of ejidos that had voting rights in ejido assemblies, but did not have formal access to ejido land.

creates identification concerns for triple difference estimates. Land rights and migration in Africa have been studied by de Brauw and Mueller (2012) who show that changes in self-reported perceived transferability of land rights was not significantly correlated with changes in probability of labor out-migration in Ethiopia. That null result of course must be considered as taking place in a context in which the land remains state-owned, and sales, mortgages and land exchanges are still illegal, making it unclear how perceived land transferability can impact migration.

Other work on property rights and labor allocation has focused on urban areas. Field (2007) finds that providing land titles to urban squatters in Peru resulted in an increase in the amount of labor allocated to work away from home, in essence due to a reduction in the need for guarding labor. In contrast, Galiani and Schargrodsky (2010) find that the provision of land titles to squatters in urban Argentina had no effect on labor market outcomes, possibly due to unconstrained labor supply prior to the reform.

Our paper complements this literature by providing theory and empirical analysis suggesting a different explanation for why households may migrate after rural land titling programs. Requirements to use land productively put households in a constrained optimum where too much labor was being used in agriculture, particularly in the least productive areas and on the smaller farms. Our model has clear predictions about what types of families should be most likely to send migrants following reform. The household-level microdata that we use allows us to test these theories. The prediction that some families should send migrants and others should not has implications for the aggregate impacts of the reform. Particularly, sorting according to land productivity suggests that average productivity could indeed increase due to migration rather than to increased investment.

The remainder of the paper is organized as follows. In Section 2 we provide further details on Procede. Section 3 develops a basic household model and derives testable implications. Section 4 discusses the data and the identification strategy. Section 5 presents the results. Section 6 provides robustness checks and section 7 concludes.

1.2 The *Procede* Land Certification Program

During the period from 1914 to 1992, Mexico's first land reform consisted in government expropriation of large private landholdings and redistribution of these tracts of land to groups of peasant farmers organized in agrarian communities called ejidos (Sanderson, 1984).[8] Once awarded, the land was managed by the assembly of farmers under the guiding hand of the state. Beneficiaries received usufruct rights to a land plot for individual cultivation, access to common-use land (for forests, pastures, and surface water), and a residential lot. With the objective of limiting land concentration, ejidatarios faced strict legal restrictions on

[8]The program also certified land in indigenous communities. In the remainder of the paper we do not differentiate ejidos from indigenous communities.

rentals and sales of land.[9] Furthermore, the Constitution itself ruled that any individual land that was not cultivated in two consecutive years was to be reassigned to a member of the community willing and able to cultivate the land, imposing a permanent "use it or lose it" restriction.

Giving access to land with obligation to use it productively has been an important instrument of land redistribution programs. For example, the United States Homestead Act of 1862 and the Reclamation Act of 1902 only awarded title to the landholder after five years of actual and continuous residence in order to guard against "dummy filings, speculation, and the accumulation of large estates" (Coman, 1911). In the Mexican ejido, the use requirement was permanent. Political scientists have argued that granting incomplete property rights with use requirements was purposefully done to create a clientelistic relationship between farmers and the party in power, in spite of the economic inefficiencies it entailed (Magaloni, 2006).[10]

This first land redistribution program, one of the largest in the world (Yates, 1981), eventually resulted in low agricultural productivity and high levels of poverty among beneficiaries (de Janvry, Gordillo, and Sadoulet, 1997). With the impending advent of NAFTA (the free trade agreement between Mexico, the United States, and Canada), the Mexican government introduced a major constitutional reform in 1992 to improve efficiency in the ejido by certifying individual land plots to current users. The reform was clearly intended to improve security of access to land in the ejido by delineating individual property boundaries within the ejido, thus encouraging long-term productive investments by ejidatarios (Heath, 1990). The reform created Agrarian Tribunals to resolve conflicts over the issuance of certificates, established an ejido National Land Registry where individuals would be assigned parcels in the ejido, allowed land rental and sales between ejidatarios, and established a well defined procedure to turn ejido certificates into full titles that could be sold to non-ejidatarios.[11] By issuing land certificates, the program effectively delinked property rights from use requirements.

The program was national in scope and took 13 years to complete. The registration process began with officials from the Agrarian Attorney's Office (PA) approaching ejido officials and providing information about Procede. An ejido assembly was called to approve initiation of the certification process. Except for a few conflict zones, the program progressed remarkably smoothly. After the first assembly, government officials from the National Institute of Statistics and Geography (INEGI) worked with the ejido to identify owners of plots and to produce GIS maps of the ejido. Any disputes over property ownership had to be resolved during this stage of the process by the agrarian courts especially created to resolve such conflicts (Deininger and Bresciani, 2001). After all conflicts had been resolved, the maps showing individual ownership were submitted for approval at a final ejido assembly. Final

[9]Although there is evidence that a black market for ejido lands existed in some parts of the country (Cornelius and Myhre, 1998).

[10]In a recent paper, we find evidence of voting behavior consistent with that hypothesis (de Janvry, Gonzalez-Navarro, and Sadoulet, 2013).

[11]See Appendini (2002) and de Ita (2006) for a description of the reforms.

approval resulted in issuance of ownership certificates by the National Agrarian Registry (RAN) *simultaneously* to all rights-holders in the ejido.

de Janvry, Gonzalez-Navarro, and Sadoulet (2013) investigate the correlates of the Procede rollout, showing that ejidos where the program was initiated earlier were on average smaller, had more land in parcels, were closer to large cities, were wealthier, had fewer posesionarios, and were more likely to be in municipalities that were politically aligned with the party of the state governor. The differences between early and late certified ejidos are not a threat to our identification strategy as long as the differences are largely time invariant or uncorrelated with changes over time in migration. As a first robustness check to address this concern we verify that changes over time in migration *prior to the program* were not correlated with the year of program completion. In our main analysis, we also interact fixed ejido characteristics with time effects to account for the possibility that migration changed over time due to these fixed characteristics that were correlated with timing of land certification.

1.3 Theory

The traditional land insecurity model treats insecurity of property rights as a tax on production. Because improving property rights in the canonical model generates a higher expected output, this naturally leads a household to optimally allocate more labor to the farm, thus reducing the equilibrium level of outmigration. Note that this result is based on the critical assumption that the household is always efficiently allocating labor between uses.

The main innovation in our model is to introduce use requirements as a condition to maintain property rights. In a context of small plot sizes (due to the prohibition of land transactions), this leads to spatial labor misallocation. The model makes clear how these two conditions can cause inefficient tying of labor to land, and how relaxing the use restriction can provoke increased outmigration. Once this is established, the model is used to generate predictions about heterogenous effects which can be taken to the data.

Setup

We use the standard agricultural production model in which farm labor h_e produces expected output Y_e according to $Y_e = \gamma A^\alpha h_e^\beta$, where $0 < \alpha, \beta < 1$, A is land, and γ is a total factor productivity parameter. We incorporate migration as households having the option of supplying labor h_m in the non-farm labor market at the wage w_m, from which they earn $w_m h_m$. Household utility is quasi-linear:

$$u(C, \ell) = C + v(\ell),$$

where C is consumption, ℓ is leisure, and utility of leisure is concave ($v' > 0,\ v'' < 0$). Households are endowed with time T which is spent working on the farm, on wage labor off the farm, and on leisure, so that $T = h_e + h_m + \ell$ is the time constraint. The household's budget constraint is $C = \gamma A^\alpha h_e^\beta + w_m h_m + I$, where I is non-labor income.

Traditional land insecurity model

Insecure property rights are usually modeled as reducing the expected product that the household reaps from farm labor (for instance Besley and Ghatak, 2010). In particular, expected farm production becomes $Y_e = (1-\tau)\gamma A^{\alpha} h_e^{\beta}$, where $\tau \in [0,1]$ reflects the degree of insecurity in property rights.

Obtaining the first order conditions of the household's problem and differentiating with respect to τ provides the following prediction:

$$\frac{\partial h_e}{\partial \tau} = \frac{-h_e}{(1-\tau)(1-\beta)} < 0.$$

Thus, in the standard setup, improving property rights results in an increase in farm labor and a corresponding decrease in migration.

When land use preserves property rights over the land

In line with the nature of property rights in Mexican ejidos, we instead incorporate land insecurity as a required minimum production level per unit of land:

$$\frac{Y_e}{A} \geq \frac{\pi_m}{s},$$

where π_m is the minimum yield, and $s \in (0,1)$ is a parameter representing the household's specific strength of property rights. The parameter s captures the idea that households with weaker property rights have to maintain a higher production level to keep their land (Goldstein and Udry, 2008). Because we do not have stochastic output, the minimum yield requirement can alternatively be thought of as a minimum labor requirement per unit of land. However, in deference to the principal-agent literature, we use the minimum yield requirement as it is more realistic.

In line with use-based ownership, there is neither a rental nor a sales markets for land, and farmers are not allowed to hire workers. Hence A is the exogenously allotted land to the household during the initial phase of land reform, and h_e can only be family labor. Lack of land markets and farm sizes below the optimal scale generate non-decreasing return to scale ($\alpha + \beta \geq 1$). Non-decreasing returns to scale can arise out of small landholdings or production indivisibilities. In any case, there is evidence for this assumption in Mexican ejidos.[12]

[12] The 1994 ejido survey was administered to around 1300 ejido households by the World Bank. We estimated a production function of the form $ln(production_{is}) = \beta_0 + \beta_1 ln(hectares_{is}) + \beta_2 ln(labor_{is}) + \alpha_s + \varepsilon_{is}$, where i indexes households and s indexes states. Standard errors were conservatively clustered at the state level. The estimates from this regression are $\hat{\beta}_1 = 0.933$ and $\hat{\beta}_2 = 0.176$. The sum of the two coefficients is significantly larger than 1 with a p-value of 0.048. While these estimates certainly cannot be interpreted causally, the results provide suggestive empirical evidence consistent with non-decreasing returns to scale in this context. See also Adamopoulos and Restuccia (2013) for estimates of the efficiency cost of small farms in developing countries.

Without constraint, the optimal allocation of labor to farm production would be:

$$h_e^* = \left(\frac{\gamma\beta}{w_m}\right)^{\frac{1}{1-\beta}} A^{\frac{\alpha}{1-\beta}}, \tag{1.1}$$

which is an increasing and convex function of A. The minimum yield constraint requires the household to allocate a minimum amount of labor ($\underline{h_e}$) to agricultural production

$$\underline{h_e} = \left(\frac{\pi_m}{s\gamma}\right)^{\frac{1}{\beta}} A^{\frac{1-\alpha}{\beta}}, \tag{1.2}$$

or else lose its land. This minimum labor requirement is an increasing and concave function of A. The restriction will bind for farm sizes that are smaller than the threshold A_0 defined by $h_e^* = \underline{h_e}$:

$$A_0 = \left[\frac{1}{\gamma}\left(\frac{\pi_m}{s}\right)^{1-\beta}\left(\frac{w_m}{\beta}\right)^{\beta}\right]^{\frac{1}{\alpha+\beta-1}}. \tag{1.3}$$

At the constrained labor allocation, the average on-farm return to labor is:

$$\frac{Y_e}{h_e} = \gamma A^{\alpha} h_e^{\beta-1} = \gamma^{\frac{1}{\beta}}\left(\frac{\pi_m}{s}\right)^{1-\frac{1}{\beta}} A^{\frac{\alpha+\beta-1}{\beta}},$$

When the restriction binds, although households allocate more time to the farm than under unrestricted optimization, it is still advantageous to allocate $\underline{h_e}$ to the farm as long as the average return to farm labor is as large as the off farm wage, i.e., $Y_e/h_e \geq w_m$. This defines a threshold A_1 below which households will prefer to relinquish their land and fully work off-farm:

$$A_1 = \left[\frac{1}{\gamma}\left(\frac{\pi_m}{s}\right)^{1-\beta} {w_m}^{\beta}\right]^{\frac{1}{\alpha+\beta-1}} = \beta^{\frac{\beta}{\alpha+\beta-1}} A_0 \tag{1.4}$$

Equilibrium Labor Allocation: *The labor allocation solution to this restricted optimization is represented in Figure 1.1 and summarized as follows*:

- Leisure is determined by: $w_m = v'(\ell)$
- On farm labor is given by:

  (i) $h_e = h_e^*$, if $A \geq A_0$

  (ii) $h_e = \underline{h_e}$, if $A_1 \leq A \leq A_0$

  (iii) $h_e = 0$, if $A \leq A_1$,

  where A_0 is defined by $h_e^* = \underline{h_e}$, and A_1 is defined by $Y_e/h_e = w_m$

- Migrant/off-farm labor is given by:

$$h_m = T - h_e - \ell \tag{1.5}$$

The results have simple interpretations since land is the key complementary input to farm labor. Households with a sufficiently small land endowment cannot obtain their opportunity cost by staying and cultivating land; they choose to surrender their land and work off-farm. Households with a large land endowment have a high marginal product of labor and are thus unaffected by the production constraint. These households optimally allocate all their labor to agriculture while at the same time producing enough output to keep their land. Only households with intermediate levels of land find themselves allocating more labor than would be optimal under unrestricted optimization.

We argue that in the context of Mexican ejidos one can think of most households as belonging to this intermediate range. First, consider that the objective of the original Mexican land redistribution program was to provide land to as many landless peasants as possible. This gave the government an incentive to minimize plot size subject to providing the household a livelihood (the opportunity cost in the model). Second, because land transactions were not allowed prior to the Procede program, farm sizes were maintained at the originally allocated size without allowing for adjustments in response to the advent of mechanization in agriculture, which is thought to increase the optimal farm size. Third, further evidence of excess labor in ejidos comes from the 1991 agricultural census which indicates that the number of workers per hectare of land in the Mexican private sector (non-ejido) was 0.041, whereas in the ejido sector it was 0.052.

Land certificates and migration

Procede certificates can be interpreted as allowing farmers to move from the restricted optimization situation to the unrestricted situation. If the minimum labor allocation restriction was binding (regime (ii) with $A_1 \leq A \leq A_0$), farm labor *decreases with land certificates*:

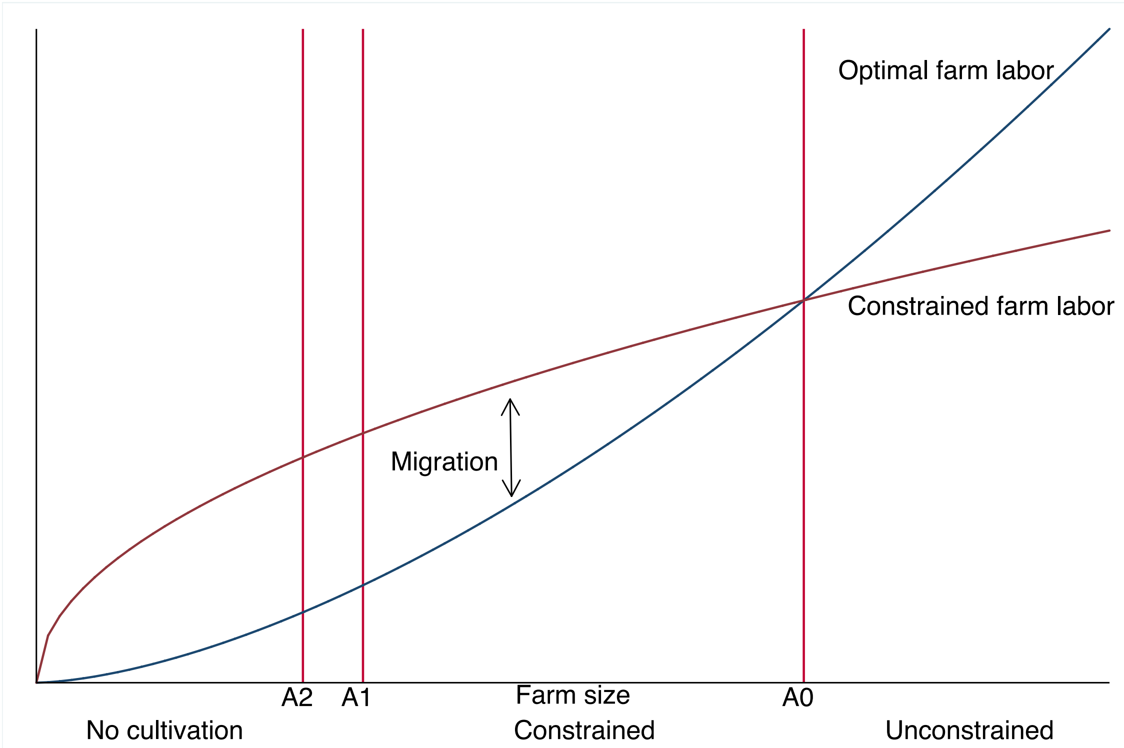
$$\Delta h_e = h_e^* - \underline{h_e}$$

And migrant labor increases by the opposite amount:

$$\Delta h_m = \underline{h_e} - h_e^* = \left(\frac{\pi_m}{s\gamma}\right)^{\frac{1}{\beta}} A^{\frac{1-\alpha}{\beta}} - \left(\frac{\gamma\beta}{w_m}\right)^{\frac{1}{1-\beta}} A^{\frac{\alpha}{1-\beta}}. \tag{1.6}$$

In Figure 1.1, certification is represented by a vertical move from the restricted to the unrestricted on-farm labor schedule. Leisure is unaffected because it is solely determined by the outside wage w_m.

Figure 1.1: Labor allocation to farm production



Notes: Figure shows the optimal agricultural labor schedule as a function of farm size. *Optimal farm labor* is labor use under the title-based regime. *Constrained farm labor* is labor use under the minimum production constraint. See Section 3 for other details on theoretical model.

Heterogeneity in migration response to certification

This simple framework can be used to obtain comparative statics predictions resulting from household level heterogeneity. Note that, while the level of migration h_m of a household depends on family size (equation 1.5), this is not the case for the out-migration Δh_m induced by the land certificate (equation 1.6). Our unitary model also does not generate predictions on which household members should migrate as a response to the program. Ejidos did not have rules on which household members should cultivate land and therefore any household member could be used to meet minimum production requirements.[13] The predicted migration response however varies with strength of the use-based property rights previously enjoyed, outside wages, farm size, and land productivity. All comparative statics results are obtained by simple differentiation of equation (1.6).

[13]In results not reported here we find marginally significant heterogeneity in program effects according to the number of young males in the household at baseline. These effects could be interpreted as either related to household size (T) or to greater potential off-farm wages (w_m).

Degree of security under use-based property rights
Heterogeneity in the degree of land insecurity under the use-based regime can be thought of as heterogeneity in the s parameter. More insecure property rights are reflected as a lower s and a higher required farm activity $\underline{h_e}$. Differentiating (1.6) with respect to s:

$$\frac{\partial \Delta h_m}{\partial s} = \frac{\partial \underline{h_e}}{\partial s} < 0$$

shows that, *ceteris paribus*, this generates a higher migration response the more insecure property rights were in the use-based regime.

Off-farm wages
Higher wages commanded higher levels of migration h_m through lower optimal leisure. They also induce a higher migration response to the land certificate:

$$\frac{\partial \Delta h_m}{\partial w_m} = -\frac{\partial h_e^*}{\partial w_m} > 0$$

Because the unrestricted on-farm labor schedule is lower the more attractive outside opportunities (w_m) are, the regime change leads to larger migration responses from households with better off-farm opportunities.

Land productivity
Differing farmland quality in the model can be understood as heterogeneity in the productivity parameter γ. Higher land quality reduces the minimum labor necessary to reach the required yield under use-based rights and increases the optimal labor that the household should allocate to the farm. Both effects contribute to a reduction in the excess labor imposed by use-based property rights:

$$\frac{\partial \Delta h_m}{\partial \gamma} = \frac{\partial \underline{h_e}}{\partial \gamma} - \frac{\partial h_e^*}{\partial \gamma} < 0$$

This suggests that farms with lower land productivity have more outmigration when moving from a use-based to a title-based property rights regime.

Farm size
Differentiation of (1.6) with respect to A gives:

$$\frac{\partial \Delta h_m}{\partial A} = \frac{\partial \underline{h_e}}{\partial A} - \frac{\partial h_e^*}{\partial A} = \left(\frac{\pi_m}{s\gamma}\right)^{\frac{1}{\beta}} \frac{1-\alpha}{\beta} A^{\frac{1-\alpha-\beta}{\beta}} - \left(\frac{\gamma\beta}{w_m}\right)^{\frac{1}{1-\beta}} \frac{\alpha}{1-\beta} A^{\frac{\alpha+\beta-1}{1-\beta}}$$

This expression can be shown to be negative for land size A greater than a threshold A_2 where the two curves $\underline{h_e}$ and h_e^* have parallel slopes.

$$A_2 = A_1 \left[\frac{(1-\alpha)(1-\beta)}{\alpha\beta} \beta^{\frac{-1}{1-\beta}}\right]^{\frac{\beta(1-\beta)}{\alpha+\beta-1}}$$

The first term in the square brackets is smaller than 1, while the second term is greater than 1, meaning that A_2 can either be greater or smaller than A_1. Hence, migration induced by relaxing the yield constraint decreases with farm size, except possibly for the smallest farms still operating with $A \in [A_1, A_2]$, if it is the case that $A_1 < A_2$. The case where $A_2 < A_1$ is depicted in Figure 1.1. In this case the vertical distance between the two curves is clearly decreasing in A. This expression suggests that if there is heterogeneity in land holding size (A) within ejidos, the larger landholders should migrate less in response to certification. This can be thought of as a sorting effect in which the larger farmers are more likely to stay behind while the smaller more marginal farmers migrate.

It is also straightforward to see that this expression implies that the differential induced migration across farm sizes is sharper in areas with higher land quality:

$$\frac{\partial^2 \Delta h_m}{\partial \gamma \partial A} < 0.$$

This prediction is economically important. It can be interpreted as saying that the migration response of larger landholders in high productivity areas is lower than the migration response of larger landholders in low productivity areas. An equivalent interpretation is that in low productivity areas, the difference in migration response between small and large landholders is not as different as that which arises in high productivity ones.

In summary, we expect that delinking property rights from use requirements allows households to allocate the optimal amount of labor to their farm activity instead of the inefficiently high level required by the "use-it or lose-it" restriction. The out-migration response is expected to be larger for households that had weaker property rights under the prior regime of incomplete property rights, that have better outside opportunities, smaller farms, and lower land quality. We also expect that the differential migration response between small and large farms is stronger in areas with better land. These are the results taken to the data in section 1.5.

Before moving to the data, we should explicitly acknowledge that the model focuses on the use constraint and its effects in an environment of small landholdings (increasing returns to scale). In doing so, it leaves out many other factors that, while relevant, cannot explain the migration responses we are interested in. The first factor left out in this model is land consolidation. Procede can be expected to allow farmers to consolidate their operations (by rentals or sales of land) in order to achieve a more efficient scale. We test for this effect when we explore land use outcomes, but note that it cannot explain increased outmigration. By increasing the productivity of labor, land consolidation works to increase labor demand. The second factor we leave out is labor markets. If Procede allowed for more efficient labor markets, we expect moving towards a separation equilibrium (as in Benjamin, 1992). We test for this in the results section, but note that it would not explain increased outmigration either.

Finally, the view we take in this model is that credit constraints were not restricting migration. Some have argued that the existence of wage differentials between urban and rural areas may be explained by credit constraints to migration (Levy and Van Wijnbergen, 1995), and this is a definite alternative mechanism. By allowing land transactions to take place, certification could have alleviated credit constraints and allowed for more migration. We investigate this alternative mechanism empirically, but fail to find evidence for it.

1.4 Data

Our source of information on the rollout of Procede is a set of ejido digital maps created during the certification process by INEGI and managed by RAN. GIS ejido boundaries are available for the 26,481 ejidos that completed the program during the period from 1993-2006.[14] The rollout of the program was quite rapid. Nearly half of all ejidos were fully certified by 1997 while all but a small subset of ejidos had completed the program by 2006. The curve in Figure 1.2 gives the share of these ejidos that had completed the program each year from 1993 to 2006. Figure 1.2 also shows the dates of the other datasets used: the Progresa surveys (ENCEL), the population censuses, the ejido censuses, and the land use maps. Figure 1.3 in the online appendix maps the rollout of Procede at the national level, helping visualize the extensiveness and national scope of the program.

We use the 1998-2000 Encuesta Evaluacion de los Hogares (ENCEL) surveys administered in the evaluation of the anti-poverty program Progresa to study individual migration behavior.[15] The ENCEL data consist of a panel of approximately 25,000 households from 506 poor localities that qualified for the program in the states of Guerrero, Hidalgo, Michoacan, Puebla, Queretaro, San Luis Potosi, and Veracruz. We matched the localities to ejidos using the coordinates of the centroid of the locality. We considered the locality to match an ejido if the centroid of the locality was located inside the boundaries of one of the ejidos in the GIS database. This process matched 200 localities to 195 different ejidos. Of these ejidos, 68 were certified in 1993-1996, 51 in 1997-1999, and 76 after 1999. Our final data consist of an unbalanced panel of 7,577 households from ejidos that were certified after 1996.[16] Approximately 2.2% of these households had a migrant leave during 1997. Between 1998 and 2000 an additional 5.9% of households sent a migrant.

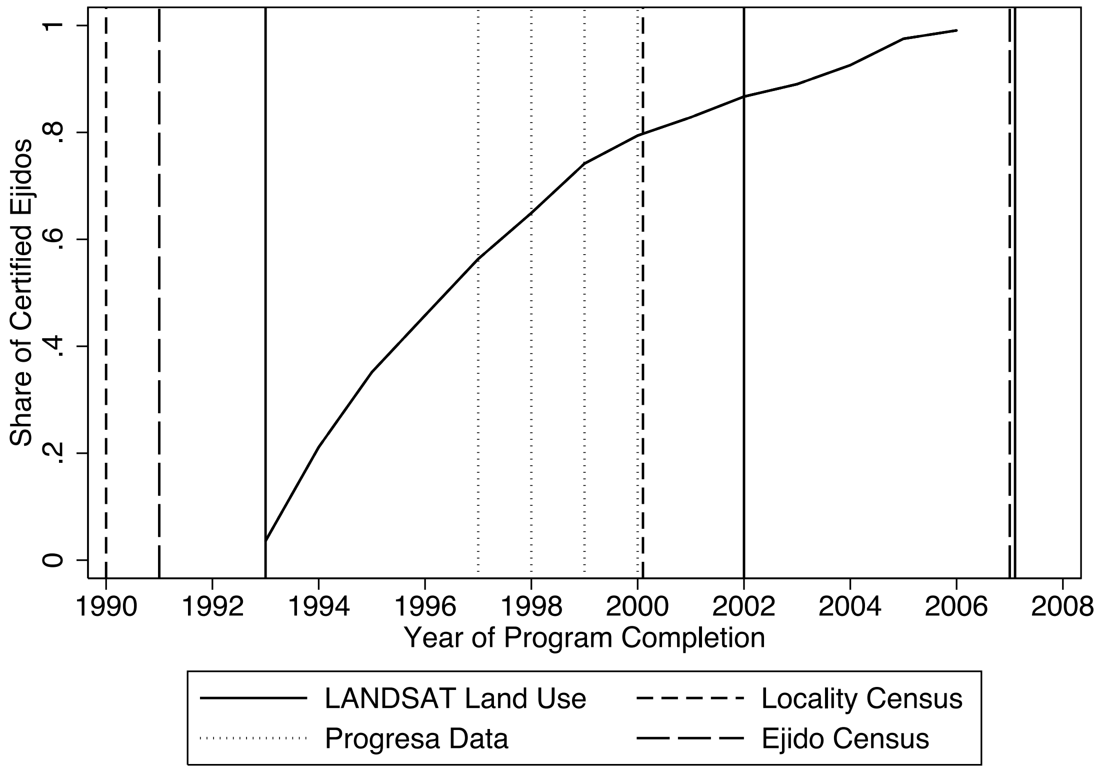
For the community level analysis, we use the 1990 and 2000 population censuses at the locality level from INEGI. Figure 1.2 shows that approximately 75% of ejidos completed

[14]These data also include 246 ejidos that were in the process of certification but had not yet completed the program during 2007. They do not include the remaining 2500 ejidos that were left to a special program after Procede closed in 2006.

[15]Progresa is the Mexican conditional cash transfer program started in 1997. The program is now referred to as Oportunidades. Progresa localities were selected to have more than 50 but less than 2,500 inhabitants and have a high marginality index as computed from the 1990 population census and the 1995 population count information. We use the 1998, 1999, and 2000 ENCEL surveys. The 1997 migration data were derived from recalls in the 1998 ENCEL survey. The 1997 ENCASEH baseline survey did not have comparable migration information.

[16]The panel is unbalanced due to attrition as well as addition of a small number of households to the sample in 1999 and 2000. Our migration result is robust to estimation with a fully balanced panel of households.

Figure 1.2: Correspondence between data and rollout of Procede



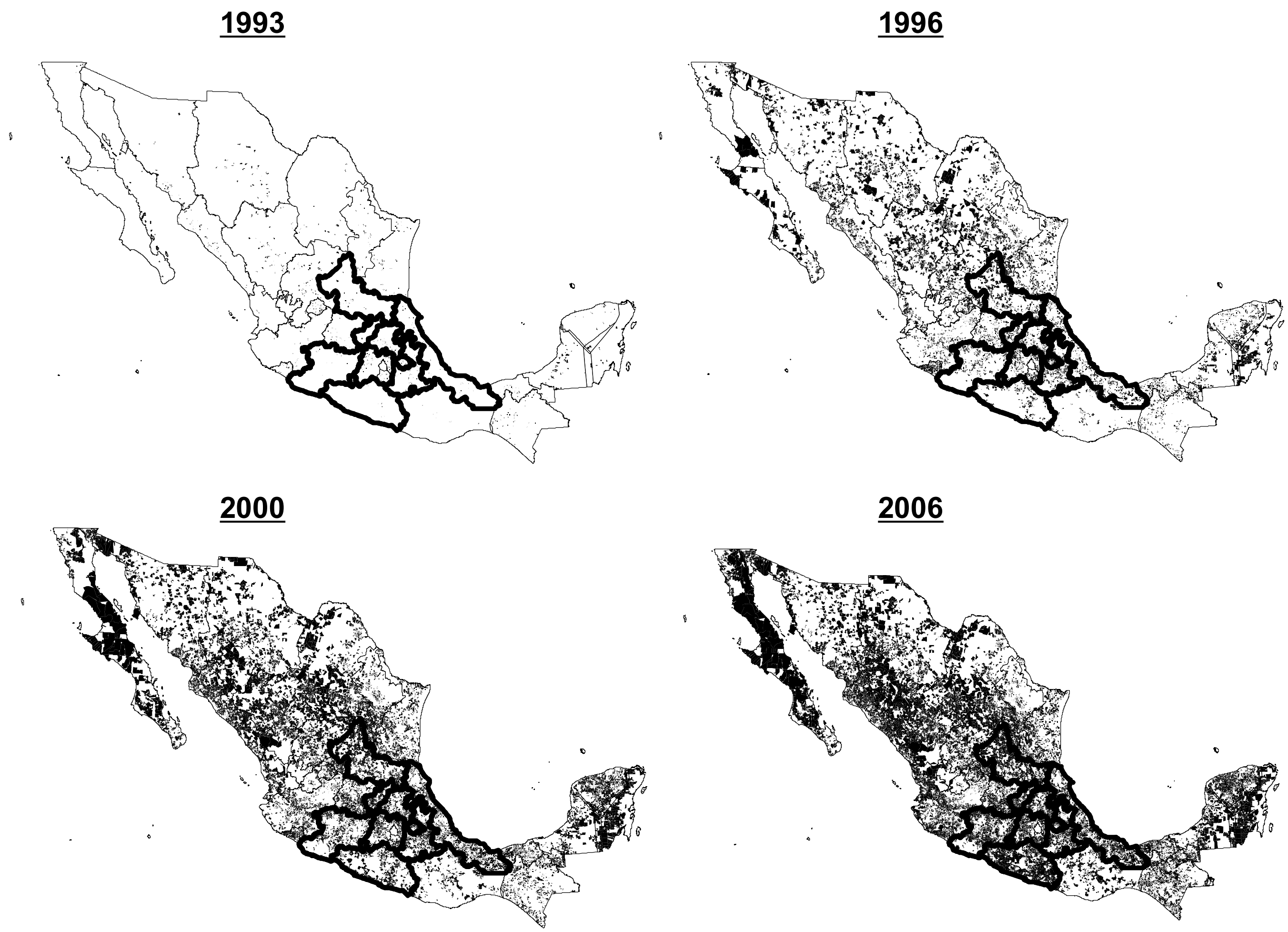
Notes: Figure shows cumulative share of ejidos certified over time. Vertical lines represent observations for each of the datasets used in the analysis. The Progresa ENCEL data are from 1998-2000. Migration recall data were used for 1997. Locality level census data are from 1990 and 2000. Ejido level census data are from 1991 and 2007. LANDSAT land use data are from 1993, 2002, and 2007.

the program between the two censuses. We matched locality centroids to ejidos using the spatial matching technique mentioned above. The final data used in the regressions is a balanced two year panel of population and certification status for 17,328 localities.[17] These data cover all states of Mexico and therefore have broader geographic coverage than the panel of Progresa households. Approximately 62% of the localities in ejidos experienced a decline in population during the period from 1990-2000.

The fourth dataset we use is the Ejido Census (Censo Ejidal) from INEGI that was administered to all ejidos in Mexico in the years 1991 and 2007. The 1991 and 2007 matched surveys are not publicly available and were merged by INEGI specifically for this study. Because the census data that were made available to us did not identify the ejido by name,

[17]All regressions at the community level exclude localities that had population of 20 or less individuals in 1990. Small localities often disappear or are regrouped over time and we therefore drop them from the analysis.

Figure 1.3: Rollout of Procede across time and space



Notes: Shaded ejidos are those that completed the Procede program during or before the listed year. States with bold outlines are 7 Progresa states for which we have migration outcomes (Guerrero, Hidalgo, Michoacan, Puebla, Queretaro, San Luis Potosi, and Veracruz).

we created a matching algorithm that builds on common variables in the two censuses and the ejido GIS maps to construct a matched dataset of 19,713 ejidos. The details of the matching algorithm are given in the online appendix.

Finally, we use INEGI GIS land use maps for the whole country. The data consist of Series II, III, and IV of the INEGI land use/land cover maps. The data are based on a combination of Landsat imagery taken during 1993, 2002, and 2007 and a series of field verifications by INEGI. The digital ejido boundaries were overlaid on the land use maps to create a panel of land use at the ejido level for the years 1993, 2002, and 2007. The median amount of agricultural land during 1993 in the ejidos certified in 1993-2006 is roughly 240 hectares, while the median share of total ejido area that is in agriculture is 27%. These figures rose slightly to 275 hectares and 32% in 2007.

1.5 Results

The impact of land certification on migration

We establish our basic result that rural land certification leads to increased outmigration in three independent datasets. First, we consider the panel of households from Progresa, which contains detailed demographic variables and migration status of household members over the four years 1997-2000. The unit of analysis is the household and the dependent variable is an indicator for whether the household has a permanent migrant that left the ejido since the onset of our observations. The main estimating equation is:

$$y_{ijt} = \delta Certif_{jt} + \gamma_j + \alpha_t + x_{ijt}\beta + \varepsilon_{ijt}, \tag{1.7}$$

where y_{ijt} is an indicator for whether household i in ejido j has a permanent migrant by year t, $Certif_{jt}$ is an indicator for whether ejido j was certified at the beginning of year t, γ_j is an ejido fixed effect, α_t is a time fixed effect, x_{ijt} is a vector of household level covariates, and ε_{ijt} is a random error term. Standard errors are clustered at the ejido level for estimation. This is a standard fixed effects regression where identification is coming from changes in migration behavior correlated to changes in certification status. Any time-invariant ejido characteristic that is correlated with the program rollout is accounted for by the ejido fixed effects. The identifying assumption is therefore that any time-varying ejido characteristic that affects migration trends is uncorrelated with the distribution of certificates. We provide support for the validity of this identification assumption in the next section, focusing first on the results.

First, we use the Progresa dataset to show that land certification led to increased migration of individual household members. In the first column of Table 1.1, the probability of a household having a migrant increases by 0.015 after being reached by Procede. The average rate of migration during the sample period is 5.3%, indicating that the effect of the program was to increase permanent migration by 28%.

The result is not sensitive to a variety of robustness checks. The second column shows that the estimated program effect is almost identical when household level covariates are

Table 1.1: Effect of Procede on household migration behavior

	Progresa Households Matched to Ejidos					
	(1) Has Migrant	(2) Has Migrant	(3) Has Migrant	(4) Has Migrant	(5) Has Migrant	(6) Has Migrant
Certified	0.0149**	0.0147**	0.0153**	0.0172***	0.0157**	0.0130**
	(0.0061)	(0.0066)	(0.0062)	(0.0059)	(0.0063)	(0.0062)
HH is Landholder		0.0136***			0.0048	
		(0.0046)			(0.0053)	
Number Males 17-30 in HH		0.0185***			0.0088**	
		(0.0046)			(0.0037)	
HH Head is Female		0.0132			0.0092	
		(0.0106)			(0.0082)	
Age of HH Head		0.0009***			0.0004***	
		(0.0002)			(0.0001)	
Agricultural value (100 USD/ha)		0.0227				
		(0.0230)				
Time Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Ejido Fixed Effects	Yes	Yes	No	Yes	Yes	Yes
HH Fixed Effects	No	No	Yes	No	No	No
State x Time Effects	No	No	No	Yes	No	No
HH Characteristics x Time Effects	No	No	No	No	Yes	No
Ejido Characteristics x Time Effects	No	No	No	No	No	Yes
Mean of Dep Variable	0.053	0.055	0.053	0.053	0.056	0.053
Number of Observations	27189	23421	27189	27189	24533	27189
R squared	0.047	0.058	0.043	0.048	0.059	0.048

Standard errors that allow for clustering at the ejido level are reported in parentheses. Asterisks indicate statistical significance at the 1% ***, 5% **, and 10% * levels. Data include observations on all households in ejidos that completed the Procede process after 1996. All regressions are linear probability models. The dependent variable is 1 if the household had a migrant leave during the year or any previous sample year. Certified indicator = 1 if ejido was certified at the start of the year. Ejido characteristics in column 6 are distance to nearest large city (population > 100,000), number of ejidatarios, number of posesionarios + avecindados, total size of the ejido, share of ejido land in parcels, locality marginalization index, longitude, and latitude.

included in the regression. This minimal change is consistent with the fact that certificates were distributed to all ejidatario households in the ejido. Importantly, the regression in column 2 also controls for an ejido-level *time-varying* measure of the value of agricultural productivity. One concern with our identification is that the opening of the Mexican economy due to NAFTA may confound our estimate. In particular, our estimate could be confounded by NAFTA if ejidos were affected differentially over time in a way that was correlated with the rollout of the land certification program. Since the influence of NAFTA on ejidos would operate through agricultural prices, we use a measure of potential agricultural revenue per hectare that proxies for the impact of prices on each ejido.[18] The limited change in our main estimate when controlling for this measure of potential agricultural value suggests that NAFTA is not a confounding factor.

The third column shows that the estimated coefficient is robust to replacing ejido fixed effects by household fixed effects. A key concern for our identification strategy is the possibility of differential time trends that would be correlated with the timing of certification. In columns (4)-(6) we show that the results are robust to controlling for specific time trends more flexibly. In column (4) we allow the time effects to be specific by state. Column (5) includes interaction terms between each time effect and the household-level covariates. In column (6) we include interactions between time effects and some ejido-level characteristics that are shown in de Janvry, Gonzalez-Navarro, and Sadoulet (2013) to be correlated with the rollout of Procede. The purpose of this robustness check is to control for the possibility that the program was initiated earlier in certain types of ejidos that experienced differential changes in migration after the program due to reasons other than land certification. For example, the program was completed on average earlier in ejidos that are located closer to large cities. The fixed effects in our specification obviously account for time invariant differences due to proximity to major cities. Allowing the time effects to depend on proximity to cities further controls for differences in migration over time that are due to earlier program areas being closer to cities rather than certification. Our main result remains economically large and statistically significant after introducing several additional controls for differential time trends. Overall, the behavior of households in the Progresa dataset firmly points to land certificates increasing the probability that a household member migrates.

Second, we study migration behavior at the locality level using the matched 1990 and 2000 population censuses. The locality level analysis captures both migration of individuals and entire families. Three key characteristics of this alternate dataset are its inclusion of localities of all sizes and levels of income, its geographical coverage (nationwide), and its longer time span (up to 7 years with a certificate). By the year 2000, 73% of the ejidos had been awarded a certificate, while the other ejidos were still in the pre-certification regime.

[18]For each ejido, we assume a fixed allocation of land to crops according the observed allocation in 1995. The crop choices of individual farmers from the farm support program PROCAMPO were used to calculate crop shares for each ejido. We then calculate the weighted average value of a hectare of farm land as $value_{it} = \sum_{k=1}^{K} price_{kt} * yield_{k,1995} * share_{ik,1995}$, where $price_{kt}$ is the price of crop k in year t, $yield_{k,1995}$ is the nationwide yield of crop k in 1995, and $share_{ik,1995}$ is the share of the crop land in ejido i that was cultivated to crop k in 1995.

We first compare the evolution of locality population over time in a standard two-period fixed effects regression:

$$Pop_{jt} = \gamma_j + \beta I(t = 2000) + \delta I(Certified\ by\ 2000_j = 1)I(t = 2000) + \varepsilon_{jt}. \tag{1.8}$$

We then allow for a linear effect of certification over time by estimating:

$$\begin{aligned} Pop_{jt} = {} & \gamma_j + \beta I(t = 2000) + (\delta_0 + \delta_1 Years\ Certified_j)I(Certified\ by\ 2000_j = 1)I(t = 2000) \\ & + \varepsilon_{jt}. \end{aligned} \tag{1.9}$$

We finally partition the ejidos certified between the two censuses into early certified and late certified groups and estimate separate effects for the two groups:

$$\begin{aligned} Pop_{jt} = {} & \gamma_j + \beta I(t = 2000) + \delta_1 I(Certified\ before\ 1997_j = 1)I(t = 2000) + \\ & \delta_2 I(Certified\ from\ 1997 - 1999_j = 1)I(t = 2000) + \varepsilon_{jt}. \end{aligned} \tag{1.10}$$

The dependent variable is the total population (or logarithm) of locality j in year t (1990 or 2000). The first specification (1.8) is a simple fixed effect regression where δ identifies the average effect of the ejido getting certification on the change in locality population. The second specification (1.9) takes into account the number of years since certification, allowing the migration response to take effect over several years in a linear way. The third specification (1.10) estimates a separate certification effect for localities in ejidos certified in 1993-1996 (δ_1) and localities in ejidos certified in 1997-1999 (δ_2).

Regression results are reported in Table 1.2, where standard errors are clustered at the ejido level. The first row in the table shows that ejido localities lost around 9.6 persons or 21% of their population between 1990 and 2000 (the time effect in the first row). The coefficients on the interaction term in the second row indicate that Procede was associated with an *additional reduction* in population of approximately 3-4 individuals, in a setting where the average locality has 99 individuals (column (1)), or 4% of its population (column (2)). Similar to Table 1.1, column (3) shows that our estimate is not meaningfully affected when controlling for the effects of agricultural prices.

While results are less statistically precise, column (4) suggests that the loss of population is progressive over time, with a decline of approximately 0.54% of the population per year after Procede certification. In column (5) we estimate separate effects for early certified ejidos (before 1997) and late certified ejidos (1997-1999). The estimated effect of certification is a 5.9% decrease in population for early certified ejidos and a 2% decrease for later certified ejidos. The difference between early and late certified ejidos is statistically significant. The large difference is consistent with certification leading to initial migration and further migration after migrant networks have been established in destination communities, as in Munshi (2003) who shows that migration networks take approximately 3-4 years to develop. As a specification check we use 12,455 localities with available population in 1980 to estimate a version of (1.8) for the period 1980-1990. The estimate in column (6) indicates

Table 1.2: Effect of Procede on locality-level population

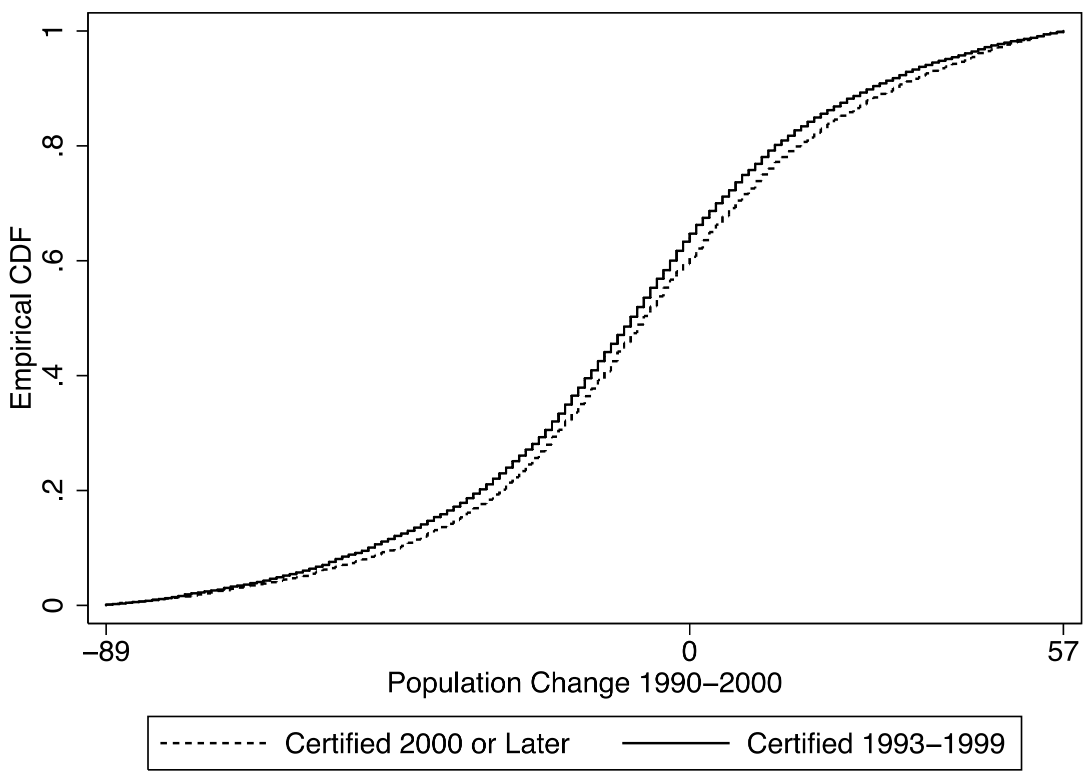
	Census Localities Matched to Ejidos					
	(1) Population	(2) ln(Population)	(3) ln(Population)	(4) ln(Population)	(5) ln(Population)	(6) ln(Population), 1980-1990
Year=2000	-9.6309***	-0.2069***	-0.1986***	-0.2069***	-0.2069***	
	(1.0014)	(0.0105)	(0.0184)	(0.0105)	(0.0105)	
Certified 1993-1999*Year=2000	-3.6893***	-0.0404***	-0.0341**	-0.0206		
	(1.1485)	(0.0128)	(0.0167)	(0.0195)		
Agricultural value (100 USD/ha)			0.0036			
			(0.0077)			
Years Certified in 2000*Certified 1993-1999*Year=2000				-0.0054		
				(0.0039)		
Certified Before 1997*Year=2000					-0.0592***	
					(0.0144)	
Certified 1997-1999*Year=2000					-0.0196	
					(0.0151)	
Year=1990						-0.2094***
						(0.0125)
Certified 1993-1999*Year=1990						-0.0082
						(0.0148)
Ejido Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Mean of Dep Variable	99.111	4.271	4.277	4.271	4.271	4.416
Number of Observations	34656	34656	24170	34656	34656	24910
R squared	0.014	0.035	0.035	0.036	0.036	0.033

Standard errors that allow for clustering at the ejido level are reported in parentheses. Asterisks indicate statistical significance at the 1% ***,5% **, and 10% * levels. Regressions in columns 1-4 based on 17,328 localities that were matched to ejidos, had population data in both the 1990 and 2000 censuses, and had a population of more than 20 individuals in 1990. Regression in column 5 is based on 12,455 localities with available population data in 1980 and with a population larger than 20 in 1980.

that the difference in population change in the 1980-1990 decade between early and late certified localities was very small and not significant. This similarity in pre-program population trends suggests that our estimate is not driven by pre-1990 differences in population change between early program and late program areas.

Ubiquity of the emigration effect across the whole distribution of change in population is illustrated in Figure 1.4. The solid black line represents the empirical distribution function for the change in population from 1990 to 2000 for localities in ejidos that were certified between the two censuses. The dashed line represents localities in ejidos certified in 2000 or later.[19] The distribution for localities in ejidos not certified by 2000 stochastically dominates that for certified localities. This indicates that the effect of certification on migration is not a feature of some specific localities but occurs throughout the distribution of population changes.

Figure 1.4: Cumulative distribution of population change, 1990-2000, by certification date



Notes: Figure displays empirical CDF of population change (in levels) from 1990-2000. Data used are the 1990 and 2000 locality-level population censuses.

How does this estimated effect of Procede on the locality population compare to what

[19] The top and bottom 5% of observations were removed for the graph.

was revealed in the selected Progresa communities? We cannot simply directly compare effects between datasets because the time periods differ. We also must be careful to measure migration effects annually, rather than over a period of several years. The Progresa data document annual emigration from 1997 to 2000, in localities that were certified from 1997 onwards. The most direct comparison can thus be drawn with column (4) of Table 1.2 where we also estimate the program effect during this time period. The time effect shows a baseline migration of 20.7% of the population over 10 years, which corresponds to an average annual rate of 2.3% (=$0.793^{0.1}$-1). The certification effect for those ejidos certified in 1997-99 is an additional effect of 1.96% over these 3 years, or an average annual effect of 0.7%. Hence Procede led to an increase of the annual loss of population of 29% (=0.7/2.3). Recall that the average annual effect in the Progresa dataset was an increase in migration by 28%. So while we looked at different measures of migration in the two datasets (households sending off one permanent migrant in the Progresa dataset and population change in the locality dataset), we find that Procede has had the same relative effect of increasing migration by an additional 28-29%.

Third, we analyze migration behavior using the 1991 and 2007 ejido censuses. By 2007, all the ejidos in our dataset had been certified. Hence we can only identify the effect of certification coming from the differential number of years an ejido has been certified in 2007. Furthermore, because the migration question was not asked in the first round, we can only perform a cross sectional regression. Our dependent variable is the response to a question from the 2007 census asking if the majority of young people leave the ejido. We estimate a cross-sectional regression of the form:

$$Y_{js} = \alpha + \gamma_s + \delta \, \mathit{Years\ Certified}_{js} + x_{js}\beta + \varepsilon_{js}. \tag{1.11}$$

where γ_s are state fixed effects and x_{js} is a vector of ejido level covariates in 1991 (before Procede). The dependent variable Y_{js} is an indicator variable for whether the majority of young people are said to emigrate from the ejido.

This is obviously a less well identified regression than those reported using the previous two datasets. However, this specification is justified by the result in Table 1.2 suggesting that the effect of certification is increasing over time. Second, the ejido census has the advantage that the unit of observation coincides perfectly with the population of interest, because questions are asked about the group of ejidatarios in each particular ejido. Finally, this is the only dataset we use that does not necessitate a geographical merge. Hence, we see this as an important verification of the results presented in the previous two tables.

Results are reported in Table 1.3. Column (1) shows a positive association between the years since certification and the probability that the majority of young people migrate from the ejido. Certified ejidos are 0.35% more likely to respond that a majority of their young people emigrate from the ejido for every year since certification. This result is robust to the addition of ejido covariates measured in 1991 (column (2)). Columns (3) and (4) suggest that most of this effect is driven by increased migration to the United States. The average ejido had been certified 9.5 years in 2007, meaning that for the average ejido, the probability that a majority of young people would be leaving the ejido increased by 7.8 percentage points due to the Procede program.

By presenting results from three independent datasets, we seek to credibly establish that delinking property rights from use requirements generated by the assignment of land certificates led to increased migration from agrarian communities. The number of households having a migrant increased by 28%, the locality population declined by 4%, and ejidos were 0.35% more likely to report that a majority of their youth were leaving the community for every year they had been certified.

Applying these migration effects to the 1.7 million population of the localities matched to ejidos (17,328 localities with average population of 99.1 as reported in Table 1.2 column(1)) suggests that Procede would have been responsible for an outmigration of about 4% of them or almost 70,000 people. This should be compared to the natural trend of a loss of 20.7% or 350,000 people in these communities over 10 years.

These results should not be interpreted as suggesting a reduction in welfare. On the contrary, as the model suggests, we interpret this as evidence that inefficient amounts of labor had been allocated to the land under the use-based property rights regime. By delinking property rights from use, the program merely allowed households to adjust from an inefficient equilibrium with too much farm labor to an efficient equilibrium with less farm labor.

Heterogeneity in pre-reform property rights security

The model predicts that the migration response to land certification should be larger when pre-reform property rights were weaker ($\frac{\partial \Delta h_m}{\partial s} < 0$). As a measure of *between ejido* security, we use a question from the 1991 ejido census on the presence of boundary problems within the ejido. Column (1) of Table 1.4 shows that the point estimate of the migration effect of certification is more than double for households in ejidos where boundary problems were present. A concern with this specification is that migration could increase over time in ejidos with boundary problems independent of certification. We control for differential time effects in column (2). The difference between ejidos with and without boundary issues becomes larger with the addition of specific time effects. The effect of certification on the probability of having a migrant household member increases from 0.008 for households in ejidos without boundary problems to 0.036 for households in ejidos with problems. This difference is significant at the 10% level.

Next, as a measure of *within ejido* insecurity, we use an indicator for a female headed household. Work by social observers indicates that, prior to Procede, female ejidatarias held low status inside the ejido (Stephen, 1996; Deere and Léon, 2001; Hamilton, 2002). For example Stephen (1996, p.291) quotes an ejidataria from Oaxaca as stating, "Women don't participate in ejido assemblies. The men in our community don't let us participate in meetings." Based on interviews conducted in four ejidos in northern and central Mexico, Hamilton (2002) points out that women were susceptible to expropriation by male relatives or friends of high-level ejido officials. This anecdotal evidence prompted the use of a female-headed household dummy as a proxy for weaker ex-ante property rights. We must however

Table 1.3: Effect of Procede on ejido-level migration of young people

	Matched Ejidos in 1991 and 2007 Ejido Census			
	(1) Migrate	(2) Migrate	(3) Migrate US	(4) Migrate US
Years Certified in 2007	0.0035***	0.0039***	0.0037***	0.0031***
	(0.0013)	(0.0013)	(0.0012)	(0.0012)
Using Improved Seeds in 1991		-0.0178*		0.0009
		(0.0100)		(0.0095)
Using Tractors in 1991		-0.0048		0.0123
		(0.0105)		(0.0104)
Electrical Lighting in 1991		0.0384***		0.0514***
		(0.0108)		(0.0110)
Log of Distance Between Ejido and PA Office		0.0528***		0.0110
		(0.0113)		(0.0113)
State Fixed Effects	Yes	Yes	Yes	Yes
Mean of Dep Variable	0.426	0.426	0.297	0.297
Number of Observations	19670	19600	19670	19600
R squared	0.086	0.092	0.128	0.131

Standard errors that allow for clustering at the municipality level are reported in parentheses. Asterisks indicate statistical significance at the 1% ***, 5% **, and 10% * levels. The question in the 2007 census identifies the ejidos where a majority of young people are integrated in the activities of the ejido or remain in the ejido but work in nearby localities. If neither of the prior conditions was true, the destination of the majority of the young people is identified. The variable "migrate" takes on a value of 1 if neither of the first two conditions was true. The dependent variable in column 3 and 4 takes on a value of 1 only if the answer to the location of the majority of young people was the United States.

Table 1.4: Heterogeneous effects of certification on migration

	Progresa Households Matched to Ejidos					
	(1) Has Migrant	(2) Has Migrant	(3) Has Migrant	(4) Has Migrant	(5) Has Migrant	(6) Has Migrant
Certified	0.0138	0.0081	0.0102	0.0096	-0.0028	-0.0041
	(0.0088)	(0.0086)	(0.0063)	(0.0064)	(0.0085)	(0.0093)
Certified*Ejido Had Boundary Problems in 1991	0.0164	0.0278*				
	(0.0142)	(0.0145)				
Certified*HH Head is Female			0.0596**	0.0648**	0.0649**	0.0702**
			(0.0247)	(0.0277)	(0.0252)	(0.0282)
Certified*Above Median Predicted Wage					0.0247*	0.0262*
					(0.0139)	(0.0155)
Above Median Predicted Wage					0.0015	0.0010
					(0.0056)	(0.0056)
Time Effects	Yes	Yes	Yes	Yes	Yes	Yes
Ejido Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Time Effects*HH Head is Female	No	No	No	Yes	No	Yes
Time Effects*Ejido Had Boundary Problems in 1991	No	Yes	No	No	No	No
Time Effects*Above Median Predicted Wage	No	No	No	No	No	Yes
Mean of Dep Variable	0.057	0.057	0.056	0.056	0.056	0.056
Number of Observations	21090	21090	24533	24533	24513	24513
R squared	0.060	0.060	0.059	0.059	0.059	0.060

Standard errors that allow for clustering at the ejido level are reported in parentheses. Asterisks indicate statistical significance at the 1% ***, 5% **, and 10% * levels. Data include observations on all households in ejidos that completed the Procede process after 1996. All regressions are linear probability models. Dependent variable = 1 if the household had a migrant leave during the year or any previous sample year. Certified indicator = 1 if ejido was certified at the start of the year. All regressions include landholder indicator, age of household head, indicator for female household head, and number of males between 17 and 30 as controls. *Ejido had boundary problems* is an indicator variable for response of yes to this question during the 1991 ejido survey. *Above Median Predicted Wage* = 1 if the household's predicted maximum off-farm wage is above the median in the sample.

interpret our result with caution since female headed households are almost certainly different for reasons other than s in our model.

Columns (3) and (4) show that indeed the effect of certification on migration of household members is significantly larger for female headed households. The magnitude of the coefficient is quite large. The subset of households with female heads is small but not trivial, consisting of around 10% of the population. The marginal effect of certification for these households represents an approximate doubling in the probability that the household has a migrant (marginal effect of Procede of 0.065 compared to the mean value of 0.056). These effects contrast with the smaller impact for male-headed households.

These results are consistent with improvements in property rights brought about by land certificates having much larger effects for households with weaker rights prior to certification. In terms of the model, we interpret this as individuals with weaker property rights (lower s) being more constrained prior to the program and thus having to dedicate more labor to the farm to maintain their land. Hence, receipt of land certificates resulted in a larger migration response for these households.[20]

Heterogeneity in off-farm wages

We derive an empirical measure of off-farm wage opportunities by using the 1994 Encuesta Nacional de Ingresos y Gastos de los Hogares (ENIGH) household survey to estimate off-farm wages as a function of gender, years of education, the interaction between gender and years of education, a quadratic function of age, and a state fixed effect. We limit this estimation to wage earners that were 18-50 years old since this population is more representative of the population of potential migrants. We then used the wage equation to predict wages for each adult in the set of Progresa households matched to ejidos. The maximum predicted off-farm wage amongst adults 18-50 was taken as the household's off-farm wage opportunity.[21] In columns (5) and (6) of Table 1.4 we estimate a separate certification effect for households above and below median values of off-farm wage opportunity. The difference in migration response to certification between households with high and low wage opportunities is statistically significant at the 10% level. Using the results from column (6), the estimated increase in the certification effect for male headed households that have above median off-farm wage opportunities is 0.026 and is statistically significant at the 10% level. These results are consistent with the theoretical prediction that the migration response should be larger for households that have higher wage opportunities outside of agriculture ($\frac{\partial \Delta h_m}{\partial w_m} > 0$).

[20] One potential issue with this interpretation is that the gender of the household head may reflect the available off-farm labor of the household. In Table 1.5, we show that households with 1-2 young males in the age range from 17-30 are if anything, more likely to respond to the program with migration. We also show that controlling for an indicator for whether the household has 1-2 young males and an interaction between this variable and the certification indicator does not change the female household head results. Thus, the result does not appear to be due to availability of potential migrants.

[21] Predicted wage was set to 0 if the household did not have any individuals in the 18-50 years old range.

Table 1.5: Heterogeneity according to gender and age composition of the household

	Progresa Households Matched to Ejidos			
	(1) Has Migrant	(2) Has Migrant	(3) Has Migrant	(4) Has Migrant
Certified	0.0077 (0.0076)	0.0098 (0.0081)	0.0003 (0.0074)	0.0019 (0.0080)
Certified*HH has 1-2 young males	0.0193 (0.0120)	0.0146 (0.0139)	0.0227* (0.0118)	0.0181 (0.0137)
HH has 1-2 young males	0.0152** (0.0060)	0.0091* (0.0049)	0.0149** (0.0060)	0.0090* (0.0049)
Certified*HH Head is Female			0.0625** (0.0246)	0.0672** (0.0275)
HH head is Female			0.0015 (0.0090)	0.0055 (0.0082)
Time Effects	Yes	Yes	Yes	Yes
Ejido Fixed Effects	Yes	Yes	Yes	Yes
Time Effects*HH has 1-2 young males	No	Yes	No	Yes
Time Effects*HH Head is Female	No	No	No	Yes
Mean of Dep Variable	0.056	0.056	0.056	0.056
Number of Observations	24533	24533	24533	24533
R squared	0.056	0.056	0.057	0.057

Standard errors that allow for clustering at the ejido level are reported in parentheses. Asterisks indicate statistical significance at the 1% ***, 5% **, and 10% * levels. Data include observations on all households in ejidos that completed the Procede process after 1996. All regressions are linear probability models. Dependent variable = 1 if the household had a migrant leave during the year or any previous sample year. Certified indicator = 1 if ejido was certified at the start of the year. All regressions include landholder indicator and age of household head as controls.

Heterogeneity in land productivity

The theory predicts that certification leads to a smaller migration response in places with higher land productivity ($\frac{\partial \Delta h_m}{\partial \gamma} < 0$). A common measure of land productivity in Mexico is rainfed corn yield. This measure has the advantage of its geographical coverage, as corn is the staple food grown all over the country. However it is only systematically available at the municipality level and since 2002 from SAGARPA (Ministry of Agriculture). We use the average corn yield over the period 2002-2008 as the measure of land productivity, and partition agricultural land as high or low productivity at the median yield of 1.29 tons/ha. Columns (1) and (2) of Table 1.6 show that, as predicted, the migration response to certification is weaker (and almost null) in ejidos where land is more productive.

Heterogeneity in land endowments

The model predicts a smaller migration effect for farmers with more land. Column 3 in Table 1.6 shows evidence that this holds in the data. The coefficient for relatively large landholders[22] is only 1/5 of that for small landholders.

The final prediction derived in the model is that large farmers in productive regions are expected to respond the least to certification with labor re-allocation ($\frac{\partial^2 \Delta h_m}{\partial \gamma \partial A} < 0$). We test for this by splitting the sample into low and high productivity areas (using the maize yield variable defined above) and estimating the effect of the program for large and small landholders (using the large landholder variable defined above). The results are striking. In low productivity areas, columns (6) and (7), larger landholders are not significantly less likely to migrate than land poor farmers. The coefficient is negative but insignificant. In contrast, in high productivity areas, columns (4) and (5), larger landholders increase their migration *significantly less* than land poor farmers. In fact, the overall effect of certification for land-rich households in high productivity areas is not statistically different from zero. In sum, these results are consistent with the prediction of the model that households are sorted according to their landholdings: larger, more productive farmers stay on the farm, whereas smaller more marginal farmers respond to the removal of use requirements by having more members migrate.

Certification and land use

The model we presented considered an autonomous household deciding how to allocate labor on and off the farm. According to the model, the freedom provided by certification makes constrained households allocate less labor to the farm. A logical byproduct of this phenomenon would be that less agricultural labor should be reflected in more land being left fallow. In reality, Procede also made land rental and sales transactions legal[23] and there

[22] We use an indicator variable which is equal to one if a family has more land per adult than the median in the ejido in 1997.

[23] Deininger and Bresciani (2001) report observing an increase in land rentals in 1997 compared to 1994.

Table 1.6: Heterogeneity in certification effect according to baseline land

	All			High Yield		Low Yield	
	(1) Has Migrant	(2) Has Migrant	(3) Has Migrant	(4) Has Migrant	(5) Has Migrant	(6) Has Migrant	(7) Has Migrant
Certified	0.0298***	0.0312***	0.0234***	0.0159*	0.0201**	0.0351***	0.0370***
	(0.0097)	(0.0101)	(0.0076)	(0.0090)	(0.0099)	(0.0129)	(0.0127)
Certified*High Maize Yield Municipality	-0.0232*	-0.0256**					
	(0.0124)	(0.0126)					
Certified * (Land per Adult > Median in Ejido [1997])			-0.0187*	-0.0240*	-0.0339**	-0.0095	-0.0145
			(0.0097)	(0.0122)	(0.0157)	(0.0157)	(0.0164)
Land per Adult > Median in Ejido [1997]	0.0061	0.0061	0.0098*	0.0134*	0.0035	0.0038	-0.0007
	(0.0051)	(0.0051)	(0.0051)	(0.0073)	(0.0057)	(0.0058)	(0.0045)
Ejido Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Time Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Time Effects*Land per Adult > Median in Ejido	No	No	No	No	Yes	No	Yes
Time Effects*High Maize Yield Municipality	No	Yes	No	No	No	No	No
Mean of Dep Variable	0.056	0.056	0.056	0.057	0.057	0.054	0.054
Number of Observations	24372	24372	24372	14533	14533	9839	9839
R squared	0.058	0.058	0.058	0.052	0.052	0.068	0.068

Standard errors that allow for clustering at the ejido level are reported in parentheses. Asterisks indicate statistical significance at the 1% ***, 5% **, and 10% * levels. Dependent variable in all regressions is 1 if the household is a migrant household. All regressions are linear probability models. Certified indicator = 1 if ejido was certified at the start of the year. All regressions include age of household head, indicator for female household head, and number of males between 17 and 30 as controls. Columns 1, 2 and 3 are for all matched ejidos in the Progresa sample. Columns 4 and 5 are for ejidos in municipalities with average maize yields above 1.293 tons/hectare. Columns 6 and 7 are for ejidos in municipalities with average maize yields below 1.293 tons/hectare.

are two reasons why land reallocation after Procede can be expected: first, it alleviates the inefficiently small farm size problem by allowing consolidation of production units in a context of increasing returns to scale; second, if some farmers are more productive than others, certification can allow for gains from trade through land markets to be realized.

We first test whether certification led to increased land concentration using a Herfindhal index of land concentration in ejidos using the Progresa data to estimate an ejido fixed-effect specification that allows us to assess changes in land concentration arising from Procede in a four year window. Column (1) in Table 1.7 reports results from a regression in which the Herfindahl index is the dependent variable. While the point estimate is positive (and reflects a 23% increase in land concentration) it is not statistically significant, possibly due to the small number of observations.[24] Given the large standard errors, we take away from this exercise that the evidence from household surveys points towards an increase in land concentration but the data is not conclusive.

Table 1.7: Effect of Procede on agricultural land use

	Progresa Data	LANDSAT Satellite Data		
	(1)	(2)	(3)	(4)
	Herfindahl	Log(Area Ag.)	Log(Area Ag.)	Log(Area Ag.)
Certified	0.0268	0.0013	-0.0080	-0.0175
	(0.0320)	(0.0093)	(0.0108)	(0.0136)
Certified * High Yield Municipality			0.0209**	0.0332*
			(0.0093)	(0.0182)
Ejido Fixed Effects	Yes	Yes	Yes	Yes
Time Effects	Yes	Yes	Yes	Yes
Time Effects*High Yield Mun	No	No	No	Yes
Mean of Dep Variable	0.116	5.718	5.714	5.714
Number of Observations	506	63392	58763	58763
R squared	0.547	0.012	0.012	0.012

Standard errors that allow for clustering at the ejido level are reported in parentheses. Asterisks indicate statistical significance at the 1% ***, 5% **, and 10% * levels. The dependent variable is the log of the area in agriculture in the ejido. *High Yield* is 1 if average maize yield in the municipality of the ejido is larger than 1.293 tons/ha.

Our second strategy to assess changes in land use is to test for aggregate changes in the amount of cultivated land in the ejido using objective data and a longer time horizon. If the certificates were used by families to simply leave the land fallow without risk of loss, we would observe a reduction in cultivated land in the ejido. Alternatively, if land was rented out or sold to other community members by households with migrants, we would observe no

[24]The small number of observations is due to the index being calculated using information from all households in a given ejido year.

changes in cultivated land. Finally, if the certification program provided better incentives to invest in agriculture, we could actually observe increases in cultivated acreage in spite of population reductions.

We test for this using panel data from Landsat providing cultivated area in 1993, 2002, and 2007.[25] At each of the three points in time we observe the amount of land allocated to agriculture, pasture, forest, jungle, and thicket in the ejido. We estimate the reduced form impact of certification on the logarithm of cultivated area in a standard fixed effects framework:

$$\log(Agland_{jt}) = \gamma_j + \alpha_t + \delta Certified_{jt} + \varepsilon_{jt}, \tag{1.12}$$

where j indexes ejidos and t refers to the year of the land use observation. Results reported in column (2) of Table 1.7 show that certification had no significant effect on the total area used for agriculture within the ejido. The coefficient is actually positive but very small (0.1%). This is a surprising result given the reduction in labor induced by Procede. If marginal farmers were abandoning land in order to migrate, then we would have observed a decrease in agricultural land after certification. Columns (3) and (4) however show a rich pattern of heterogenous effects by land quality. Column (3) shows that cultivated land actually increased with certification in agriculturally favorable regions but decreased in lower land quality areas. In column (4), we add controls for differential time trends in high and low yield areas. The estimated coefficient shows that certification is associated with an insignificant decline of cultivated land in low-yield regions. Point estimates range from -0.8 to -1.8%. In contrast, agricultural land increases with certification in high agricultural productivity areas. The point estimate ranges from 1.3 to 1.6%, and the difference between favorable and non-favorable areas is significant.[26]

We conclude our land use analysis by verifying that there is a correlation between population changes and cultivated area changes. In order to do this, we consider the overall change in log agricultural land between 1993 and 2007 using the Landsat data. The median change in log of agricultural land in these data is .0001 while the mean is 0.111. To limit the influence of outliers, we use the rank of the ejidos in the distribution of change in cultivated land.[27] The first two columns of Table 1.8 repeat the fixed effects regression of locality population on whether the ejido has been certified separately for the localities with agricultural land use change below and above the median value. The table shows that the negative effect of certification on population size is much stronger in localities that also saw the largest decreases in agricultural land. Column (3) shows that localities with the most pronounced declines in agricultural land ($rank$ =0) experienced a decline in population of

[25] INEGI GIS land use series II, III and IV.

[26] As a robustness check on the resolution of the Landsat images, we ran all the regressions in Table 1.7 after dropping the smallest 5% of ejidos. The coefficients change only minimally and statistical significance is unaffected (results not reported).

[27] The value of the variable Rank corresponds to the empirical distribution function of the change in the logarithm of agricultural land.

9.2% in response to certification, while ejidos with the largest increases in agricultural land saw no significant effect of certification on population.

Table 1.8: Population regressions by change in agricultural area

	Rank>0.5	Rank<0.5	All
	(1)	(2)	(3)
	ln(Population)	ln(Population)	ln(Population)
Year=2000	-0.2285***	-0.1936***	-0.1765***
	(0.0143)	(0.0195)	(0.0239)
Certified 1993-1999*Year=2000	-0.0230	-0.0760***	-0.0924***
	(0.0183)	(0.0232)	(0.0292)
Rank of Ag Change * Year=2000			-0.0705*
			(0.0368)
Rank of Ag Change * Certified 1993-1999 * Year=2000			0.0876*
			(0.0461)
Ejido Fixed Effects	Yes	Yes	Yes
Mean of Dep Variable	4.240	4.324	4.278
Number of Observations	15200	12420	27624
R squared	0.035	0.041	0.038

Dependent variable in all regressions is log of locality population. Standard errors that allow for clustering at the ejido level are reported in parentheses. Asterisks indicate statistical significance at the 1% ***,5% **, and 10% * levels. Data come from the 1990 and 2000 locality population censuses. Localities located in ejidos with no agricultural land during either 1993 or 2007 are excluded from the regressions, thus explaining the difference in observations from Table 1.2. The first column limits to localities in ejidos that experienced above the median change in log agricultural area from 1993-2007. The second column limits to localities in ejidos that experienced below the median changes. The final column is for all localities in ejidos that had nonzero agricultural land area in both 1993 and 2007.

In summary, in areas of low land quality, certification induced a strong migration response accompanied by a decline in cultivated land. In more favorable land quality regions, only the less well endowed households responded with migration, while the larger farmers did not migrate, and total land in agriculture increased slightly.

Effects of land certification on household consumption expenditures

How does land certification affect household-level consumption? If more secure rights to land allow households to allocate labor more optimally, then this could translate into increased consumption. To investigate this, we use the consumption modules in four rounds of the follow-up surveys from Progresa. The specific surveys were carried out in May 1998, October 1998, June 1999, and November 2000. Each survey had a detailed consumption module that allows us to calculate monthly expenditures per household member for both food and nonfood items.[28] 43% of the households from our main sample had the program completed in this interval.

The data show that land certification only led to increases in aggregate consumption per capita in areas of low land quality where the migration effects were the strongest. Column (1) of Table 1.9 shows that the overall immediate effect of certification is small and statistically insignificant. However, column (2) shows that certification led to a 7.5% increase (p=0.07) in monthly consumption per capita in low-yield areas where the program also led to increased migration. This suggests that reallocating household labor is an important mechanism through which land certification increases household welfare.

Turning to the breakdown of expenditures by category, land certification did not lead to increased food expenditures. Column (3) shows a small and statistically insignificant effect of certification. Column (4) separates certified households into those that were certified within the last six months and those that had been certified for more than six months at the time of the survey. Being certified for a longer period of time is unimportant for food consumption.

There is some evidence that certification led to increased consumption of nonfood items, particularly for households in ejidos that had the program completed more than six months before the survey date. The overall effect on nonfood consumption in column (5) is a 4.7% increase, however the coefficient is not statistically significant. This is largely due to timing. Column (6) shows that households in ejidos that had been certified for at least six months increase consumption of nonfood items by 16.7%.

Overall, the results on consumption suggest that our migration results are not driven by other mechanisms where migration happens as an undesired outcome of the program. For instance, we would not expect to see changes in consumption if the program allowed the most powerful ejidatarios to obtain a disproportionate share of ejido lands during the land registration process. While we can't definitely say that the increased consumption in low productivity areas is due to migration, the result is consistent with land certification allowing households to better allocate labor and thus increase consumption.

Alternative mechanisms from certification to migration

While the view taken in this paper has been that the increased migration caused by land certification is a result of relaxing the land use constraint, there is an alternative mechanism that would also be consistent with increased migration. Namely, land certification could have relaxed liquidity constraints by allowing poor households to sell or rent their land and use those funds to finance migration.[29] While this would not invalidate the link between certification and migration, it refers to a completely different cause of increased migration.

[28] Nonfood items include transportation, medicine, fuel and electricity, hygiene products, clothing, and home accessories.

[29] In the context of Mexico, McKenzie and Rapoport (2007) have shown that migration to the U.S. is related to wealth.

Table 1.9: Effect of Procede on household consumption

	Log consumption per capita		Log food consumption per capita		Log nonfood consumption per capita	
	(1)	(2)	(3)	(4)	(5)	(6)
Certified	0.003	0.075*	-0.012		0.047	
	(0.032)	(0.040)	(0.033)		(0.054)	
Certified*High Yield Municipality		-0.122**				
		(0.047)				
Certified 6 months or less				-0.028		-0.036
				(0.039)		(0.059)
Certified more than 6 months				0.011		0.167***
				(0.037)		(0.063)
Ejido Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Time Effects	Yes	Yes	Yes	Yes	Yes	Yes
Mean of Dep Variable	4.838	4.838	4.470	4.470	3.435	3.435
Number of Observations	25482	25482	25373	25373	24965	24965
R squared	0.234	0.235	0.226	0.227	0.186	0.187

Standard errors that allow for clustering at the ejido level are reported in parentheses. Asterisks indicate statistical significance at the 1% ***, 5% **, and 10% * levels. The dependent variable is the log of monthly expenditures – either for all goods or food and nonfood items separately. *High Yield Municipality* is 1 if average maize yield in the municipality of the ejido is larger than 1.293 tons/ha.

In particular, it would imply that credit constraints were the critical factor holding people in agriculture, not the land use requirement.

One way to distinguish between these two competing explanations is by taking advantage of the Progresa experiment. Progresa randomly allocated cash transfers across villages in our sample to poor households equivalent to 140% of monthly food consumption per adult (Angelucci and De Giorgi, 2009). Because the cash payments were awarded to the poorest families, Progresa would have alleviated liquidity constraints in households where the restriction was more likely to be binding. Crucially, because these large cash inflows to poor families were occurring *before* certification, liquidity constraints would have been less binding in Progresa treatment villages when Procede arrived. Hence if the liquidity constraint story is correct, we should observe *smaller* effects of certification in Progresa treatment villages. We test for this by estimating the following regression:

$$y_{ijt} = \delta_1 Certif_{jt} + \delta_2 Certif_{jt} * Progresa_j + \gamma_j + \alpha_t + \varepsilon_{ijt}. \quad (1.13)$$

The specification is similar to our main specification in Table 1.1. The only difference is that in this specification we test whether the migration response is differentiated according to Progresa treatment status. Note that the because of the ejido fixed effects, the specification allows for Progresa to have a direct effect on migration (these are explored in Stecklov et al. (2005) and Angelucci (2012a)).[30] The specification does impose the same time trends for both groups. We relax this assumption by also showing a specification which adds Progresa-treament specific time trends.

$\delta_2 < 0$ would be evidence that liquidity is the mechanism causing certification to increase migration. The results in columns (1) and (2) of Table 1.10 do not support this story. If anything, the migration effect is *larger* in Progresa treatment villages. In Column 1, the certification effect in Progresa control villages is .006 and that in treatment villages is .021. The difference between control and treatment villages is economically large, *positive*, but not statistically significant (p =0.25). The same story holds in column (2) where we allow for differential time trends in Progresa treatment villages. We hence reject the hypothesis that liquidity was the factor holding back households in ejidos.

Rural labor markets

Hiring outside labor was technically illegal prior to Procede. It is possible then that ejido households substituted hired labor for family labor, thus allowing family members to migrate. However, recall that the estimates from the locality level regressions (which correspond to net migration) were of similar order of magnitude to the individual household estimates. This suggests that substituting hired labor for family labor was not an important phenomenon in the data.

We can nonetheless inquire whether Procede led to more efficient rural labor markets. Benjamin (1992) shows that frictions in rural labor markets generate non-separation between production and consumption decisions. In our context, the correlation between household size and the amount of land cultivated can be expected to decrease after the completion of Procede. The intuition is that a large labor endowment was necessary to cultivate a large amount of land prior to the program. If the program had a significant impact on the labor market, then this correlation should decrease after completion of Procede. We estimate:

$$Hectares_{ijt} = \beta_0 Certif_{jt} + \beta_1 Adults_{ijt} + \beta_2 Adults_{ijt} * Certif_{jt} + \gamma_j + \mu_t + x_{ijt}\alpha + \varepsilon_{ijt}, \tag{1.14}$$

where $Hectares_{ijt}$ is land cultivation and $Adults_{ijt}$ is the time-varying number of adults in the household. A negative and significant estimate of β_2 would suggest an increased separation between the household as a firm and the household as a consumer, which would be interpreted as working through rural labor markets. The estimate in Column (3) of Table 1.10 shows that the correlation between household size and cultivation does not decrease significantly after program completion. Thus the data are consistent with the certificates liberating family labor from the farm, but not with hiring in of workers to substitute for family labor.

[30] We use ejido fixed effects to maintain consistency with our previous specification. The direct effect of Progresa on migration is fully absorbed when locality fixed effects are used. Since the match between localities and ejidos is near one to one, the results of these two specifications are very similar.

Table 1.10: Alternative explanations of migration effect of land certification

	(1) Migration	(2) Migration	(3) Hectares
Certified	0.0065	0.0056	-0.1135
	(0.0085)	(0.0086)	(0.1670)
Certified*Progresa locality	0.0143	0.0153	
	(0.0123)	(0.0123)	
Adults in HH			0.4314***
			(0.0393)
Certified*Adults in HH			0.0296
			(0.0414)
HH Head is Female			-0.4606***
			(0.0741)
Age of HH Head			0.0268***
			(0.0027)
Ejido Fixed Effects	Yes	Yes	Yes
Time Effects	Yes	Yes	Yes
Time Effects*Progresa Locality	No	Yes	No
Time Effects*Adults in HH	No	No	Yes
Mean of Dep Variable	0.054	0.054	2.121
Number of Observations	26690	26690	24211
R squared	0.047	0.047	0.288

Standard errors that allow for clustering at the ejido level are reported in parentheses. Asterisks indicate statistical significance at the 1% ***, 5% **, and 10% * levels. Column 3 drops the top 1% of observations with hectares cultivated of 25 or more.

1.6 Internal validity checks

We present several tests that support the validity of the identifying assumptions of the paper. The main threat to identification in the Progresa dataset is correlation between the timing of Procede and the time-path of migration in the ejido. The estimated average program effect would be biased if completion of Procede were correlated with pre-program changes in migration. To investigate the possibility of bias in program timing, we use a standard regression of *pre-program* changes in ejido level migration rates on indicators for the year Procede was completed:

$$\Delta y_{jt} = \gamma + \alpha_t + \sum_{k \geq t} \delta_k I(Procede\ Year_j = k) + \varepsilon_{jt} \quad \forall t \leq Procede\ Year_j. \tag{1.15}$$

The dependent variable Δy_{jt} is the change in the average level of the migrant indicator in ejido j from year $t-1$ to year t. The key independent variables are a set of dummy variables, $Procede\ Year_j = k$, for the year in which the program was completed in the ejido. Since the data cover the years 1997 to 2000, only three such variables are necessary for the ejidos certified in 1999, 2000, or after 2000.[31] *Procede Year* effects that are jointly significant would indicate that year of program completion was correlated with pre-program changes in migration. The results are reported in the online appendix in Table 1.11 where we report results separately for changes in migration from 1997-1998, 1997-1999, and 1997-2000. Year of program completion does not significantly explain pre-program changes in migration in either of the three regressions. Lack of a significant correlation between the year of Procede completion and changes in ejido level migration rates over time provides evidence that pre-program time trends in migration were not correlated with completion of the program.

Another possibility is that the timing of Procede is correlated with sharp changes in migration prior to the program. If Procede was rolled out in response to sharp declines in migration prior to the program, then our estimate would simply reflect reversion to mean migration levels. Perhaps more likely, if households anticipated the program and reduced migration to oversee the certification process, then post-program returns to normal migration rates would confound our estimate. We estimate the following specification to consider this potential Ashenfelter dip effect (Ashenfelter, 1978):

$$y_{jt} = \gamma_j + \alpha_t + \beta_0 \cdot \textit{(Year of)}_{jt} + \beta_1 \cdot \textit{(Year before)}_{jt} + \beta_2 \cdot \textit{(2 Years before)}_{jt} + \varepsilon_{jt}, \quad (1.16)$$

where y_{jt} is average migration at the ejido level, and other variables are indicators for the year of, year before, and two years before program completion. The β coefficients indicate whether migration levels were significantly different than average in the ejido during the years directly before the program. Column (4) of Table 1.11 gives the results of estimating (1.16). The point estimates are very small and statistically insignificant (the smallest p-value is 0.84), yet the standard errors are large. An ideal result of the regression would be a set of precisely estimated zeros on the three indicator variables. While we cannot reject large coefficients, it is reassuring that there are no obvious significant changes in migration in the years leading up to completion of the program. We interpret the combined results in the table as providing no clear evidence that our identification strategy is biased by correlation between program completion and pre-program migration.

A similar concern with our identification strategy is that anticipation of being certified in the future would lead to a decrease in migration in uncertified ejidos. Our observed increase in migration would then reflect an anticipation effect and not a true migration effect. The results in column (4) of Table 1.11 are not consistent with a large decrease in migration during the years immediately prior to the program.

Finally, another potential issue of concern is attrition of households from the ENCEL survey. 11.2% of households with an interview completed in 1998 did not have an interview completed in 1999. The percentage rose slightly to 12.7% in 2000.[32] In Table 1.12 we run the

[31] The base group is composed of ejidos certified in 1998 since we require the ejido to be certified at the start of the year to be considered as certified for that year.

[32] We define attrition as the interview not being conducted for any purpose.

Table 1.11: Relationship between Procede and pre-program migration

	Progresa Households Matched to Ejidos, Pre-Program Period			
	(1) Δ Migration,97-98	(2) Δ Migration,97-99	(3) Δ Migration,97-00	(4) Migration,97-00
Procede Completed in 1999	-0.0011			
	(0.0113)			
Procede Completed in 2000	-0.0040	-0.0087		
	(0.0110)	(0.0092)		
Procede Completed After 2000	-0.0131	-0.0102	0.0015	
	(0.0090)	(0.0086)	(0.0046)	
Year Procede Completed (0/1)				0.0018
				(0.0150)
Year Before Procede (0/1)				-0.0021
				(0.0107)
2 Years Before Procede (0/1)				-0.0015
				(0.0089)
Time Fixed Effects	No	Yes	Yes	Yes
Ejido Fixed Effects	No	No	No	Yes
Mean of Dep Variable	0.022	0.020	0.018	0.050
Number of Observations	111	187	225	406
Number of Ejidos	111	94	76	127
R squared	0.047	0.019	0.002	0.774
Pvalue of joint test	0.190	0.493		

Standard errors are reported in parentheses. Robust standard errors are used in column 1. In columns 2-4, standard errors are clustered at the ejido level. Asterisks indicate statistical significance at the 1% ***, 5% **, and 10% * levels. The dependent variable in columns 1-3 is the change in ejido migration rate. The dependent variable in column 4 is the ejido migration rate. Both regressions are for the pre-treatment period. Columns 1 is for for 1998. Column 2 is for 1998-1999. Column 3 is for 1998-2000. Column 4 is for 1997-2000.

basic regression used to identify the role of Procede on migration, equation (1.7), on attrition. The coefficient of the certified variable is both insignificant and very small. There is therefore no evidence that the migration effect we estimate could be due to selective attrition.

Table 1.12: Regressions of attrition on certification status and household covariates

	(1) Attrition
Certified	-0.003
	(0.025)
HH is Landholder	-0.043***
	(0.010)
Number Males 17-30 in HH	0.005
	(0.004)
HH Head is Female	0.030**
	(0.012)
Age of HH Head	-0.000
	(0.000)
Ejido Fixed Effects	Yes
Time Fixed Effects	Yes
Mean of Dep Variable	0.112
Number of Observations	12895
R squared	0.115

Standard errors that allow for clustering at the ejido level are reported in parentheses. Asterisks indicate statistical significance at the 1% ***, 5% **, and 10% * levels. Data are for all households that were surveyed in the Fall 1998 ENCEL survey. Observations are from 1999 and 2000. Dependent variable = 1 if household did not have survey completed. Certified indicator = 1 if household had a certificate at the start of the year. 446 households attrited in 1999 but not in 2000. 331 households attrited in both 1999 and 2000. 554 households attrited in 2000 but not in 1999.

1.7 Conclusions

Delinking land rights from land use has been the focus of a number of large land certification programs. In this paper, we showed that if property rights were tied to land use requirements in the previous regime, these policy reforms can induce increased outmigration from agricultural communities. We provided evidence on this phenomenon by analyzing the Mexican ejido land certification program which, from 1993 to 2006, awarded ownership certificates to farmers on about half the country's farm land.

We used three independent datasets to document a strong migration response in agricultural communities where certificates were issued. Families that obtained certificates were subsequently 28% more likely to have a migrant household member and the locality's overall population fell by 4%. The estimated effect increased over time. We documented heterogeneity in migration response according to the ex-ante level of property rights insecurity, level of off-farm opportunities, initial plot size, and land quality.

There is also evidence of sorting within the community: larger farmers stayed, whereas land-poor farmers left, and this effect was starker in high productivity areas. This prompts the question of whether total acreage under cultivation decreased with the program. We found that, on average, cultivated land was not reduced because of the program. However, this masks an interesting heterogeneity. While in low land quality regions agricultural land was reduced, in high land quality regions the certification program led to increases in agricultural land, which we attribute to gains in agricultural labor productivity or increased incentives to invest. Overall, the evidence shows that certification of ownership increases the efficiency of labor allocation across space by inducing low productivity farmers to migrate, while leaving higher productivity farmers in place and allowing them to consolidate land. Because smallholder farmers are the ones most likely to leave after certification, efficiency gains are accompanied by immediate benefits for them. These results are most consistent with a model where the key constraint imposed by insecure property rights is the requirement of continued presence. The empirical evidence is not consistent with alleviation of liquidity constraints being the mechanism explaining the increase in migration.

The literature on property rights focuses on investment and increased access to credit as key pathways between rural land reform and economic growth (Galiani and Schargrodsky, 2011). Outmigration from rural areas has only recently received attention. Our results suggest that the permanent reallocation of labor between sectors of the economy can be an equally important pathway resulting in the effects of agricultural land reform extending beyond rural areas. The importance of low agricultural labor productivity in explaining low aggregate output across countries suggests that enhancing agricultural labor productivity can possibly have large effects (Restuccia, Yang, and Zhu, 2008). Removing the barriers to migration through property rights reforms is one way to achieve this. An important policy implication of our results is thus that improvement of property rights through formal land certification can have not only direct effects on investment, land markets, and land use patterns but also indirect effects on the spatial performance of labor markets, resulting in particular in large flows of rural migrants.

Chapter 2

Behavioral effects of risk-reducing seed varieties

2.1 Introduction

Exposure to uninsured risks is a major factor contributing to poverty, low productivity, and slow growth. This occurs through two mechanisms. The first is the direct cost of shocks that periodically destroy assets and alter growth trajectories, with short run exposure to adverse events eventually having very large long-term consequences (Alderman, Hoddinot, and Kinsey, 2006; Maccini and Yang, 2009). The second is self-insurance whereby risk-averse producers who anticipate the occurrence of shocks shift management decisions toward activities with lower expected returns in exchange for lower exposure to risk (Binswanger and Rosenzweig, 1993; Morduch, 1995). The latter effect presents a major setback for growth. Donovan (2014) estimates that uninsured production risk leads to less use of intermediate inputs and thus increases the difference in agricultural productivity between poor and rich countries by 40%.

While the first mechanism is well-studied, we focus on the latter mechanism by asking how does risk affect the choices made by small farmers? To study this, we focus on flooding risk faced by farmers in rural Odisha India. We use the randomized dissemination of a new flood-tolerant rice seed that effectively decreases risk by reducing the damages caused by prolonged flooding. The technology is well-suited for the experiment because it offers flood tolerance – thus reducing risk – without any other meaningful differences compared to the existing popular seed variety. Our intervention can therefore be thought of as partially reducing the production risk faced by farmers.

This leads to our main finding: reducing risk with technological progress causes farmers to shift towards behaviors that are more risky, but lead to higher expected productivity. As examples, farmers with access to the new technology use more inputs such as fertilizer and land, shift away from cheap and less productive planting techniques, and increase the use of agricultural credit. These changes translate into increases in productivity – measured by output per hectare – in years when flooding is absent.

There are two implications of this finding. First, risk is an important factor that discourages farmers from investing in inputs and thus reduces overall productivity. Second, we are the first to show that *new technologies* that reduce weather risk have meaningful impacts on the choices made by farmers. By decreasing the need to self-insure and thus encouraging farmers to take more risk, the benefits of these technologies extend beyond stabilizing yields during extreme events. This finding is important given the substantial scientific effort to create new seed varieties that are more tolerant to weather extremes.[1]

How does the technology that we rely on allow us to study risk and economic behavior? Swarna-Sub1 was developed by use of modern plant breeding techniques to insert the SUB1A (flood tolerance) gene into Swarna – eastern India's most popular rice variety.[2] Excluding flood tolerance, the technology is otherwise genetically similar to Swarna – a fact that has been extensively documented by agricultural scientists (Neeraja et al., 2007; Bailey-Serres et al., 2010; Mackill et al., 2012). Therefore, the technology increases yield during flooding while leaving it unaffected during normal years when flooding is absent. This has been shown in agronomic trials (Singh, Mackill, and Ismail, 2009). The genetic similarity between the two varieties also means that other important characteristics such as responsiveness to fertilizer are unaffected by the introduction of flood tolerance.

We randomized the distribution of this new technology across 128 villages of Odisha India. Prior to the 2011 wet season, a random subset of five farmers in each of 64 treatment villages were provided a minikit of Swarna-Sub1 seed. The minikit contained only a small amount of seed and a short information sheet on the variety. The comparison farmers in the remaining villages were not provided with any seeds, as the base variety of Swarna is widely grown in the area.

Spatial variation in flooding intensity during September 2011 allowed us to verify the yield properties of the variety. That is, Swarna-Sub1 significantly increases yield during flooding without any differences in yield when flooding does not occur (Dar et al., 2013). The results presented in this paper are from the following 2012 wet season – a year when flooding did not occur.

We show three main findings from the second year of the experiment. First, farmers given access to the flood-tolerant variety use more inputs on rice, which is the only wet season crop for over 90% of the sample.[3] Treatment farmers increase *overall* cultivated area, sowing on average 0.7 additional plots, implying an increase in area of 10%. Part of this response is driven by treatment farmers being less likely to take land out of production,

[1]These technological advancements include development of rice varieties that are more tolerant to flooding (Xu et al., 2006; Hattori et al., 2009), development of rice and maize varieties that are tolerant to drought (Capell, Bassie, and Christou, 2004; Karaba et al., 2007; Nelson et al., 2007), and identification of genes in rice that confer tolerance to extreme cold temperatures (Fujino et al., 2008).

[2]The SUB1A gene confers flood tolerance by restricting the natural elongation response of the plant, thus preventing the crop from failing by losing stored energy (Xu et al., 2006; Fukao and Bailey-Serres, 2008).

[3]Due to a lack of irrigation in the dry season, most farmers grow a single rice crop only in the wet season. Only 20% of the sample cultivates a dry season crop.

particularly lower-lying and lower quality plots. Access to the technology also causes farmers to be 33% less likely to use a traditional planting technique known as broadcasting, which is cheaper, less labor intensive, but less productive. In addition, treatment farmers use approximately 11% more fertilizer, conditional on total area cultivated. The increase in fertilizer use is consistent with a risk explanation because it is concentrated on the types of fertilizer typically used for soil conditioning at the time of planting or early in the growing season, i.e. the fertilizers that are at the most risk of loss to flood.

Our second set of findings are on credit usage and savings behavior. Treatment farmers are 36% more likely to utilize credit. These loans are primarily agricultural loans distributed from local cooperatives early in the growing season. This effect can plausibly be explained by either demand or supply side responses in the credit market. That is, the technology decreases the probability of the low production state, which could increase the demand for credit. Alternatively, reducing risk could increase the supply of credit by decreasing default risk. Our design does not allow us to distinguish between these alternative mechanisms.[4]

Farmers given access to Swarna-Sub1 also reduce the share of the harvest that is saved for future consumption by five percentage points. Given that farmers in the sample set aside excessive amounts of rice, a plausible interpretation of this finding is that reducing risk crowds out these precautionary savings. Also, the effect is smaller for households that are at least partially insured by access to government subsidized rice, which is also consistent with this precautionary savings explanation.

Third, we show that productivity increases as a result of having access to the risk-reducing variety. Specifically, rice yield on plots cultivated by treatment farmers increases by approximately 10%. This occurred during a year when there was no flooding. Therefore, the effect is most likely driven by changes in management. Along these lines, we show evidence that the productivity effect decreases significantly when conditioning on fertilizer use, planting technique, and credit access. This result is consistent with reducing risk leading to changes in farm management which in turn lead to increases in productivity.

Our results could be alternatively explained by two different mechanisms. First, wealth effects could explain the results if the success of Swarna-Sub1 during the first season increased wealth and thus increased demand for inputs during the second season. We show that this explanation is infeasible because the amount of seed provided to treatment farmers during the first year was sufficient for testing and seed multiplication, but not large enough to generate meaningful wealth effects. Second, we show in a simple model that the results could be alternatively explained by the technology increasing the marginal products of inputs during flooding. This explanation is less consistent with our results on savings and credit because these actions obviously don't enter the production function. Also, the key effects on fertilizer and planting techniques persist on plots that are not cultivated with Swarna-Sub1, which suggests that reducing risk is an important channel.

We are the first to study how technological progress aimed to reduce risk in agriculture

[4]The supply side explanation seems more likely given that liabilities are often waived by cooperatives after years of heavy flooding or drought.

affects decisions made by farmers. In contrast to technology, insurance is another tool that can transfer risk away from farmers and can affect their decisions. Thus, our results are most related to a recent literature on index-based weather insurance and decision-making in agriculture (Mobarak and Rosenzweig, 2012; Cole, Giné, and Vickery, 2013; Karlan et al., 2013). The common finding among these papers is that access to insurance causes farmers to increase the use of risky inputs.

While this literature clearly suggests potential for insurance to improve agricultural productivity, several barriers have limited wide-scale adoption of index insurance. These barriers include lack of trust, credit constraints, and risk from being far from the station where weather is measured (Giné, Townsend, and Vickery, 2008; Cole et al., 2013). A range of empirical studies have found adoption of index insurance to vary from 5-50% (Giné, Townsend, and Vickery, 2008; Mobarak and Rosenzweig, 2012; Cole et al., 2013; Dercon et al., 2014; Karlan et al., 2013). One advantage of new technologies is that some of these barriers to adoption become less relevant for new seed varieties.

In addition to barriers to adoption, another important difference is that new seed varieties reduce risk rather than transfer it to insurers. If variability in agricultural productivity affects non-farmers through prices or wages, as in Mobarak and Rosenzweig (2014), then risk-reducing technologies would also be expected to benefit those that are not farmers.

We also add to a literature on the barriers to technological progress in agriculture (Duflo, Kremer, and Robinson, 2011; Suri, 2011). We point to production risk as a key barrier that limits investment in agriculture. Past work that has arrived at a similar conclusion has relied on spatial variation in exposure to risk (Binswanger and Rosenzweig, 1993; Kurosaki and Fafchamps, 2002; Dercon and Christiaensen, 2011). In contrast, the uniqueness of the new technology that we study allows us to experimentally vary the level of production risk faced by farmers.

The rest of this chapter is organized as follows. In section 2.2 we provide additional background on the technology, with particular emphasis on how it reduces risk. This section also provides more details on the experiment and outlines data collection. Section 2.3 develops a basic model of input use and savings with production risk. The model guides our empirical analysis and particularly emphasizes the two main mechanisms through which the technology changes decision-making. Section 2.4 presents results, including analysis that rules out some alternative non-risk mechanisms. Section 2.5 concludes.

2.2 Background, experimental design, and data

In this section we summarize the evidence showing that the technology used in the experiment offers reduced risk to farmers. We then discuss the details of the experimental design and sampling. Finally, we outline the timing of data collection and present summary statistics on both village and household characteristics.

Background on technology

The genetic similarity between Swarna and Swarna-Sub1 is tremendously important for interpreting our results. This genetic similarity was indeed an important target of breeding efforts to create Swarna-Sub1. While scientific effort to create submergence-tolerant seed varieties dates back to the 1970's, it was not until more modern plant breeding techniques became popular that the SUB1 genes were successfully bred into popular varieties without introducing other undesirable traits (Xu et al., 2006; Mackill et al., 2012). The result is a new variety that offers a decrease in risk without any losses in yield when fields are not flooded.

We use data from the first year of this experiment to verify that this property holds in farmers' fields. Specifically, Swarna-Sub1 offered a significant advantage in rice yield during flooding of approximately 6 to 14 days (Dar et al., 2013). The magnitude of this agronomic impact is meaningful: farmers with Swarna-Sub1 experienced an approximate 65% increase in yield when fields were submerged for 13 days. At the same time, yield differences were small and statistically insignificant when flooding did not occur. Combining these results, the technology both reduces the variance of output and increases the expected level of output.

Experimental design

Our sample is drawn from villages in flood-prone areas of the Bhadrak and Balasore districts of northern Odisha. This area is suitable for the study because flood risk is high, Swarna is widely grown, and Swarna-Sub1 was still unknown and unavailable to farmers in May 2011 when the project was initiated. The villages were identified from two sources. In Bhadrak, RADARSAT satellite imagery was matched to a GIS database of villages to identify villages that were affected by flooding during recent floods. A random subset of 64 affected villages was selected for inclusion in the study. We used a list of flood-prone villages from our local NGO partner to randomly select 64 villages in Balasore district.[5] Figure 2.1 displays a map of the study area and the villages included. As is seen in the map, almost all of the study villages are in low-lying coastal areas.

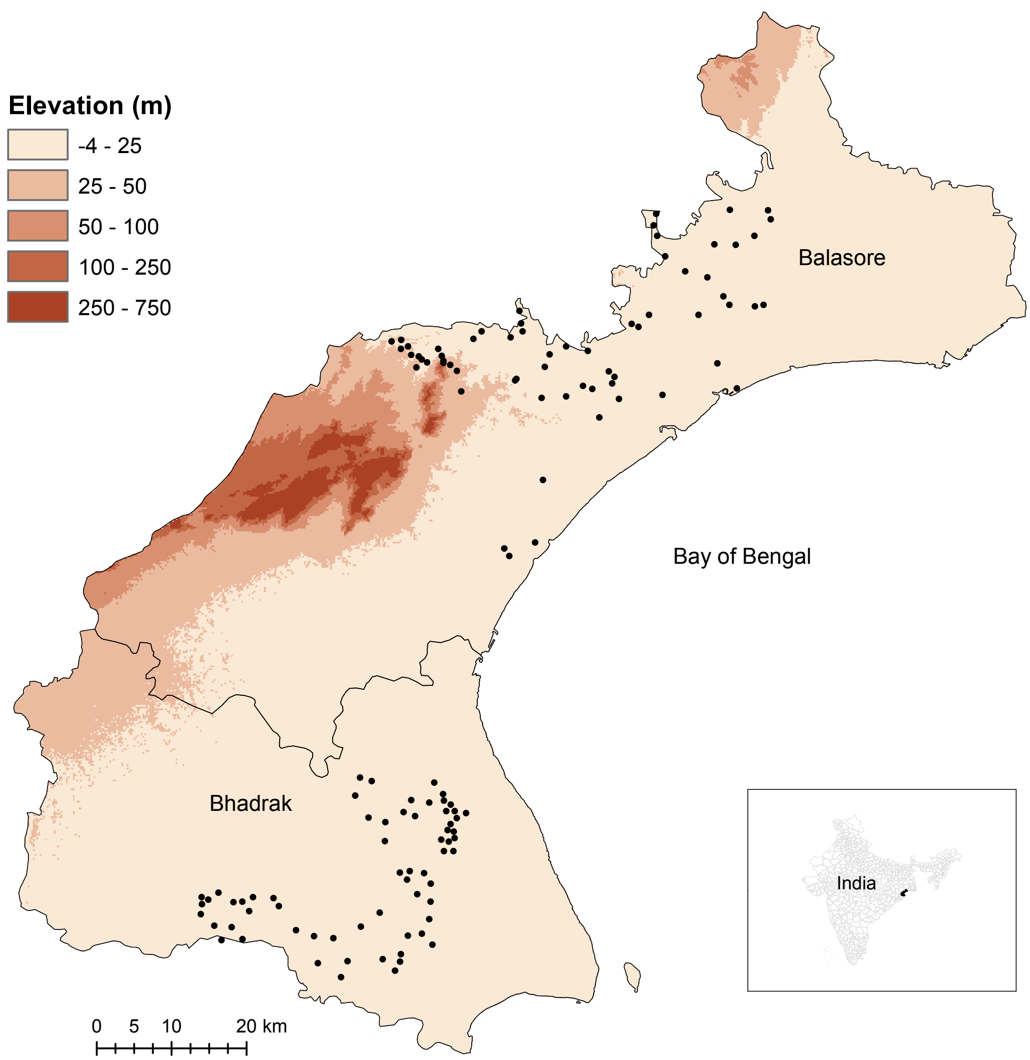
The timing of the experiment is as follows. In the 64 treatment villages, five farmers were randomly selected to receive minikits containing five kilograms of Swarna-Sub1 seeds.[6] Each minikit was delivered during June 2011, which corresponds to the time of planting for year one of the experiment. The comparison group consists of ten randomly selected non-recipients in treatment villages and five randomly selected farmers in the 64 control villages.

A large flood in September 2011 affected several districts in coastal Odisha. Many of the villages in our sample were affected by the flood, with the most severe inundation occurring in the southern-most district of Bhadrak. Flood damage was high, as the crop was entirely

[5]The satellite imagery of historic floods in Balasore was not available at the time of village selection.

[6]Five kilograms of seed is sufficient to cultivate 0.1 to 0.2 hectares, or approximately 10-20% of average cultivated area. This small size of the minikit is standard, but also helps eliminate large wealth effects.

Figure 2.1: Location of villages



lost on approximately 11% of the plots in our sample. Importantly, the flood allowed for the effects of the technology on yield to be estimated for different levels of flood severity.

Data collection

Our first follow-up survey was conducted in March 2012 after the first year's crop was harvested. A total of 1,248 farmers were reached, making a response rate of 97.7%.[7] The main objective of the initial follow-up was to establish the effects on productivity that are reported in Dar et al. (2013).

The second follow-up survey was carried out one year later after the growing season for year two. A total of 1,237 of the farmers surveyed during 2012 were reached again during this survey. The important outcomes of interests are farm-level information on fertilizer use, allocation of output across uses, and credit utilization. In addition, a plot-level module on seed variety choice, planting methods, and production was administered. All of the analysis in this paper uses these data from the second follow-up, unless explicitly noted otherwise.

Compliance with the design – defined as continued cultivation of Swarna-Sub1 during year two – was high. 76% of minikit recipients cultivated the technology during year two.[8] The average number of plots sown with Swarna-Sub1 amongst minikit recipients was 1.5. Average land area cultivated with Swarna-Sub1 was 0.33 hectares, or approximately a third of average landholdings. Conversely, only 10.1% of of control farmers cultivated Swarna-Sub1 during year 2. This was a direct result of seed transfers from original recipients: 13.3% of control farmers cultivated Swarna-Sub1 in treatment villages and only 3.3% did so in control villages. Given this fairly low level of overall non-compliance, we take a conservative approach and report intention-to-treat (ITT) results throughout the paper.

Summary statistics

Villages in the sample are fairly representative of the low-lying villages in the flood-prone states of Bihar, Odisha, and West Bengal. Table 2.1 shows village characteristics from the 2001 census.[9] With the exception of village size, the sample villages are roughly similar to all villages in low-lying areas of the three states.

Turning to the microdata from our sample, Table 3.1 reports summary statistics of variables that are likely predetermined, but were collected during the first follow-up survey. Most importantly, treatment and control households look similar on most observable characteristics. Farms in the sample are small. Average landholdings are less than one hectare. While

[7]This small level of attrition is balanced across treatment and control. Enumerators were not able to contact farmers in one control village due to disagreement with local village leaders about participation in the program. The results reported are for the remaining 127 villages.

[8]The most common reason for disadoption was crop loss during year one. Swarna-Sub1 is not suitable for low areas where water remains stagnant after flooding (Singh, Mackill, and Ismail, 2011). Swarna-Sub1 that was planted in these areas during year one was lost.

[9]Villages in the other three states were included if the elevation was below 56 meters, the maximum elevation in the sample.

Table 2.1: Characteristics of sample villages and other low-lying villages in Odisha, West Bengal, and Bihar

	In experiment	Other villages in 3 states
Number households	177.72	307.66
	(175.536)	(408.872)
Household size	5.31	5.26
	(0.891)	(0.843)
Share Scheduled Caste	0.20	0.23
	(0.202)	(0.246)
Share Scheduled Tribe	0.09	0.07
	(0.181)	(0.183)
Share cultivating land	0.12	0.09
	(0.069)	(0.067)
Share agricultural laborers	0.06	0.09
	(0.066)	(0.082)
Literacy rate	0.60	0.51
	(0.110)	(0.182)

All data are taken from the 2001 census. Column 1 pertains to 125 villages in Odisha that were part of the experiment. 3 of the 128 sample villages were not successfully matched to the 2001 census. Column 2 pertains to the other 55,324 villages in Odisha, Bihar, and West Bengal that have an elevation of less than 56 meters (the maximum elevation of the sample villages). Numbers in parentheses are standard deviations.

electricity is fairly widespread, few households have access to piped water. Most households rely on either village or private tubewells for water. Approximately 56% of households hold Below the Poverty Line (BPL) cards, which provide access to government supports such as a monthly allotment of subsidized rice.

2.3 A basic model of risk and decisions

In this section we develop a household-level model of optimal input choices under production uncertainty. We set up the model to realistically include the properties of Swarna-Sub1. Most importantly, the model clarifies the two mechanisms through which introducing this new technology can change input use: by reducing risk and by increasing the expected marginal product of the input. We use this result when distinguishing between these two mechanisms in the empirical analysis.

Table 2.2: Mean values of household characteristics by treatment status

	Control	Treatment	P-value of difference
Land owned in hectares	0.810	0.868	0.18
HH has private tubewell	0.332	0.325	0.82
HH has piped water	0.035	0.057	0.14
HH has refrigerator	0.078	0.076	0.92
HH has television	0.628	0.605	0.50
Education of farmer	6.896	6.946	0.83
Age of farmer	51.191	51.783	0.46
HH has thatched roof	0.557	0.548	0.79
HH has latrine	0.289	0.354	0.08
HH has electricity	0.843	0.822	0.42
HH has below poverty line card	0.574	0.559	0.67
ST or SC	0.189	0.176	0.59

Data are from year 1 follow-up. Values in columns 1 and 2 are means. P-values in column 3 are based on t-tests of equality of means. Standard errors are adjusted for clustering at the village level. ST refers to Scheduled Tribe and SC refers to Scheduled Caste.

Model setup and comparative statics

The model has two periods. Investment and savings decisions are made in the first period and output is realized in the second period.[10] We assume that the farmer holds an exogenously determined amount of rice in the first period. This amount is denoted as h. The farmer chooses between saving an amount s for the next period and consuming or selling the remainder immediately. The farmer can invest in a continuous amount of input x at a cost of r. Non-farm income is denoted as w.

After the crop is planted, there are two states of nature that can be realized in the second period. The farmer experiences a flood with probability α and no flood with probability $1-\alpha$. Consumption in the second period takes place after the state of nature is realized and the crop is harvested.

Introducing the flood-tolerant variety results in a change in the production technology.

[10]In reality the decision making process of the farm household occurs in three periods. The savings decision is made after harvest, the input decision is made at or near the time of the next planting, and the harvest is realized at the end of the growing season. We simplify the model by assuming that the first two events occur during the same period.

We denote ϕ as the amount of flood-tolerant seeds used. Usage of flood-tolerant seeds is exogenous, matching the randomization in our experiment. The production function in a non-flood year is $f_0(x)$. This formulation explicitly makes the assumption that $\frac{\partial f_0}{\partial \phi} = 0$ for any value of x. This assumption is consistent with the property that the productivity of Swarna-Sub1 is indistinguishable from Swarna when flooding does not occur. The production technology in the event of flood is $f_1(x, \phi)$.

We assume that investment does not enhance productivity more in a flood year than a non-flood year, that is $\frac{\partial f_0}{\partial x} > \frac{\partial f_1}{\partial x}$ for all values of x. In addition, the level of production under flooding is larger when more flood-tolerant seeds are used. More formally, $\frac{\partial f_1}{\partial \phi} > 0$.

Finally, consumption in the first period is c, consumption in the second period if there is flood as c_1, and consumption in the second period without flood is c_0. The discount factor is δ.

The maximization problem of the farmer is,

$$\max_{c,c_0,c_1,s} u(c) + \delta\left(\alpha u(c_1) + (1-\alpha)u(c_0)\right) \tag{2.1}$$

subject to,

$$\begin{aligned} s &\leq h \\ c &\leq w - rx + h - s \\ c_0 &\leq w + f_0(x) + s \\ c_1 &\leq w + f_1(x, \phi) + s. \\ x &\geq 0 \\ s &\geq 0 \end{aligned}$$

Assuming that the constraints on consumption bind with equality, the two first order conditions for x and s are

$$ru'(c) = \delta\left[\alpha u'(c_1)\frac{\partial f_1}{\partial x} + (1-\alpha)u'(c_0)\frac{\partial f_0}{\partial x}\right] \tag{2.2}$$

$$u'(c) = \delta\left(\alpha u'(c_1) + (1-\alpha)u'(c_0)\right). \tag{2.3}$$

Both savings and input use are chosen such that the expected marginal benefits in the future period are equal with the marginal cost in terms of foregone consumption in the present.

Application of the implicit function theorem delivers,

$$\frac{\partial x}{\partial \phi} = \frac{\delta\alpha\left[u''(c_1)\frac{\partial f_1}{\partial \phi}\left(U_{xs} - U_{ss}\frac{\partial f_1}{\partial x}\right) - U_{ss}u'(c_1)\frac{\partial^2 f_1}{\partial x \partial \phi}\right]}{U_{xx}U_{ss} - U_{xs}^2} \tag{2.4}$$

where U_{xs} is the cross-partial derivative of the objective function with respect to x and then s. The sufficient conditions for maximization deliver that $U_{ss} < 0$, $U_{xx} < 0$ and $U_{xx}U_{ss} - U_{xs}^2 > 0$. Therefore, access to Swarna-Sub1 increases use of the input if $U_{xs} - U_{ss}\frac{\partial f_1}{\partial x} < 0$. The existence of "under-investment" in the input during a flood year (i.e. $\frac{\partial f_0}{\partial x} > r$) is sufficient to generate $\frac{\partial x}{\partial \phi} > 0$.

There are two effects driving increased input use. The first is a "production effect" where the flood-tolerant variety increases the marginal product of the input under flood conditions ($\frac{\partial^2 f_1}{\partial x \partial \phi} > 0$). This effect is represented by the final term in brackets in Equation 2.4. The second is an "insurance effect", represented by the first term in square brackets. By increasing the *level* of production during flooding ($\frac{\partial f_1}{\partial \phi} > 0$) , the technology acts as insurance, hence leading to increases in input use, particularly for risk averse farmers. The insurance effect is represented as the first term in brackets.

Little is known empirically about whether the new technology enhances the marginal product of inputs during flooding. If the results are to be interpreted as driven by risk, it must be the case that the production effect is not the only channel through which the technology affects input use. This is easy to see in equation (2.4): even a farmer with a low level of risk aversion will use more inputs if $\frac{\partial^2 f_1}{\partial x \partial \phi}$ is positive and large. We show empirical evidence in sections 2.4 and 2.4 suggesting that the insurance channel is important.

A similar derivation to the above gives

$$\frac{\partial s}{\partial \phi} = \frac{\delta \alpha \left[u''(c_1) \frac{\partial f_1}{\partial \phi} \left(U_{xs} \frac{\partial f_1}{\partial x} - U_{xx} \right) + U_{xs} u'(c_1) \frac{\partial^2 f_1}{\partial x \partial \phi} \right]}{U_{xx} U_{ss} - U_{xs}^2}. \tag{2.5}$$

If the technology increases the marginal product of the input during flooding, then savings is more likely to decline after adoption of the new technology because increased input use substitutes for precautionary savings. Similar to the case of input use, increasing the level of production during flooding also crowds out savings due to the insurance effect.

Introducing credit

We do not formally model the demand for agricultural credit. Rather, we instead discuss how introducing the new technology can affect utilization of agricultural credit in a credit environment similar to the one faced by farmers in our sample.

There are two important characteristics of the credit market in our sample. First, local agricultural cooperative societies are the most popular source of credit. 45% of loans during year one of the experiment came from cooperative societies. Since cooperatives have limited resources, borrowing constraints are likely to be relevant. Second, limited liability is a feature of a substantial amount of these loans. In particular, 40% of loans from year one were renegotiated or had liability fully waived.[11]

Credit could therefore be realistically introduced into the model by allowing both borrowing constraints and limited liability. Specifically, the household borrows an amount b, where an exogenous borrowing constraint forces $b \leq \bar{b}$. Assuming that γ is the degree of

[11]Loans from agricultural cooperative societies in areas where heavy flooding occurred were nearly twice as likely to have their terms changed compared to loans from other sources. The probability of renegotiation is also increasing in the duration of flooding.

limited liability, the household must pay back $(1-\gamma)(1+v)b$ during the second period, where v is the interest rate. Since loans are most likely to be forgiven after flooding, it is plausible to assume that $\gamma = 0$ in the event that flooding does not occur.

Under this setup, there are two plausible mechanisms which could explain how introducing the new technology will influence credit usage. First, by increasing production during the flood state, the technology increases consumption, thus decreasing the marginal utility of consumption. This makes it less painful to have liabilities and therefore increases demand for credit. This effect becomes less relevant as limited liability increases because limited liability effectively acts as insurance by further increasing consumption during flooding. Second, making production less risky could induce cooperatives to make more credit available to treatment farmers – an increase in $\bar{b}$. This supply effect would increase credit utilization as long as credit constraints were binding prior to introduction of the technology.[12]

While the frequency of limited liability in our data suggests that the credit supply mechanism is most relevant, our empirical analysis does not completely distinguish between these two mechanisms.

2.4 Results

This section presents results supporting the argument that reducing risk with agricultural technology changes the management decisions of farmers. We first outline the estimation approach and then present the main results on input use, savings and credit, and productivity. We then present further evidence supporting the risk interpretation of the results. Finally, we present evidence against some key non-risk explanations.

Estimation Approach

Our main approach is to use the random distribution of Swarna-Sub1 seeds to explain management choices at both the farm and plot level. The baseline specification is therefore

$$y_{ivb} = \beta_0 + \beta_1 treatment_{ivb} + x_{ivb}\delta + \alpha_b + \varepsilon_{ivb}, \tag{2.6}$$

where y_{ivb} is an outcome observed for farmer i in village v and block b, x_{ivb} is a vector of covariates, and α_b is a fixed effect for the block, which was a stratification variable for village-level randomization. The error term is clustered at the village level since this corresponds to the first tier of randomization. We continue to use the farmer-level treatment indicator when outcomes are observed at the plot level. The estimate of β_1 in the plot-level regressions therefore represents an average effect across all plots, not just the plots cultivated with Swarna-Sub1.

[12] Boucher, Carter, and Guirkinger (2008) show theoretically that uninsured risk can induce lenders to offer loan terms that effectively crowd out a large share of the credit market. Carter, Galarza, and Boucher (2007) use simulations to suggest that weather insurance can indeed crowd in credit supply in rural Peru.

This intention to treat (ITT) estimate of β_1 is obviously attenuated if there is either significant disadoption by minikit recipients or adoption by non-recipients. Neither of these patterns are present in the data.

First, continued adoption by treatment farmers was large: 76% of minikit recipients cultivated the technology during year two.[13] The average number of plots sown with Swarna-Sub1 amongst minikit recipients was 1.5. Average land area cultivated with Swarna-Sub1 was 0.33 hectares, or approximately a third of average landholdings.

Second, there was fairly limited sharing of the seeds with non-recipients. 10.1% of non-recipients cultivated Swarna-Sub1 during year two. This was a direct results of transfers from treatment farmers, as 13.3% of non-recipients cultivated Swarna-Sub1 in treatment villages and only 3.3% did so in control villages.

These statistics suggest that our focus on ITT estimates is conservative. Instrumental variables estimates of treatment on the treated (TOT) effects would be larger by a factor of approximately 1.5. However, the ITT estimates carry the most policy relevance when dissemination of the new technology naturally leads to some dis-adoption due to imperfect targeting. Our design also allows us to drop non-recipient farmers in treatment villages to reduce the impacts of adoption spillovers. The pattern of results is similar in this smaller sample, suggesting that spillovers due to adoption by non-recipients are minimal.

Input usage

Reducing production risk with new seed varieties leads to significant extensification. Column 1 of Table 2.3 shows that access to Swarna-Sub1 led to an additional 0.68 plots being cultivated. This represents a 19% increase in the number of plots. Columns 2 and 3 show that the corresponding increase in cultivated area due to the technology is 0.1 hectares, or a 9-10% increase. Overall, the results suggest that some land gets left uncultivated due to uninsured risk and that reducing this risk brings the land into production.

Some minimal background on fertilizer use in rice is necessary in order to understand our estimates of the effect of the technology on fertilizer demand. Nitrogen, phosphorous, and potassium are the three nutrients commonly supplied to the rice plant by fertilizers. Most importantly for our analysis, these nutrients accomplish different things and are thus generally applied at different times of the growing season. In particular, phosphorous (DAP) and potassium (MOP) contribute mostly to soil conditioning and root development. In contrast, nitrogen – in the form of urea – contributes to healthy plant and leaf development and therefore is mostly applied later in the season.

Fertilizer application during year one by farmers in our sample generally followed these patterns. Figure 2.2 shows the cumulative share of the total fertilizer applied as a function

[13] The most common reason for disadoption was crop loss during year 1. Swarna-Sub1 is not suitable for low areas where water remains stagnant after floodwaters recede. Swarna-Sub1 was equally productive to Swarna in these areas.

Table 2.3: Effects on land cultivation

	(1) Number plots	(2) Rice area	(3) Log rice area
Original minikit recipient	0.675***	0.102*	0.090**
	(0.125)	(0.055)	(0.044)
ST or SC	-0.252	-0.032	-0.048
	(0.156)	(0.052)	(0.055)
HH has BPL card	-0.001	-0.147***	-0.105**
	(0.115)	(0.051)	(0.045)
HH has thatched roof	-0.391***	-0.208***	-0.207***
	(0.119)	(0.052)	(0.047)
Block Fixed Effects	Yes	Yes	Yes
Mean of Dep Variable	3.57	1.00	-0.20
Number of Observations	1235	1235	1173
R squared	0.112	0.136	0.186

Dependent variable is number of plots cultivated with rice (column 1), total rice area in ha (column 2) and log of rice area (column 3). Standard errors that are clustered at the village level are reported in parentheses. Asterisks indicate statistical significance at the 1% ***, 5% **, and 10% * levels.

of the time in days between planting and application.[14] Only 18% of urea was applied at planting (i.e. as a "basal application"), while nearly 47% of DAP and 41% of MOP were applied at the time of planting. More broadly, DAP and MOP are used earlier in the growing season.

The implication of fertilizer timing for our analysis is that fertilizers applied later should be less affected by the technology because they are not subject to losses from early floods. Floods in Odisha generally occur from the time of planting to up to 60 days before harvest. We test this idea explicitly by estimating effects separately by fertilizer type.

Table 2.4 shows that the risk-reducing technology led to an increase in overall fertilizer use. In column 2, receiving the Swarna-Sub1 minikit led to an increase in total fertilizer use of 24.6 kg, representing 11% more fertilizer. This effect is conditional on total area cultivated, which is important given the above results on extensification.[15] The results in columns 2-5 are striking. The effect is entirely concentrated on potassium and phosphorous fertilizers. This result is consistent with the technology providing flood-tolerance, but inconsistent with other explanations such as wealth effects.[16]

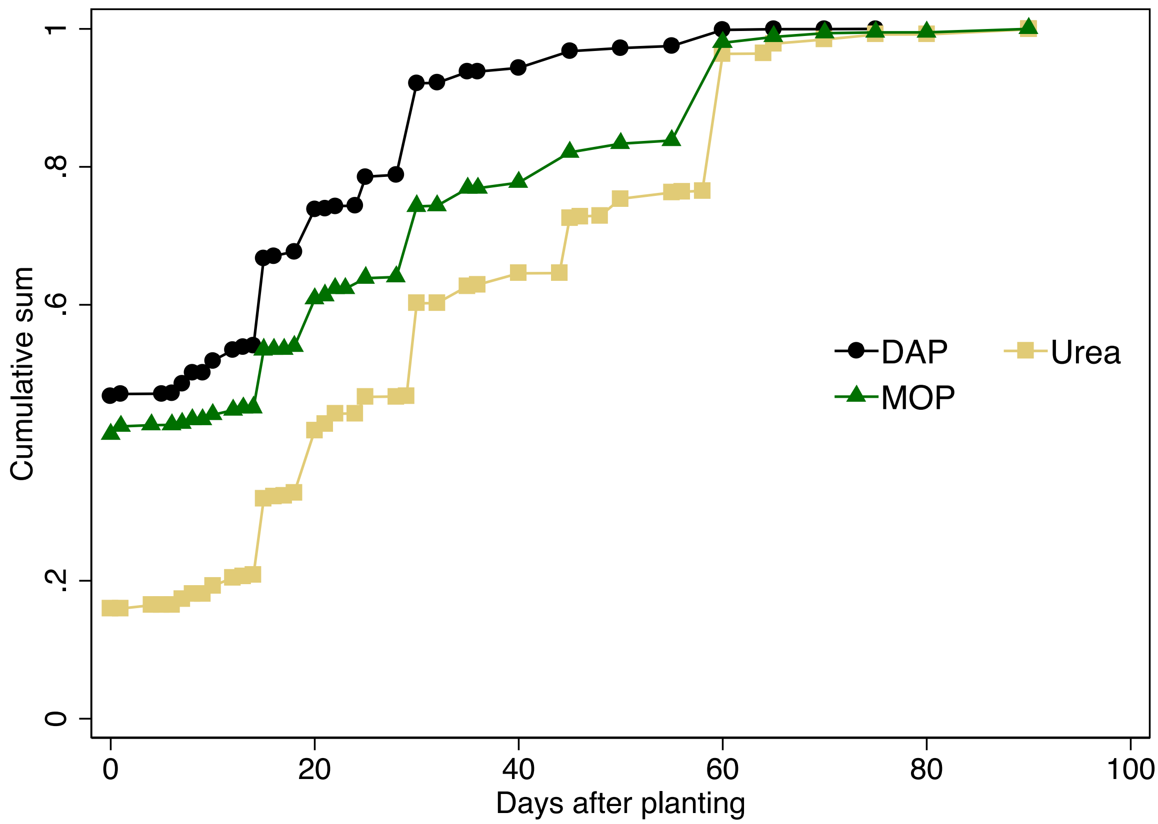
At the plot level, accessing to Swarna-Sub1 led to significant shifting away from other

[14]These data were collected for a single plot for each respondent. The chosen plot was the Swarna-Sub1 plot for treatment farmers and the largest Swarna plot for control farmers.

[15]Some farmers did not cultivate land during year 2. Values of 0 are inserted for fertilizer use and planted area for these farmers. The results are similar when using only the sample of farmers that cultivated land.

[16]Note that we include the compound fertilizer Gromor in column 5 for completeness. While we see no effect for this fertilizer, it accounts for a small share of overall use and is used by only 20% of the sample.

Figure 2.2: Timing of fertilizer applications during first year of study



Notes: Figure displays cumulative share of each fertilizer applied by each day in the growing season, where timing is measured in days after planting. Data are for farmers surveyed during the follow-up after year one. Urea is source of nitrogen (N), DAP is primary source of phosphorous (P) and MOP is the source of potassium (K).

types of seeds. Not surprisingly, column 1 of Table 2.6 shows that plots cultivated by treatment farmers were less likely to be sown with Swarna by 10.2 percentage points, an approximate 28% effect. More interestingly, treatment farmers were 4.1 percentage points (or 14.6%) less likely to choose traditional seed varieties. These traditional varieties are local varieties that were never formally released as part of the modern Green Revolution.

The crowding out of traditional seed varieties is consistent with risk motivations because traditional varieties are often chosen as an insurance strategy in areas prone to flooding and stagnant water. These varieties have an increased ability to survive due to their ability to rapidly elongate when submerged (Voesenek and Bailey-Serres, 2009).[17] The downside

[17]We have yet to distinguish between flash flooding, where Swarna-Sub1 performs well, and stagnant water accumulation. Flash flood areas are those where flooding occurs and water recedes after a period of

Table 2.4: Effects on aggregate fertilizer use

	(1) All	(2) Urea	(3) DAP	(4) MOP	(5) Gromor
Original minikit recipient	24.622**	3.329	18.347***	5.940*	-2.993
	(10.002)	(4.235)	(5.282)	(3.213)	(2.637)
Rice area (hectares)	217.515***	86.801***	89.468***	32.445***	8.802***
	(19.379)	(13.401)	(9.272)	(7.027)	(2.567)
ST or SC	-17.964*	-3.952	-5.363	-5.040*	-3.609
	(9.574)	(3.748)	(4.579)	(2.947)	(2.602)
HH has BPL card	-1.348	-0.933	-4.993	1.480	3.098
	(9.021)	(4.041)	(4.306)	(2.440)	(2.456)
HH has thatched roof	-18.801**	-7.489**	0.247	-6.292**	-5.267**
	(8.400)	(3.572)	(4.502)	(2.653)	(2.481)
Block Fixed Effects	Yes	Yes	Yes	Yes	Yes
Mean of Dep Variable	215.49	83.33	79.99	37.15	15.01
Number of Observations	1235	1235	1235	1235	1235
R squared	0.615	0.508	0.518	0.288	0.074

Dependent variable is total fertilizer use (across all rice plots), measured in kg. Column labels indicate type of fertilizer. Standard errors that are clustered at the village level are reported in parentheses. Asterisks indicate statistical significance at the 1% ***, 5% **, and 10% * levels.

of this strategy is that the yield of traditional varieties is significantly lower under normal conditions. This is shown in Figure 2.3. During year one, traditional varieties were lower yield, but significantly more likely to survive when submerged for more than one week.[18] Consistent with this, Figure 2.4 shows that ability to survive in flooding and stagnant water are two of the frequently stated reasons why farmers choose traditional varieties. A plausible explanation of the result is therefore that Swarna-Sub1 substitutes for traditional varieties in providing insurance against losses during flooding.

Returning to column 3 of Table 2.6, farmers relied significantly less on the broadcasting planting method after receiving Swarna-Sub1. Fields cultivated by treatment farmers were 6.3 percentage points (33.1%) less likely to use broadcasting. Consequently, 4% fewer seeds were used because broadcasting has a larger seed requirement (column 4). Broadcasting involves manually throwing dry seeds on puddled soil. This contrasts with transplanting – the other popular method. The transplanting method involves raising seedlings on a small portion of land, pulling and bundling the seedlings after approximately three weeks,

one day to around two weeks. Stagnant water areas are those where water remains partially submerging the crop even after floodwaters recede. Our findings in Dar et al. (2013) and other research (i.e. Singh, Mackill, and Ismail (2011)) show that Swarna-Sub1 does not tolerate stress due to stagnant water of more than approximately 15 days.

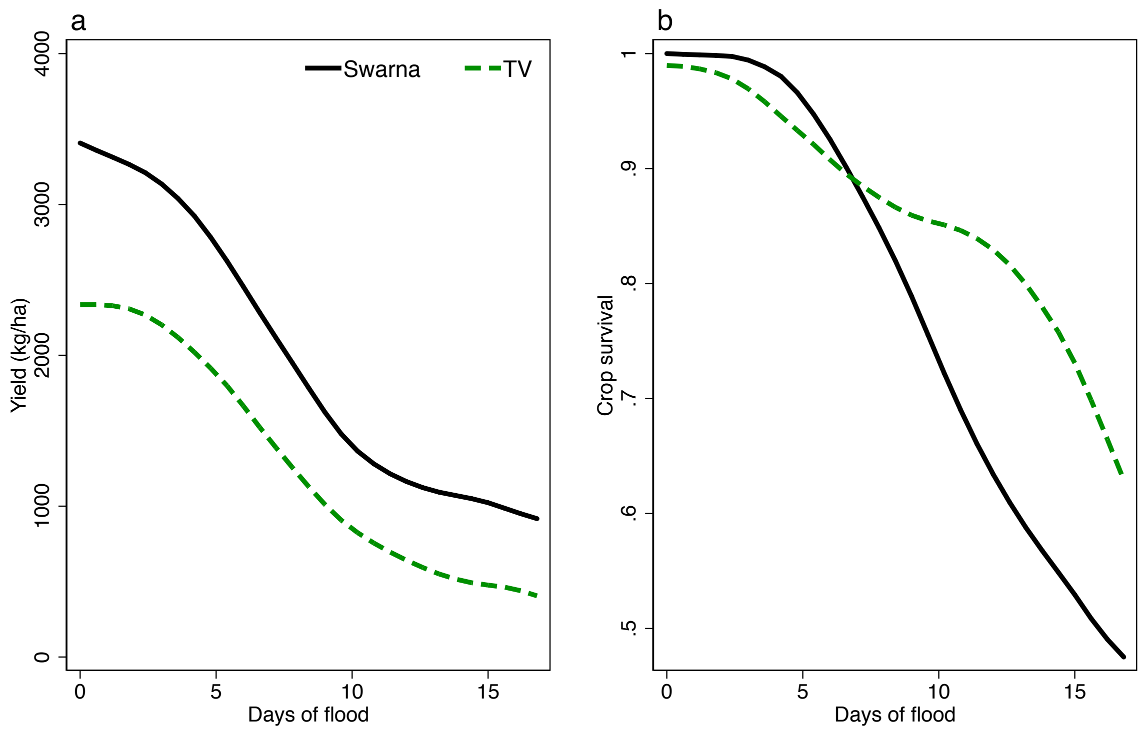
[18]We show in Table 2.5 that the declining yield gap with flood duration is statistically significant. Also, the increase in survival probability is significantly larger as flooding worsens.

Table 2.5: Relationship between variety type and crop performance, Kharif 2011

	Yield (kg/ha)		Survival(0/1)	
	(1)	(2)	(3)	(4)
Days flood	-92.095***	-83.440***	-0.019***	-0.017***
	(18.198)	(17.138)	(0.005)	(0.005)
Traditional variety	-973.533***	-751.871***	-0.031	0.015
	(202.782)	(166.845)	(0.044)	(0.030)
Traditional variety*Days flood	55.267***	51.555***	0.018**	0.017**
	(18.358)	(17.022)	(0.007)	(0.007)
Swarna-Sub1	-118.004	-185.347	-0.022	-0.025
	(161.621)	(160.159)	(0.016)	(0.018)
Swarna-Sub1*Days flood	63.887***	71.113***	0.004	0.004
	(21.945)	(22.339)	(0.004)	(0.004)
Other modern variety	-348.682***	-288.421**	-0.021*	-0.015
	(119.816)	(111.699)	(0.012)	(0.012)
Other modern variety*Days flood	21.332*	18.919*	-0.000	-0.001
	(10.951)	(10.622)	(0.003)	(0.003)
Broadcasted		-294.999*		-0.083
		(159.815)		(0.068)
Irrigated		517.381***		0.029*
		(148.531)		(0.015)
Low land		108.668		-0.023
		(126.913)		(0.028)
Medium land		351.455***		0.013
		(123.564)		(0.025)
Constant	3500.883***	2988.497***	1.032***	1.011***
	(294.703)	(282.359)	(0.025)	(0.028)
Block Fixed Effects	Yes	Yes	Yes	Yes
Mean of Dep Variable	2214.26	2221.76	0.88	0.89
Number of Observations	4182	4138	4182	4138
R squared	0.451	0.471	0.332	0.333

Dependent variable in Columns 1 and 2 is rice yield in kg/ha. Dependent variable in Columns 3 and 4 is an indicator for nonzero yield. Data are for the first year of the experiment (Kharif 2011). Standard errors are clustered at the village level. Asterisks indicate statistical significance at the 1% ***, 5% **, and 10% * levels.

Figure 2.3: Nonparametric regressions of yield and crop survival on duration of flooding during year one



Notes: Figure displays fan regressions of yield and crop survival (0/1) on duration of submergence. Estimates are for year one when flooding occurred in part of the sample area.

and planting the seedlings manually in the main field. While broadcasting is a popular planting technique in flood prone areas due to the decreased labor investment, the increased competition from weeds reduces yields (Khush, 1997; Rao et al., 2007).

Decreased reliance on broadcasting is also consistent with a risk explanation. Since transplanting requires a significant labor investment at the time of planting, it carries substantial risk that the investment will not pay off in the event of flooding. The new technology reduces this risk and induces farmers to opt for the costlier, but more productive method.

Column 5 shows that while minikit recipients were slightly more likely to irrigate their fields, the effect is not statistically significant. It is also noteworthy that 74% of plots were irrigated during year two while only approximately 20% of plots were reported to have irrigation during year one. This is attributable to use of supplemental sources when rainfall is not sufficient (i.e. electric or diesel pumps from rivers/canals).

Providing access to the risk-reducing technology also led to a decrease in the likelihood of taking land out of production – a result that likely drives part of the result on overall

Figure 2.4: Stated reasons for choosing rice varieties during year 2

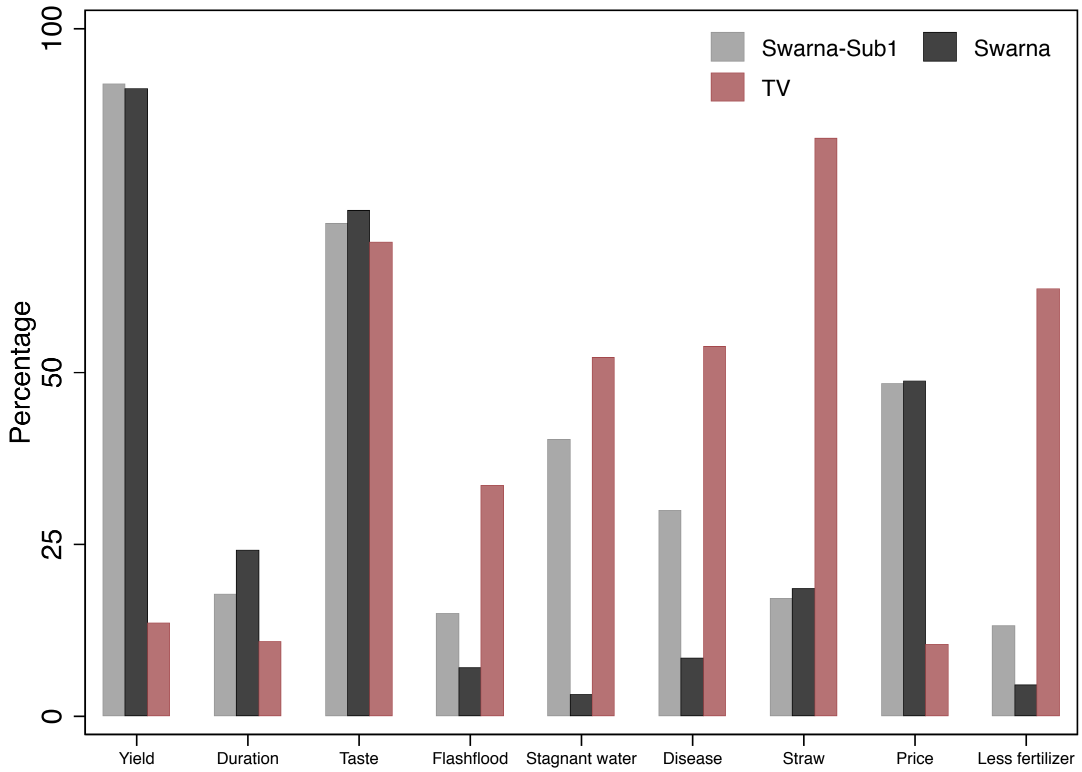


Figure displays percentage of farmers (Swarna and Swarna-Sub1) and percentage of farmer-variety pairs (TV) for which the characteristic on the horizontal axis is a reason the variety was chosen during year 2. For example, over 90% of farmers cultivating Swarna stated that high yield was one of the reasons for this choice (first grey bar above yield).

area cultivated. Columns 6 shows that access to Swarna-Sub1 caused farmers to be around 2.3 percentage points – or 28% – less likely to take land out of production. While the heterogeneity in column 7 suggests that the effect is larger on plots considered to be low land, the interaction term is not statistically significant (p=0.15). Column 8 shows that the effect is significantly larger on plots that farmers consider to be below average in terms of land quality.[19]

Savings and credit

Farmers in the sample save excessively large amounts of rice after each harvest. The average total rice harvest amongst cultivators in our sample was 2,945 kg. An average of 1,711 of

[19]This result suggests that any differences in output per hectare we observe are unlikely attributable to composition effects. If anything, the additional land cultivated by treatment farmers is *lower quality*.

Table 2.6: Effects on plot-level production practices

	(1) Use Swarna	(2) Use TV	(3) Broadcast	(4) ln(seedrate)	(5) Irrigated	(6) Not cult.	(7) Not cult.	(8) Not cult.
Original minikit recipient	-0.102***	-0.041**	-0.063***	-0.040**	0.036	-0.023*	-0.018	-0.018*
	(0.017)	(0.016)	(0.017)	(0.019)	(0.029)	(0.012)	(0.014)	(0.010)
ST or SC	0.041	-0.008	-0.029	-0.001	-0.031	-0.013	-0.012	-0.010
	(0.025)	(0.022)	(0.028)	(0.024)	(0.041)	(0.016)	(0.016)	(0.015)
HH has thatched roof	-0.020	0.023	0.009	-0.008	0.025	0.004	0.007	0.005
	(0.017)	(0.019)	(0.023)	(0.021)	(0.028)	(0.013)	(0.013)	(0.012)
HH has BPL card	-0.020	0.013	0.020	-0.028	-0.026	-0.006	-0.004	-0.003
	(0.018)	(0.020)	(0.021)	(0.024)	(0.029)	(0.013)	(0.013)	(0.012)
Minikit*Low land							-0.030	
							(0.020)	
Low land							0.016	
							(0.016)	
Minikit*Low quality land								-0.072**
								(0.036)
Low quality land								0.120***
								(0.025)
Block Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Mean of Dep Variable	0.36	0.28	0.19	4.16	0.74	0.08	0.07	0.07
Number of Observations	4578	4577	4571	4569	4546	5057	5036	5001
R squared	0.116	0.270	0.243	0.198	0.080	0.020	0.021	0.038

Estimation data are at the plot level. Dependent variables are an indicator for sowing plot with Swarna (column 1), an indicator for sowing plot with a traditional rice variety (column 2), an indicator for planting using the broadcasting technique (column 3), logarithm of kilograms seed used per hectare (column 4), an indicator for irrigation (column 5), and indicator for plot not being cultivated (columns 6-8). Standard errors that are clustered at the village level are reported in parentheses. Asterisks indicate statistical significance at the 1% ***, 5% **, and 10% * levels.

the harvest was consumed or set aside for future consumption.[20] This amount is enough to feed roughly 11 adults.[21] Average household size in our sample is 7 persons. While we acknowledge that stored rice is a liquid asset that can be sold upon family need, it does appear that a precautionary motive can be behind the excessive amounts of rice being set aside for consumption.

Farmers receiving Swarna-Sub1 save a smaller share of their harvest for future consumption. Column 1 in Table 2.7 shows that treatment farmers save 5 percentage points less output for consumption. The magnitude of the decrease is not trivial. It amounts to approximately 150 kilograms, or enough to meet the consumption needs of one household member for the entire year. Column 2 shows that this effect is concentrated amongst households that *do not* hold Below the Poverty Line (BPL) cards. BPL cards serve as consumption insurance because they entitle households to purchase 30 kilograms of rice per month at highly subsidized rates, an amount that is enough to feed approximately two adults.[22] Therefore, the savings result is consistent with reduced incentives for precautionary savings. In Column 3 we show that the decrease in savings corresponded with an increase in selling the harvest as output.

In addition to decreasing precautionary savings, reducing risk by adding flood tolerance leads to increased uptake of agricultural credit. Columns 4-8 of Table 2.7 focus on binary indicators of household access to credit. Column 4 shows that Swarna-Sub1 led to a 6.8 percentage point increase in the probability of having any credit during the growing season, which represents a 36% increase. We disaggregate loans by timing in columns 5 and 6. Most of the effect is driven by early loans during the first four months of the growing season. Columns 7 and 8 show that a large portion of the effect is driven by loans originating from agricultural cooperatives.

This credit result contrasts with the literature on insurance. Giné and Yang (2009) show that access to rainfall insurance led to decreased take-up of credit for the purchase of hybrid maize in Malawi. While not focused on agricultural insurance, Banerjee, Duflo, and Hornbeck (2014) have a similar finding that adding mandatory health insurance decreased the demand for microfinance in India. Finally, Karlan et al. (2013) find no effect of insurance on borrowing behavior amongst their sample of farmers in Ghana. The technology we study is therefore a unique case where reducing risk crowds in credit.

In contrast to inputs, the results on savings and credit are unlikely to be explained by shifts in the marginal productivity of inputs. Rather, savings of grain offers insurance against low or zero production during flooding. By increasing the overall level of production during flooding, the technology substitutes for precautionary savings. Utilization of agricultural credit increases either due to a demand effect where the increase in expected production in

[20]Since our survey was conducted shortly after the harvest and post-harvest production practices, most of the amount indicated for consumption had yet to be consumed at the time of the survey.

[21]This calculation is based on an average annual consumption in rural Odisha of 158 kgs per year, as reported in the 64th round of the National Sample Survey.

[22]The price of BPL rice is 1-2 Rupees per kilogram while the market price of similar rice is 20 Rupees or higher.

Table 2.7: Effects on precautionary savings and credit access

	Share of harvest:			Credit				
	(1) Saved	(2) Saved	(3) Sold	(4) All	(5) Early	(6) Late	(7) From cooperative	(8) Other source
Original minikit recipient	-0.050***	-0.087***	0.045***	0.068**	0.054**	0.017	0.050**	0.023
	(0.017)	(0.026)	(0.016)	(0.027)	(0.025)	(0.017)	(0.024)	(0.019)
ST or SC	-0.014	-0.014	-0.036**	-0.048**	-0.050**	0.008	-0.033*	-0.007
	(0.019)	(0.020)	(0.016)	(0.023)	(0.024)	(0.020)	(0.019)	(0.023)
HH has BPL card	-0.001	-0.018	-0.020	-0.032	-0.013	-0.016	-0.020	-0.016
	(0.018)	(0.020)	(0.015)	(0.024)	(0.022)	(0.015)	(0.021)	(0.016)
HH has thatched roof	0.032*	0.032*	-0.067***	-0.014	-0.032	0.018	-0.035*	-0.002
	(0.019)	(0.019)	(0.016)	(0.024)	(0.022)	(0.014)	(0.021)	(0.015)
Original minikit recipient*HH has BPL card		0.067**						
		(0.030)						
Block Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Mean of Dep Variable	0.70	0.70	0.17	0.19	0.14	0.05	0.12	0.08
Number of Observations	1165	1165	1165	1235	1235	1235	1228	1235
R squared	0.073	0.077	0.122	0.063	0.062	0.028	0.063	0.016

Dependent variable is share of rice harvest consumed or saved for consumption (columns 1 and 2), share of rice harvest sold (column 3), indicator for taking a new loan during the 2012 agricultural season (column 4), indicator for taking a loan during the first 4 months of the season (column 5), indicator for taking a loan during the remaining 5 months of the season (column 6), indicator for taking loan from agricultural cooperative (column 7) and indicator for taking loan from another source (column 8). Other sources are banks, input sellers, Self-Help groups (SHG's), MFI's, friends/relatives, or money lenders. Standard errors that are clustered at the village level are reported in parentheses. Asterisks indicate statistical significance at the 1% ***, 5% **, and 10% * levels.

the future makes borrowing more desirable, or a supply effect where credit constraints are a function of the borrower's expected production.

Indirect effects on productivity

The results indicate strongly that reducing risk causes several changes in management. A natural next question is to ask how these changes translate to changes in productivity.

There is a noticeable increase in yield for plots cultivated by farmers with access to Swarna-Sub1. Figure 2.5 displays the estimated density of yield by treatment status. There is a clear rightwards shift in the distribution for treatment farmers. In addition, the effect occurs throughout the distribution of yield, suggesting that it is not concentrated on the largest or smallest farmers. The effect is quantified in the first column of Table 2.8. The average yield effect is 283 kilograms per hectare. This represents a 10% increase and also is about 45% the magnitude of the maximum yield advantage we observed in heavily flooded areas during year one of the study. Importantly, the effect is not driven by agronomic properties of the seed because these properties are identical to those of Swarna when there is no flooding.

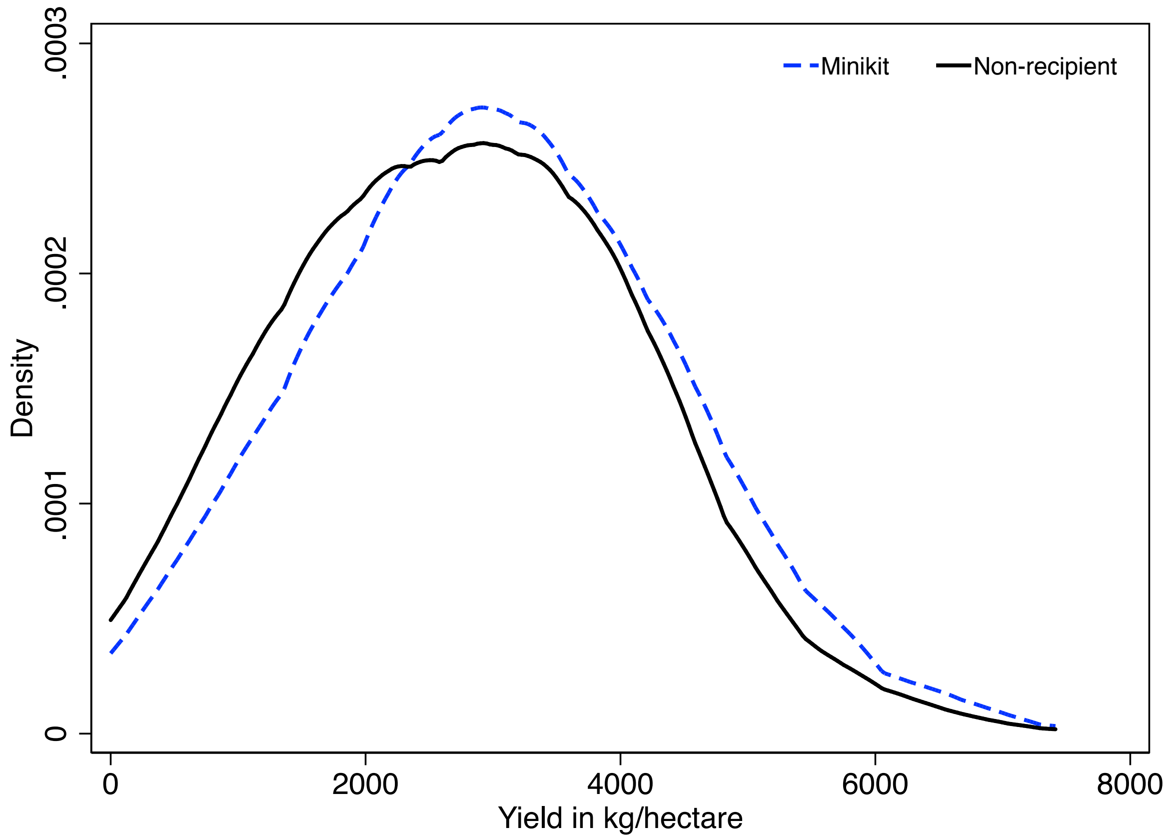
As a method of investigating whether the management changes we observe are channels through which reducing risk affects productivity, we sequentially add these outcome measures to the yield equation. Adding these endogenous outcomes as controls should attenuate the reduced form productivity effect if the effect is partly operating through these channels. This approach has been used to investigate the channels through which education affects voting (Milligan, Moretti, and Oreopoulos, 2004) and the channels through which early-life rainfall shocks affect outcomes later in life (Maccini and Yang, 2009). Of course, the coefficients on the controls represent merely correlations and can not be interpreted causally.

The conditional productivity effect is substantially lower than the unconditional effect. Columns 2-5 of Table 2.8 show that accounting for the main sources of behavioral change reduce the productivity effect. Specifically, the overall conditional effect in column 5 decreases by 40% relative to the unconditional effect in column 1. In addition, the goodness of fit of the regression approximately doubles. Therefore, the changes in management that we observe appear to explain part of the overall productivity effect.

Effects on fields not cultivated with Swarna-Sub1

Returning to the theoretical model, we now investigate whether the production effect can possibly be the only explanation for the results. We estimate effects on yield, planting technique, and fertilizer use at the plot level for the plots that were not cultivated with Swarna-Sub1. If the results are explained by how the technology changes the marginal productivity of these inputs, then there is no reason to observe effects for these plots. Conversely, if the risk channel is important then some effect should persist on these plots because this channel is driven by how the technology increases the overall level of production during flooding.

Figure 2.5: Kernel density of plot-level yield by treatment status



Notes: Figure displays estimated kernel densities of yield (kg/ha) for plots cultivated during year two of the study. Density for minikit recipients is across all plots, regardless of choice of variety.

While the effects are smaller, the effects on productivity, broadcasting, and fertilizer use all persist on plots where Swarna-Sub1 was not used. Column 1 of Table 2.9 shows a yield effect of 170 kilograms per hectare, or approximately 6%. Also, as shown in column 2, there is a 3.6 percentage point decrease in the probability of using broadcasting. Finally, the fertilizer effect also persists on non-Sub1 plots (column 3).[23] These results are inconsistent with the new technology being more responsive to fertilizer or the transplanting method. Instead, the results provide some suggestive evidence that is more consistent with the risk explanation.

[23]The results are complicated by plot selection issues if farmers place Swarna-Sub1 on land that is systematically different. We find that controlling for stated land quality and whether the plot is low land – two factors that are likely important in this decision – does not influence the results in Table 2.9.

Table 2.8: Effects on productivity

	(1)	(2)	(3)	(4)	(5)
Original minikit recipient	283.449***	230.301***	196.545***	180.561***	167.732**
	(77.484)	(73.730)	(68.059)	(66.174)	(64.672)
Broadcast planting		-801.222***	-679.361***	-414.951***	-416.833***
		(129.450)	(117.529)	(116.348)	(116.084)
Tons fertilizer per hectare			4350.392***	3305.200***	3263.775***
			(997.705)	(824.647)	(829.248)
Tons fertilizer per hectare^2			-4025.260**	-2984.097**	-2967.389**
			(1628.518)	(1259.866)	(1275.298)
Traditional variety				-434.933***	-436.087***
				(70.351)	(70.411)
Irrigated				713.510***	708.489***
				(91.662)	(91.979)
Log seed rate				1.878	-12.881
				(114.168)	(114.812)
Has credit					150.020**
					(70.133)
Block Fixed Effects	Yes	Yes	Yes	Yes	Yes
Mean of Dep Variable	2817.97	2819.53	2819.53	2821.17	2821.17
Number of Observations	4573	4568	4568	4507	4507
R squared	0.159	0.200	0.236	0.300	0.302

Estimation data are at the plot level. Dependent variable in all regressions is yield in kg/hectare. All independent variables are measured at the plot level, except for fertilizer per hectare, which is measured at the farmer level. The regression in column 6 is limited to data from all plots except those cultivated with Swarna-Sub1. Standard errors that are clustered at the village level are reported in parentheses. Asterisks indicate statistical significance at the 1% ***, 5% **, and 10% * levels.

Heterogeneity according to stated risk preferences

As an additional test of whether the risk-reducing property of the technology is important for our results, we investigate heterogeneity with respect to risk preferences. We elicited risk preferences during the first follow-up by providing respondents with a sequence of hypothetical choices between two lotteries. Each choice involved a fixed relatively "safe" option and a riskier option with higher expected return. We define the highly risk averse as those selecting the safe option over the risky option with the lowest expected return.[24]

[24] The safe option was a 50-20 lottery. The lowest return riskier option was 60-15. The expected value of the safe lottery (35 Rs) represents about 1/20th of monthly consumption per capita in rural Odisha. Assuming CRRA preferences, the implied range for the coefficient of relative risk aversion is $(2.04, \infty)$. For this calculation, baseline consumption was set to half of the minimum daily wage. We must also be careful to

Table 2.9: Effects estimated for sample of fields not cultivated with Swarna-Sub1

	(1) Yield	(2) Broadcast	(3) Fertilizers
Original minikit recipient	169.853**	-0.036*	4.926**
	(77.557)	(0.019)	(2.437)
ST or SC	-154.534	-0.033	-1.375
	(111.273)	(0.030)	(2.208)
HH has thatched roof	-134.968	0.010	-0.290
	(86.014)	(0.024)	(1.924)
HH has BPL card	-23.595	0.026	-2.671
	(77.930)	(0.024)	(1.890)
Area of field			43.797***
			(4.708)
Block Fixed Effects	Yes	Yes	Yes
Mean of Dep Variable	2757.71	0.20	33.96
Number of Observations	4091	4086	4091
R squared	0.17	0.26	0.41

Data consist of entire sample of plots not cultivated with Swarna-Sub1 during the second year of the study. The dependent variables are rice yield in kg/hectare (column 1), an indicator for planting using the broadcasting technique (column 2), and total kilograms of DAP and MOP fertilizers used (column 3). Standard errors that are clustered at the village level are reported in parentheses. Asterisks indicate statistical significance at the 1% ***, 5% **, and 10% * levels.

Some of the effects are larger for the population that is highly risk averse. Table 2.10 shows our main estimates where the treatment indicator is interacted with an indicator for high risk aversion. While power is limited, the effects are larger for the highly risk averse for all outcomes except for fertilizer use and credit utilization. However, the interaction term between minikit receipt and high risk aversion is only statistically significant for planting method. It is however noteworthy that the effects on number of plots cultivated, planting method, and grain savings are much larger for the highly risk averse and it is these three outcomes where there is a statistically significant correlation between high risk aversion and the outcome.

Alternative Explanations

A plausible alternative explanation of our results is that the provision of the new seed variety during year one led to a boost in agricultural income which then affected decision-making during the following year. One aspect of our design that minimizes the likelihood of this

only interpret the results suggestively since risk preferences were collected after treatment was administered and the lotteries were not incentivized.

Table 2.10: Heterogeneity of main effects according to risk preferences

	(1) Number plots	(2) Fertilizer	(3) Use TV	(4) Broadcast	(5) Share saved	(6) Credit
Original minikit recipient	0.442**	29.340**	-0.033	-0.005	-0.030	0.106**
	(0.188)	(14.031)	(0.028)	(0.031)	(0.024)	(0.041)
Original minikit recipient*High risk aversion	0.393	-9.403	-0.013	-0.102**	-0.031	-0.070
	(0.268)	(20.746)	(0.034)	(0.046)	(0.034)	(0.054)
ST or SC	-0.255	-17.654*	-0.008	-0.027	-0.014	-0.046**
	(0.156)	(9.443)	(0.022)	(0.028)	(0.019)	(0.023)
HH has BPL card	-0.004	-1.635	0.013	0.020	-0.000	-0.033
	(0.114)	(8.948)	(0.020)	(0.021)	(0.018)	(0.024)
HH has thatched roof	-0.402***	-18.498**	0.024	0.012	0.032*	-0.011
	(0.121)	(8.596)	(0.019)	(0.023)	(0.019)	(0.024)
High risk aversion	-0.233*	-2.847	0.005	0.048**	0.036**	0.010
	(0.135)	(10.418)	(0.020)	(0.021)	(0.017)	(0.024)
Rice area (hectares)		217.522***				
		(19.430)				
Block Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Mean of Dep Variable	3.57	215.49	0.28	0.19	0.70	0.19
Number of Observations	1235	1235	4577	4571	1165	1235
R squared	0.115	0.615	0.270	0.247	0.077	0.065

Dependent variable is number of rice plots (column 1), total fertilizer use in KG (column 2), indicator for use of traditional variety on plot (column 3), indicator for plot being planted using the broadcasting technique (column 4), share of total rice harvest consumed or saved for future consumption (column 5), and access to credit (column 6). Standard errors that are clustered at the village level are reported in parentheses. Asterisks indicate statistical significance at the 1% ***, 5% **, and 10% * levels.

channel is the small amount of seed that was provided in the minikits. Treatment farmers cultivated an average of 10-15% of their land with Swarna-Sub1 during year one of the study. The average yield gain across all areas during year one was 10%. Combining these two facts, the overall effect on aggregate income after the first year was clearly small.

To more directly rule out wealth effects, we use data on the total amount of rice harvested during the first year as a measure of the wealth that would have been influenced by being treated. If increases in wealth due to the new technology led to the behavioral changes we observe during year two, then controlling for the year one harvest should attenuate our main estimates. We show in Table 2.11 that none of our main estimates are affected when conditioning on the year one total harvest. This is consistent with the minikits providing only a small amount of seed for testing and not a substantial increase in income.

Are differences in output prices responsible for our results? We collected information on prices received by variety from each farmer that sold any output after the second year of the study. The average price received for Swarna was 10.29 Rupees per kilogram and the average price received for Swarna-Sub1 was 10.76.[25] This 4.6% difference in prices is statistically significant at the 5% level. Since the eating quality of the two varieties is similar (see Figure 2.4), this difference in prices could possibly be due to increased value of output as seed rather than grain for consumption.

There is no evidence that output prices drive the results. Descriptively, only 40% of farmers sold any rice following the second year and on average only 17% of the harvest was sold as grain for consumption, suggesting that effects of output price differences are likely to be small. We consider this possibility further by separate estimation of the main results for the subsample of farmers that did not sell *any* rice following the second year.[26] If prices are explaining the results, then the effects of the technology should be smaller in this sub-sample. We show in Table 2.12 that this is not the case. This evidence is not consistent with output prices being the relevant channel for our results.

2.5 Conclusion

This chapter has shown the importance of risk as a barrier that impedes investment by small farmers in the developing world. In short, uninsured risk induces farmers to make safe decisions that result in reduced income variability at the cost of lower productivity. Therefore, our work shows that despite the efforts of governments to help farmers deal with risk, there is substantial room for new instruments that protect farmers from weather-induced variability in production.

In addition to identifying risk as a key barrier, we have presented the first evidence on how technological progress can serve as one of these instruments. Specifically, adding flood tolerance to a commonly used seed variety induced farmers to cultivate more land, increase fertilizer use, rely more on a costly but more productive planting method, and increase the uptake of agricultural credit. Importantly, these changes increased productivity by 10%, indicating the importance of risk as a barrier to agricultural development.

Our results suggest that efforts to reduce the susceptibility of commonly used seeds to weather extremes can go a long way in increasing agricultural productivity. While the first major Green Revolution was successful at increasing agricultural productivity throughout the world, production risk increased with the spread of modern Green Revolution seed varieties (Feder, Just, and Zilberman, 1985; Foster and Rosenzweig, 2010). Using flood tolerance as

[25] The government's paddy procurement program set the minimum support price (MSP) for the 2012 harvest at 12.5 Rupees per kilogram. Many farmers in our sample sell instead to private traders at prices below this level.

[26] A problem with this approach is that the sample is being split according to an endogenous outcome. The most plausible effect of sample selection in this case is that the group of farmers selling rice are the largest and wealthiest farmers that have the most capacity to respond after having access to Swarna-Sub1. For instance, average landholdings of farmers that do not sell output are 50% less than those selling output. This would then work against us finding any effects in the subsample of farmers not selling output.

Table 2.11: Estimation of main results conditioning on 2011 (year 1) kharif rice harvest

	Plot level				Farmer level			
	(1) Yield	(2) Use TV	(3) Broadcast	(4) Not cult.	(5) Fertilizer	(6) Number plots	(7) Share saved	(8) Credit
Original minikit recipient	275.834***	-0.039**	-0.061***	-0.024*	21.536**	0.626***	-0.042**	0.065**
	(77.658)	(0.016)	(0.017)	(0.012)	(10.146)	(0.123)	(0.017)	(0.027)
2011 rice harvest (tons)	21.122	-0.005	-0.004	0.001	14.730**	0.142***	-0.022***	0.007
	(16.863)	(0.004)	(0.003)	(0.003)	(6.212)	(0.049)	(0.004)	(0.006)
ST or SC		-0.010	-0.030	-0.013	-14.342	-0.211	-0.019	-0.046*
		(0.022)	(0.029)	(0.016)	(10.283)	(0.158)	(0.019)	(0.023)
HH has thatched roof		0.021	0.007	0.004	-16.305*	-0.327***	0.022	-0.010
		(0.019)	(0.023)	(0.013)	(8.309)	(0.114)	(0.018)	(0.024)
HH has BPL card		0.011	0.018	-0.006	1.544	0.054	-0.009	-0.029
		(0.020)	(0.022)	(0.013)	(8.803)	(0.113)	(0.017)	(0.024)
Rice area (hectares)					197.895***			
					(21.294)			
Block Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Mean of Dep Variable	2817.97	0.28	0.19	0.08	215.49	3.57	0.70	0.19
Number of Observations	4573	4577	4571	5057	1235	1235	1165	1235
R squared	0.161	0.271	0.244	0.020	0.626	0.134	0.104	0.065

Estimation data are at the plot level in columns 1-4 and the farmer level in columns 5-8. Dependent variables are rice yield in kg/ha (column 1), an indicator for sowing plot with a traditional rice variety (column 2), an indicator for planting using the broadcasting technique (column 3), an indicator for plot not being cultivated (column 4), total fertilizer use (column 5), number of plots cultivated with rice (column 6), share of harvest consumed or saved for consumption (column 7), and indicator for access to credit (column 8). Standard errors that are clustered at the village level are reported in parentheses. Asterisks indicate statistical significance at the 1% ***, 5% **, and 10% * levels.

Table 2.12: Estimation of main results for sample that did not sell rice following year 2 harvest

	Plot level				Farmer level		
	(1) Yield	(2) Use TV	(3) Broadcast	(4) Not cult.	(5) Fertilizer	(6) Number plots	(7) Credit
Original minikit recipient	465.781***	-0.031	-0.071**	-0.026	8.574	0.664***	0.067**
	(102.602)	(0.020)	(0.032)	(0.021)	(8.923)	(0.159)	(0.030)
ST or SC		0.013	-0.033	-0.049**	-4.226	0.062	-0.051**
		(0.026)	(0.031)	(0.023)	(7.988)	(0.152)	(0.025)
HH has thatched roof		0.000	0.023	-0.023	-20.018**	-0.104	0.009
		(0.028)	(0.026)	(0.021)	(8.661)	(0.149)	(0.027)
HH has BPL card		-0.010	0.013	-0.020	6.524	-0.057	0.011
		(0.021)	(0.027)	(0.022)	(7.434)	(0.128)	(0.026)
Rice area (hectares)					139.609***		
					(17.306)		
Block Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Mean of Dep Variable	2362.44	0.30	0.22	0.12	137.92	2.97	0.13
Number of Observations	2297	2302	2297	2696	747	747	747
R squared	0.138	0.326	0.298	0.026	0.402	0.116	0.057

Data are limited to 749 farmers that did not sell any rice after the kharif 2012 harvest. Estimation data are at the plot level in columns 1-4 and the farmer level in columns 5-7. Dependent variables are rice yield in kg/ha (column 1), an indicator for sowing plot with a traditional rice variety (column 2), an indicator for planting using the broadcasting technique (column 3), an indicator for plot not being cultivated (column 4), total fertilizer use (column 5), number of plots cultivated with rice (column 6), and indicator for access to credit (column 7). Standard errors that are clustered at the village level are reported in parentheses. Asterisks indicate statistical significance at the 1% ***, 5% **, and 10% * levels.

an example, we have shown that technological advances that improve these varieties can further enhance agricultural productivity by inducing farmers to take more risk.

At the same time, the technology we have studied here addresses flooding – only one type of risk faced by farmers. The new technology is equally susceptible to drought and it obviously does not account for other non-weather shocks. Nonetheless, simply reducing this one type of production risk causes farmers to make decisions that lead to productivity gains.

Chapter 3

The efficiency of informal seed exchange

3.1 Introduction

The identities of agents are usually considered to be irrelevant in the classic marketplace because buyers and sellers come together at "arm's length" to make efficient transactions. While this abstract definition of the marketplace constitutes the ideal textbook scenario, a broad set of goods are exchanged bilaterally between agents that are connected in networks (Jackson, 2009). This broad set includes informal insurance in numerous contexts (Fafchamps and Lund, 2003; De Weerdt and Dercon, 2006; Mazzocco, 2012; Attanasio et al., 2012), electronics in Japan (Nishiguchi, 1994), and fish in southern France (Vignes and Etienne, 2011).

This chapter asks whether trading in networks – a common nonmarket institution – can allocate a new technology efficiently. In the absence of capacity constraints, the specific notion of efficiency is adoption by all potential buyers that have positive expected benefits from using the technology. Despite the importance of networks as a mode of exchange, the growing theoretical literature on network-based exchange (Kranton and Minehart, 2001; Elliott, 2013), and several laboratory experiments (Gale and Kariv, 2009; Cassar, Friedman, and Schneider, 2010), there is little evidence from the field on how effectively networks allocate goods (Jackson and Zenou, 2013). I present the first field experiment to measure whether network-based exchange allocates a product to everyone with demand. Ex-ante, the answer to the question is uncertain. On the one hand, the costs of adopting from suppliers coming from different social groups may create a friction and limit exchange to closely linked individuals (Elliott, 2013). Conversely, if demand is competitive, then buyers with high valuations of the technology may be induced to bear the costs of making links with sellers (Kranton and Minehart, 2001).

I exploit the flood tolerance property of Swarna-Sub1 to characterize ex-ante the potential adopters with the highest expected returns. Given that flood severity depends heavily on local topography, this flood tolerance property creates variation in expected benefits that can be used to compare the efficiency of different allocations.

While I rely on this particular agricultural technology, its most important feature is that flood exposure – the key determinant of returns – is observable. This property creates a rare opportunity to characterize the relative efficiency of different modes of exchange. Most technologies simply do not have such a property.

Overall, I find that trading in social networks does not efficiently allocate the technology. I first compare adoption in networks with two benchmarks for demand: the allocation where every farmer with positive expected returns adopts and a revealed preference measure from door-to-door sales. With respect to either benchmark, the adoption rate in networks is inferior. In contrast to this reduced adoption, trading in networks does result in a small improvement in targeting of buyers with larger expected gains. However, this improvement in targeting is not large enough to offset the reduction in adoption. The limiting of transactions to close peers is one factor that limits the ability of networks to allocate the technology. To establish this, I show that existing social relationships between buyers and suppliers – defined by caste and surname association – have significant influence on adoption in networks.

Three sources of experimental variation are used in the analysis. First, five farmers were randomly chosen in each of 82 villages to receive a small amount of the new seed variety. After a single year of production, this small amount produces a large amount of output that can potentially be used as seeds by other farmers in the village. The selection of the initial recipients of the technology is therefore akin to selection of "suppliers" because these initial recipients were effectively endowed with more than enough seeds for their own cultivation. The random selection of suppliers allows for causal identification of whether social relationships with suppliers affect adoption in networks.

Following the first year of production, the second source of variation was village-level randomization of the mode of exchange. In half of the villages nothing further was done, effectively forcing adopters to rely on suppliers for taking up the technology. I refer to this system of exchange as the "network" because trading is decentralized, non-anonymous, and thus requires at least some link between buyers and sellers.[1]

The seed was *additionally* made available via door-to-door sales in the remaining half of villages. The purpose of this intervention was to generate a revealed preference measure of demand in an environment with minimal transaction costs and no frictions due to identities of buyers and sellers. Importantly, the door-to-door intervention is not meant to simulate a potential policy, but rather to generate a benchmark of demand. Therefore, the allocation produced in these villages is a benchmark measure that can be compared to both the allocation in networks and the perfect allocation where the technology reaches all farmers with positive expected gains. Since exchange via networks could still occur in villages where sales were offered, the design allows me to address whether networked trade alone meets demand. If so, then the additional adoption resulting from access to door-to-door sales should be

[1] The term "link" is used to refer to links used for the purpose of making one-shot transactions, not necessarily links for more repeated interactions such as mutual insurance.

small.

The third source of variation was randomization of prices at which sales offers were made. Since transaction prices in networks were beyond the control of the experiment, I rely on price randomization to ensure that a comparison between the two modes of exchange can be made while holding prices constant. Thus, the design ensures that price differences can not explain the results.

The experiment produced four main results. First, the overall rate of adoption is 83% lower with networks alone. Only 7% of farmers adopted in network villages, while 40% did in door-to-door villages. Considering that 84% of farmers are expected to benefit from the technology, only about half of the optimal adoption rate is achieved in door-to-door sales. Nonetheless, the magnitude of the difference between adoption in networks and revealed demand suggests that a significant share of farmers that otherwise have positive demand for a product do not adopt when exchange occurs in social networks. An alternative explanation of this effect is that door-to-door sales met demand by simply eliminating scarcity. I show that this explanation is unlikely because the amount of seeds available to suppliers was sufficient to meet the potential demand of *more* than an entire village.

The second result speaks to how social relationships restrict trading in networks. Specifically, farmers relying on networks are more likely to adopt when the suppliers in their village belong to the same sub-caste or share the same surname. In my preferred specification, having the same surname as an additional supplier results in a 106% increase in the probability of adoption. Similarly, being part of the same sub-caste as an additional supplier leads to a 53% increase in adoption probability. Relationships with suppliers are less important for demand in door-to-door sales. An equivalent interpretation of the finding is that introducing an outside buying opportunity increases adoption, but particularly for those that are not connected to suppliers and thus would have otherwise faced barriers to adopting in networks. The result provides micro-level evidence that is consistent with the cross-country result that the diffusion of technology is slower in countries where networks are organized into distinct sub-networks or collectives (Fogli and Veldkamp, 2012). Additionally, the result empirically demonstrates the importance of network structure for trading outcomes – something that is consistent with results from laboratory experiments (Charness, Corominas-Bosch, and Frechette, 2007; Gale and Kariv, 2009).

Third, I show that while there is a moderate improvement in targeting by social networks, it is insufficient to offset the large gap between adoption and demand. I exploit the flood-tolerance property of Swarna-Sub1 to generate estimated gains in revenue using impact estimates from a randomized experiment (Dar et al., 2013). In particular, I calculate a farmer-specific measure of expected gains from the new technology. I find that trading in social networks is moderately effective at concentrating adoption amongst farmers that have above-median gains in expected revenue. In addition, the average return of adopters decreases by approximately 23% when door-to-door sales are offered. Nonetheless, this moderate improvement in targeting by networks only offsets a small amount of the inefficiency due to reduced adoption.

As an extension to this result, I exploit the random variation in sales prices to show that increasing prices does little to improve targeting. Most simply, the average return of adopters is no larger at higher prices. If anything, the results suggest that higher prices are less effective at targeting buyers with high returns. This finding adds to the literature on the allocative efficiency of prices for distributing technologies in developing countries (Ashraf, Berry, and Shapiro, 2010; Cohen and Dupas, 2010; Cohen, Dupas, and Schaner, 2013).

Building on these first three results, my final result quantifies the magnitude of the losses resulting from trading in networks. I define these losses as the percentage of the total gains in expected revenue in the door-to-door villages that are not achieved in network-based exchange. This measure is not a measure of overall social welfare, but is a measure of the losses to buyers of having to rely on networks for adoption. The total expected gain in one-year revenue due to the new technology is almost three times larger in villages where farmers were offered door-to-door sales. More precisely, the loss due to missed trading opportunities in networks represents 63% of the total gains achieved with door-to-door sales. The magnitude of the revenue effect implies substantial losses to farmers due to trading in networks.

These results are based on the short-run allocation of the technology after a single growing season. One caveat of the results is therefore that the allocation observed over a longer time period may differ from that in the short run. In short, the proper interpretation of the findings is that in the short-run, exchange in social networks is unable to meet demand.

This finding that trade in networks does not meet demand adds new empirical evidence helping to distinguish between different models of decentralized trade in networks. If demand is competitive and the costs of making links with suppliers are sufficiently low, then the model in Kranton and Minehart (2001) shows that efficient allocations are a unique equilibrium to a noncooperative game of network formation. In contrast, there may be frictions that limit exchange. As one example, not all efficient transactions will be made if sellers have some bargaining power and links between buyers and sellers are considered as relationship-specific investments (Elliott, 2013). My results suggest that there are indeed frictions that restrict the flow of goods across social groups in networks.

The results also contribute to the literature on the barriers to the adoption of agricultural technologies in developing countries. Important barriers that limit adoption of profitable agricultural technologies include limitations to demand such as self control (Duflo, Kremer, and Robinson, 2011) and risk (Dercon and Christiaensen, 2011). In line with Suri (2011), this paper suggests that constraints on the supply side are also important for limiting technological progress in agriculture.

An important policy implication of the results is that although seemingly desirable as a low-cost method of diffusing a new technology, social networks alone may not efficiently allocate the technology. Given the push to make development interventions sustainable (Kremer and Miguel, 2007), relying on decentralized exchange through social networks seems ideal because of its low cost. Indeed, farmer-to-farmer seed exchange is common throughout the developing world (Sperling and Loevinsohn, 1993; Almekinders, Louwaars, and De Bruijn, 1994). My results suggest that this approach significantly limits the diffusion of new technologies.

The rest of this chapter is organized as follows. In section 3.2, I provide a description of how the experiment was specifically designed to measure the efficiency of exchange in networks. Section 3.3 provides a model of technology adoption that lays the groundwork for the empirical analysis in section 3.4. After establishing the inefficiency of exchange in networks, section 3.5 provides further analysis that points to network structure and the tendency to transact with only close peers as one important explanation of this result. Section 3.5 also presents evidence that is inconsistent with some alternative explanations of the findings. Section 3.6 concludes.

3.2 Experimental Design

In this section I describe the approach to create random variation in the identities of suppliers, the mode of exchange, and transaction prices in door-to-door sales. Motivated by the questions of whether exchange in networks allocates to everybody with demand and whether social relationships with suppliers influence adoption in networks, I discuss how the three sources of variation can be used to answer these questions. Finally, I also discuss the timing of data collection.

The experiment was carried out in 82 villages in three blocks of Bhadrak district of Odisha (see Figure 3.1 for a map of the villages).[2] The villages were selected using satellite imagery of flooding during 2008 and 2011. The villages are located in a low-lying coastal area adjacent to the Bay of Bengal. The median elevation of the district is approximately 10 meters, and rivers flowing from adjacent higher-elevation districts make flooding frequent during the June-October rainy season. Most recently, heavy flooding occurred in 2008, 2009, and 2011.

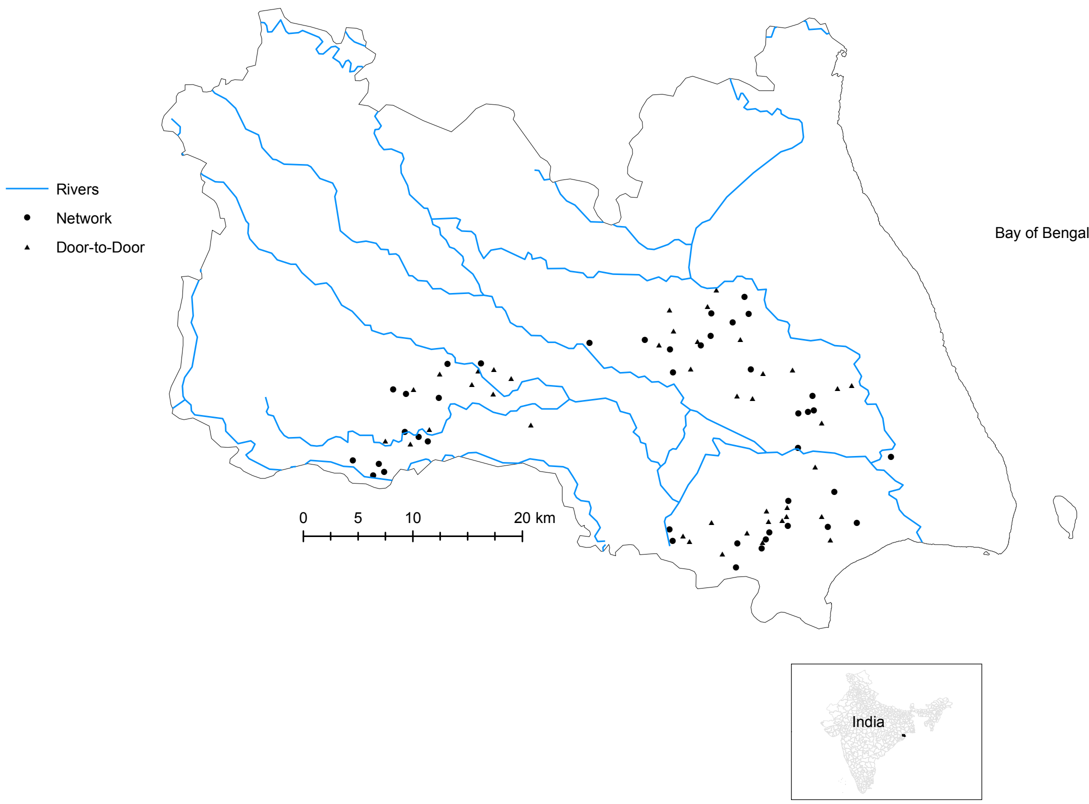
Suppliers were randomly selected at a village meeting carried out in May 2012. Each village was visited and farmers were informed that there would be a meeting to discuss a new flood-tolerant rice variety. The meeting was open to any farmers cultivating rice. Participants were informed that five farmers would be chosen via lottery to receive a five kilogram minikit of Swarna-Sub1.[3] The meetings were attended by anywhere from 15 to 41 farmers, with average attendance being 22 – or approximately 22% of cultivating households in the village.[4] During each meeting, enumerators provided a brief overview of the characteristics of Swarna-Sub1, described its similarity to the known variety Swarna, and pointed to flood tolerance as its only known benefit. After the information was provided, each farmer provided responses to a short baseline social network survey before placing their name in a bucket for

[2]The total number of villages is 84. Two villages were used for piloting of surveys and interventions and are therefore not used in the analysis.

[3]Minikits are a common approach to introducing a new seed variety in India (Bardhan and Mookherjee, 2011). Each minikit contained only five kg of Swarna-Sub1 seeds, which is enough to cultivate approximately 0.1-0.2 hectares. The minikits were identical to those provided in Dar et al. (2013).

[4]The households in the sample are fairly representative of the villages. The average share of the population that is Scheduled Caste is 20% in both the sample and the matched 2001 census of villages. Average household size and male literacy are also similar between the sample and the census.

Figure 3.1: Location of villages in Bhadrak district



the lottery. After all data were collected, the names of the five recipients were drawn and minikits were provided. Planting occurred upon the onset of the southwest monsoon, which occurred around the second week of June.

The selection of five original recipients is akin to random selection of the "suppliers" since their role in the experiment is to multiply the seed and sell/exchange with other farmers after the harvest but prior to the following growing season. Importantly, the identities of suppliers were known to all farmers attending, thus eliminating the possibility that lack of information on identities of suppliers affected the experiment. By randomly selecting suppliers, I can compare adoption outcomes between non-recipients (henceforth "buyers") that are more or less connected to suppliers.

Most suppliers complied with the experiment by planting the seeds contained in the minikit. To verify this, enumerators returned to all villages during harvesting in November/December to collect information about production. 396 of the 410 farmers were surveyed.[5] Of the farmers surveyed, 87% indicated that the minikit had been planted.[6]

The amount of seed provided to suppliers produced enough output to eliminate any concern that demand could not be met with the harvest. The average harvest of Swarna-Sub1 at the village level was approximately 1.8 tons. This amount was sufficient to meet potential demand because most farmers use approximately 5-10 kg of seed during their first year of cultivation and there is an average of 103 farmers per village. As I discuss in further detail in Section 3.5, alternative uses of output were no more profitable to suppliers, indicating that suppliers had no incentives to use the output in other ways.

Prior to randomization of the mode of exchange, a survey was administered to 1,151 randomly selected potential buyers during February-April 2013.[7] There were four purposes of this survey. First, a plot-level record of the duration of past flooding events during the previous five years was collected in order to estimate the expected returns of the new technology. I return to the estimation of expected returns using these data below. Second, farmers were also reminded about the new variety and the potential to obtain it from other farmers in the village. These reminders limit the possibility that farmers chose not to adopt simply because they had forgotten or did not know about the technology. Third, all potential buyers were again informed about the flood tolerance property of Swarna-Sub1 and that it is most effective during flooding of 5-15 days. Fourth, another social network survey was administered, thus allowing for analysis of whether stated network relationships responded to selection of suppliers.

The mode of exchange was randomized at the village level prior to planting for the 2013 season. In half of the villages, no intervention was carried out and thus decentralized trade between farmers was the only channel for adoption. The transactions between farmers in these villages could be sales, exchanges, or outright gifts – the latter likely occurring with some expectation of future reciprocity.[8] The randomization of the mode of exchange was stratified by block – an administrative unit two levels above villages – and the relative importance of suppliers to buyers.[9]

In the remaining half of villages, farmers were additionally given the opportunity to

---

[5]14 of the 410 suppliers could not be reached because either the household had moved from the village or household members were away for work during survey visits.

[6]The most common reason reported for not cultivating the minikit was that the seedbed was damaged by drought or cows. The common method of planting rice in the area is transplanting, which involves preparing a small seedbed and uprooting the small seedlings approximately 3-4 weeks after emergence. The uprooted seedlings are then bundled and planted in the main field. Lack of water is particularly problematic for the seedbed.

[7]In villages with more than 15 potential buyers, a random sample of 15 names was drawn from the list of remaining farmers from the original village meeting. All buyers were selected if there was less than 15 names remaining.

[8]The ability to exchange seeds is an advantage of the networked market if farmers face liquidity constraints at the time right before planting.

[9]Suppliers were defined as being relatively more important when the ratio of average degree of suppliers to the average degree of buyers was larger than the sample median. The degree is simply the number of links of a farmer, where two farmers are defined to be linked if either farmer stated that they would go to the other farmer for seeds, fertilizers, or other inputs.

purchase the technology from NGO staff. The NGO staff went directly to the homes of farmers to make sales offers at pre-determined village-level prices. Except for telling the farmers about availability of the technology, the staff gave farmers no additional details about its benefits. Since farmers knew about the technology from the village meeting and previous surveys, there is little chance that increased awareness could drive the results. Nonetheless, I return to this possible explanation in Section 3.5.

Since five suppliers were selected in all villages, network-based exchange was equally possible in all villages. Therefore, taking door-to-door sales as a method for eliciting demand, the question being addressed by random provision of door-to-door buying opportunities is whether exchange through networks alone leaves significant demand unmet. If so, then a large number of farmers will be "crowded in" when door-to-door sales are available.

Returning to prices, the prices were randomized in order to approximate the prices paid in transactions between farmers. Farmers often exchange seeds directly or sell at prices that are approximately equal to prices of harvested rice. Since the opportunity cost to the seller of such a transaction is the value of output, a sensible benchmark is the output price of rice for consumption. The minimum support price set by the Indian government for the 2012-2013 season was 12.5 Rs per kg (1 USD $\approx$ 58 Rs). Many farmers also sell to private traders at prices ranging from 10-12 Rs. Using these values as a benchmark, prices were randomly set at 3 levels: 10, 12, and 14 Rs per kg. Since most network transactions are one-to-one exchanges of Swarna-Sub1 for a different variety of seeds, these prices are reasonable proxies for prices paid in network transactions. Therefore, I can effectively hold prices constant by estimating the main treatment effects at the average price in farmer-to-farmer transactions.

A final endline survey was carried out in all villages in July 2013 to track adoption and area planted. The survey was administered to all farmers in order to verify transactions from both buyers and suppliers. A total of 1,150 of the previously surveyed buyers and 394 of the previously surveyed suppliers were reached. I use adoption from this survey as the main outcome variable throughout the remainder of the paper.

Summary statistics indicate that the experimental groups are comparable on observable characteristics. Panel A of Table 3.1 shows mean values of baseline observable characteristics for the suppliers and randomly selected buyers. Observable characteristics of suppliers appear similar to those of buyers, suggesting that the randomization in the field was successful at generating a random group of suppliers. Focusing on the social network measures, two farmers are defined to have an information link if either farmer indicated they would go to the other farmer to talk about rice farming. Similarly, two farmers have a sharing link if either farmer indicated that the other farmer is somebody they would hypothetically go to for seeds, fertilizers, or other inputs. Each farmer has on average 5 information links and 4.25 sharing links.

Village-level statistics from the 2001 census are presented in Panel B of Table 3.1. The villages are fairly small, with an average of 165 households, 103 of which are engaged in cultivation. The average elevation of five meters shows that the villages are located in a coastal low-lying area. Importantly for the design, the share of suppliers not cultivating the seeds provided and the aggregate Swarna-Sub1 harvest are balanced across network and

Table 3.1: Summary Statistics

	(1) Buyer	(2) Supplier	(3) p-value: (1)-(2)
Panel A: Farmer Level Statistics (N=1584)			
Rice acres in Kharif 2011	3.88	3.80	0.53
Acres flooded 4 days or less in Kharif 2011	1.25	1.25	0.94
Acres flooded 5 days or more in Kharif 2011	2.63	2.56	0.52
Acres grown with Swarna in Kharif 2011	1.95	1.88	0.34
Farmer is Scheduled Caste (SC)	0.20	0.18	0.46
Age of farmer	48.96	49.07	0.86
Farmer is lead farmer	0.09	0.11	0.29
Information degree	4.89	5.02	0.40
Sharing degree	4.19	4.37	0.21
Information in-degree	2.31	2.44	0.36
Sharing in-degree	1.94	2.16	0.08*
Panel B: Village Level Statistics (N=82)	Network	Door-to-door	
Total households	149.68	180.60	0.26
Total cultivators	89.41	117.33	0.13
Total Ag. laborers	46.80	55.42	0.46
Persons per household	5.84	5.90	0.64
Share Scheduled Caste (SC)	0.21	0.17	0.29
Literacy Rate	0.63	0.65	0.26
Approximate elevation (m)	5.29	4.28	0.19
Share of farmers not cultivating minikit	0.11	0.14	0.54
Estimated village harvest of Swarna-Sub1	1647.71	2066.22	0.20

Asterisks indicate statistical significance at the 1% ***, 5% **, and 10% * levels. Column 1 in Panel A is for buyers. Column 2 in Panel A is for suppliers. Column 1 in Panel B is for network villages, while column 2 in Panel B is for villages where door-to-door sales were made. All farmer level statistics are from the baseline survey in May-June 2012. Information degree is the number of links (undirected) where a link occurs if either farmer lists the other farmer as somebody with which they talk about rice farming. Sharing degree is the number of links (undirected) where a link occurs if either farmer lists the other farmer as somebody with which they would go to if they needed seeds, fertilizers, or other inputs. Information in-degree is the number of *other* farmers in the village naming this farmer as an information contact.

door-to-door villages, suggesting that any differences in adoption can not be attributed to differences in production of suppliers.[10]

3.3 Model of Technology Adoption in Networks

In this section I formulate a model of adoption of a new technology in networks. In contrast to a model where networks function to spread information, I focus on how network relationships create variation in costs of adopting across the population. I then use the model to help show how targeting of buyers with high expected benefits may vary between networks and a setup where these costs are eliminated in door-to-door sales.

Simple Example

Before formulating the adoption choice of buyers, I present a simple example that is meant to convey the ways in which trading in networks may vary from the outcome when door-to-door sales are offered. Figure 3.2 displays the network structure for one of the sample villages, where two farmers are assumed to have a link if they share a common surname, an assumption I provide support for in Section 3.5. The dark nodes (S1-S5) represent the five farmers that were selected as suppliers and the remaining nodes (B1-B15) are potential buyers. Since the harvest of suppliers is enough to meet demand, and there are no alternate uses of the output that are more profitable, the first-best allocation would require each buyer with positive demand to adopt. As an example, if B5 has a high valuation for the technology, then she faces a tradeoff of going to a supplier outside of her network, or not adopting. As the theoretical literature suggests, it is not obvious as to whether these transactions will take place (Kranton and Minehart, 2001; Elliott, 2013).

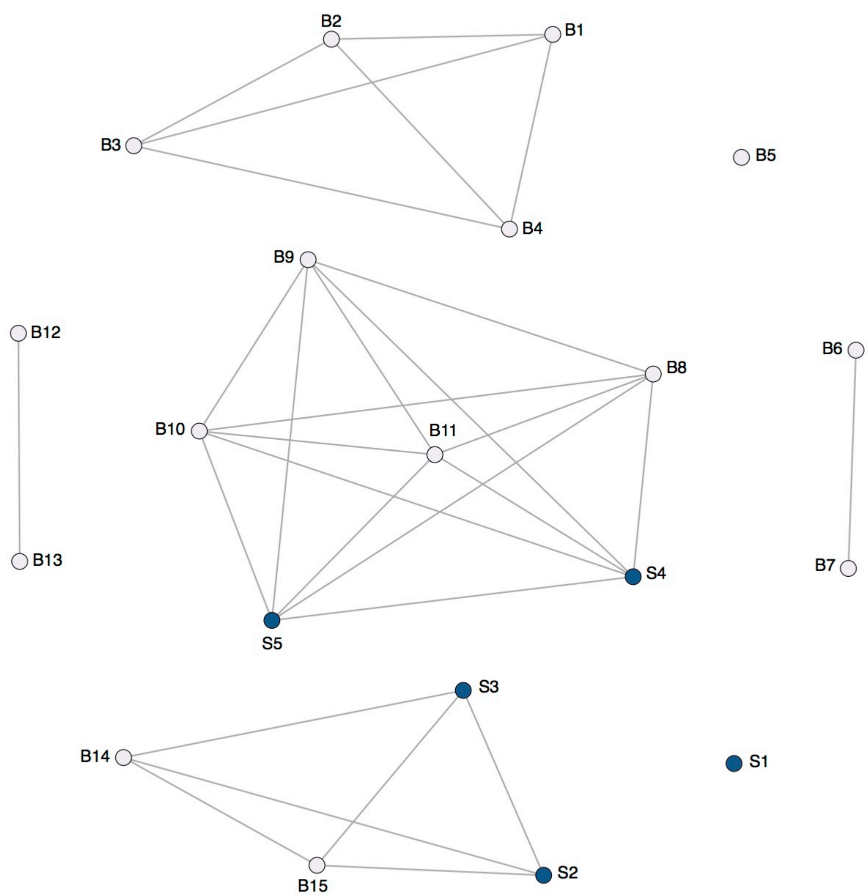
The link pattern is inconsequential when demand is revealed via door-to-door sales because all potential costs of transacting are eliminated. As a result, network structure imposes no barriers to B5 adopting the technology. If network relationships present barriers to adoption, then the amount demanded from door-to-door sales will be significantly larger for those farmers that are unconnected to suppliers. In contrast, if networks work efficiently for exchange, then B5 should adopt regardless of the mode of exchange.

Model Setup

The main benefit to the farmer of adopting the new technology is improved flood tolerance. To formalize this, denote α_i as the probability that farmer i is affected by flooding. The

[10] Another useful test is the test of whether any differences between suppliers and buyers are greater in door-to-door villages as compared to network villages. In results not reported, I regress each characteristic in Panel A of Table 3.1 on village-level treatment, a supplier indicator, and the interaction of these two variables. The F-statistics of these 11 regressions range from 0.29 to 1.19 and thus the three variables do not jointly explain variation in any of the farmer characteristics.

Figure 3.2: A Sample Network



Notes: Figure displays a network diagram for one of the 82 sample villages. Dots (nodes) represent individual farmers and edges (lines) represent connections, where connections are assumed if the farmers share a common surname. The shaded nodes, marked S1-S5 are farmers that were randomly selected as suppliers. The unshaded nodes, marked B1-B15, were randomly selected as buyers.

agronomic return of the technology when flooding occurs is $r_i > 0$. Conversely, the return under non-flood conditions is zero – an assumption consistent with the experimental results in Dar et al. (2013). Therefore, the expected return of the technology is $R_i = \alpha_i r_i$.

In addition to the returns due to flood exposure, there is an idiosyncratic term, u_i, which measures benefits that are observed to the farmer, but not to the econometrician. For instance, some farmers may have stronger preferences for trying new technologies. I assume that u_i is mean zero and independent of both R and c. Since R_i can be approximated with data on past exposure to flooding and u_i is unobservable, the door-to-door sales treatment serves the purpose of generating a measure of overall demand that is based on both terms.

The difference in prices between the old and new technologies is v. I consider prices to be fixed. While there is a large literature on bargaining between buyers and sellers in networks (Corominas-Bosch, 2004; Manea, 2011; Abreu and Manea, 2011), I focus on network structure and costs of exchange as potential barriers to adoption. The lack of significant variation in prices and the prevalence of exchanging seeds at a rate of one-for-one suggests that bargaining is not an important consideration in this context.

The costs of exchange in networks can be decomposed into two terms. The first term, $\underline{c}$, is the inconvenience of having to leave the house to obtain seeds. While likely important in many contexts, $\underline{c}$ is likely to be less important in this sample due to the close proximity of houses. Half of the households in the sample have a supplier that lives within 42 meters of their household. Also, over 90% of buyers are located within 300 meters of a supplier.

The second term, c_i, denotes the costs to the buyer of making a trading link with a supplier. The value of c_i varies across the population because of varying degrees of connectedness to suppliers. As an example, a low caste farmer may find it very costly to adopt from a higher caste supplier. However, it need not be the case that $c_i > 0$. For instance, a farmer may *benefit* from trading in networks if peers extend credit or allow for other types of flexible payments.[11]

In contrast to a conventional market, door-to-door sales eliminate both $\underline{c}$ and c_i by making transactions anonymous and bringing seeds directly to farmers. A standard market with some transportation costs would eliminate c_i, but not $\underline{c}$. I use various measures of connectedness to suppliers to show empirically that c_i is quantitatively important.

I assume that c, R, and u are distributed multivariate normally where the means of c and R are μ_c and μ_R. The parameter ρ denotes the correlation between R and c. The idiosyncratic term u has a mean of zero and is uncorrelated with both R and c.

Combining all benefits and costs, the probability of adopting the new technology in networks is $P\{R + u > v + c + \underline{c}\}$. Introducing door-to-door sales causes the adoption probability to be $P\{R + u > v\}$. In addition to differences in adoption, there may be a targeting effect where the quality of adopters varies between the two modes of exchange.

[11] See Kranton (1996) and Aoki and Hayami (2001) for discussion of some of the benefits of reciprocal exchange through networks.

Targeting

The expected return of adopters is a natural measure of targeting effectiveness. Conditional on the overall rate of adoption, the average return of adopters is a direct measure of how efficiently the technology is allocated. As shown in the appendix, the expected return of adopters when trading occurs in networks is

$$E(R|R+u-c>v+\underline{c}) = \mu_R + \frac{\sigma_R(\sigma_R-\rho\sigma_c)}{\sqrt{\sigma_R^2+\sigma_u^2+\sigma_c^2-2\rho\sigma_R\sigma_c}} * M\left(\frac{v+\underline{c}-\mu_R+\mu_c}{\sqrt{\sigma_R^2+\sigma_u^2+\sigma_c^2-2\rho\sigma_R\sigma_c}}\right), \tag{3.1}$$

where $M(z) = \frac{\phi(z)}{1-\Phi(z)}$ is the inverse Mill's ratio. The expected return of adopters is above the average in the population if $-1 \leq \rho < \frac{\sigma_R}{\sigma_c}$. Intuitively, if returns and costs of adopting are negatively correlated, then the farmers facing the fewest barriers to adopting in networks are those with the highest returns. Therefore, on average, the adopters have higher returns than the overall population. On the other hand, if the correlation between costs and returns is sufficiently large, then targeting in networks is less effective because the farmers with high costs are those with high returns.

The expected return of adopters in door-to-door sales is

$$E(R|R+u>v) = \mu_R + \frac{\sigma_R^2}{\sqrt{\sigma_R^2+\sigma_u^2}} * M\left(\frac{v-\mu_R}{\sqrt{\sigma_R^2+\sigma_u^2}}\right). \tag{3.2}$$

Comparing equations (3.1) and (3.2), the average return of adopters will be larger in door-to-door sales if $\rho > \frac{\sigma_R}{\sigma_c}$. In this case, door-to-door sales crowd in farmers with high returns that did not adopt in networks. Conversely, if ρ is held constant, and μ_c is large relative to σ_c, then adoption in networks sends a stronger signal that returns are large because the farmer is willing to make the costly investment of adopting from other farmers.[12] Expected returns of adopters in networks will be larger in this case.

Overall, the model predicts that the difference in the average return of adopters between the two modes of exchange will depend upon the magnitudes of μ_c, σ_c^2, and ρ. However, the experimental design causes c and R to be uncorrelated when conditioning on network size. This results because random selection of suppliers generates random variation in connectedness to suppliers. Applying this to the model, the targeting effectiveness of exchange in networks will depend only on the distribution of costs. If barriers to exchange in networks are irrelevant, then costs will be small and the average returns of adopters in networks will be similar to that in door-to-door sales. This prediction is considered in detail in the empirical analysis.

[12] To see this, note that $M(z)$ increases monotonically with z. Therefore, as μ_c increases and σ_c decreases, both M and $\frac{\sigma_R(\sigma_R-\rho\sigma_c)}{\sqrt{\sigma_R^2+\sigma_u^2+\sigma_c^2-2\rho\sigma_R\sigma_c}}$ increase.

3.4 Results

I first report results showing that exchange in social networks results in lower adoption, crowding out of farmers with fewer connections to suppliers, and a small improvement in targeting. These results build up to an overall measure of efficiency losses. I then consider whether adoption effects vary across the population. Finally, I take advantage of the random variation in prices to show that increasing prices is not effective at screening the pool of adopters.

Adoption, Peer Effects, and Targeting

I first show that exchange via social networks alone results in significantly lower adoption when compared to villages where door-to-door sales were used to reveal demand. In order to estimate the magnitude of this effect while holding prices constant, I rely on random price variation to estimate the effect at the average price observed in network transactions between farmers. Formally, the regression specification is

$$adoption_{ij} = \beta_0 + \beta_1 door\ to\ door_j + \beta_2 door\ to\ door_j * (price_j - 12.4) + \varepsilon_{ij}, \qquad (3.3)$$

where $adoption_{ij}$ is an indicator for adoption by farmer i in village j, $door\ to\ door_j$ is an indicator for door-to-door villages, and $price_j$ is the random offer price in door-to-door villages.[13] Since the average price of transactions between farmers is 12.4 Rs per kg[14], the coefficient β_1 measures the gap between network adoption and revealed demand at a price equivalent to an average network transaction.

The estimates in column 1 of Table 3.2 show that when holding price constant, the demand revealed by door-to-door sales is higher by 33 percentage points. The rate of adoption of 40% in door-to-door sales is larger than the adoption in networks by over five times. Focusing on the ratio of the two estimated coefficients in column 1, the price charged in door-to-door sales would need to approximately double to result in the same adoption rate observed in networks alone. The coefficient changes little when including control variables (column 2). Further, as shown in column 3, adoption in networks fell far below demand at all three price levels, even at the highest price, which is larger than the prices of almost all farmer-to-farmer transactions.

One potential explanation for the low adoption in networks is that exchange tends to be limited to farmers from the same social groups, effectively crowding out farmers without links to suppliers. I rely on the random selection of suppliers to test whether relationships

[13] I focus on a binary adoption rate throughout the paper because the amount used is only relevant for a single year. After one year, the harvest produced from only 1-2 kg of seed is enough to cultivate the average farmer's entire landholdings. In door-to-door villages, the adoption indicator is set to 1 if *either* the farmer purchased from an NGO representative, or adopted from a peer.

[14] The dominant transaction type is direct exchanges of Swarna-Sub1 for a different variety of rice. The price for these exchanges is valued at the output price of rice. I use the most conservative estimate, which is the government-supported price of 12.5 Rs per kg.

Table 3.2: Estimated difference between adoption in networks and demand revealed in door-to-door sales

	(1)	(2)	(3)
Door-to-door treatment	0.327***	0.328***	
	(0.042)	(0.041)	
Door-to-door treatment*(Price-12.4)	-0.026	-0.026	
	(0.024)	(0.024)	
Door-to-door and Price=10			0.385***
			(0.078)
Door-to-door and Price=12			0.351***
			(0.067)
Door-to-door and Price=14			0.280***
			(0.059)
Farmer is SC		-0.058	-0.057
		(0.041)	(0.040)
Farmer has BPL card		-0.056*	-0.056*
		(0.031)	(0.030)
Land cultivated in 2012		0.005	0.005
		(0.007)	(0.007)
Ag. cooperative member		-0.019	-0.019
		(0.023)	(0.023)
Swarna user in 2012		0.087***	0.086***
		(0.033)	(0.032)
Strata Fixed Effects	Yes	Yes	Yes
Mean of Dep Variable: Network	0.07	0.07	0.07
Number of Observations	1150	1134	1134
R squared	0.190	0.208	0.209

Dependent variable is 1 if farmer adopted Swarna-Sub1 for 2013 wet season. *Door-to-door treatment* is 1 for villages where farmers had the opportunity to obtain the technology from a door-to-door salesperson. Land cultivated in 2012 is measured in acres. Standard errors are clustered at the village level. Asterisks indicate statistical significance at the 1% ***, 5% **, and 10% * levels.

between buyers and suppliers are more important in networks. The estimating equation is

$$\begin{aligned} adoption_{ij} =& \beta_0 + \beta_1 door\ to\ door_j + \beta_2 suppliers_{ij} + \beta_3 degree_{ij} + \beta_4 suppliers_{ij} * door\ to\ door_j \\ &+ \beta_5 degree_{ij} * door\ to\ door_j + \varepsilon_{ij}, \end{aligned} \tag{3.4}$$

where $suppliers_{ij}$ is the number of peers of farmer i that were selected as suppliers and $degree_{ij}$ is the total number of peers of farmer i. Peers are defined using either the baseline social network survey, common surnames, or belonging to the same sub-caste.[15] Importantly for identification of β_2 and β_4, the random introduction of the technology guarantees that the number of suppliers that are connected to a given farmer is as good as randomly assigned when conditioning on the total number of connections, thus avoiding the classic reflection problem discussed in Manski (1993).

The results in Table 3.3 show that stated relationships with suppliers from the baseline social network survey have little impact on adoption in both types of villages. The effect of being linked to an additional supplier in social networks is small and not statistically significant across all specifications in columns 1-3. As seen by the estimate of β_4, adding door-to-door sales does little to change this effect.

In contrast, sharing surnames with suppliers is significantly more important for obtaining the technology when trading occurs in networks. In column 4, sharing the same surname with an additional supplier results in a 3.5 percentage point, or 50%, increase in the probability of adoption in networks. Adding door-to-door sales causes this effect to decrease significantly by 7.5 percentage points. The negative effect of 4 percentage points in door-to-door sales represents an approximate 10% decrease in adoption, but the effect is not statistically significant (p=0.24). The results do not change substantially when adding household controls (column 5). Turning to column 6, the effects are somewhat larger when village fixed effects are added.[16] Having the same surname as a single additional supplier results in a 106% increase in the likelihood of adoption in networks. Again, the effect in door-to-door sales is slightly negative, but not statistically significant. Holding network size constant, a farmer would need to share the same surname as an additional 4.5 suppliers in order to have the same likelihood of adopting as when door-to-door sales are available. A natural explanation for the difference between surname association and stated network links is that farmers have some flexibility to adopt from others that are not their closest peers, but that establishing a trading link with another farmer from a different social group is too costly.

Belonging to the same subcaste as suppliers is also a significant determinant of adoption in networks. In column 7, having one additional supplier from the same subcaste leads to

[15] There is substantial variation in surnames within villages. The average number of unique surnames per village is 5.6. Therefore, each farmer in the sample shares a surname with approximately 3.3 other farmers in the sample.

[16] This likely occurs because the villages with little variation in adoption and where most farmers share the same surname receive less weight in the identification. In Table B.1 I show that the estimated peer effects are much larger in the sample of villages where there was at least one adopter (columns 1 and 2). This is mostly due to very low adoption in one of the three blocks (columns 3 and 4). The results are also more similar to fixed effects results when discarding the 5% of observations where over 15 of the farmers in the village have the same surname (not shown).

Table 3.3: Effects of social relationships with suppliers on adoption

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Door-to-door Treatment	0.346***	0.341***		0.332***	0.327***		0.367***	0.362***	
	(0.068)	(0.067)		(0.056)	(0.055)		(0.065)	(0.062)	
Door-to-door Treatment * Baseline links with suppliers	0.002	-0.000	-0.026						
	(0.033)	(0.033)	(0.031)						
Door-to-door Treatment * Baseline degree	-0.003	-0.001	-0.002						
	(0.016)	(0.016)	(0.013)						
Baseline links with suppliers	-0.007	-0.006	-0.004						
	(0.011)	(0.012)	(0.013)						
Baseline degree	0.005	0.003	0.003						
	(0.008)	(0.008)	(0.007)						
Door-to-door Treatment * Number suppliers w/ same surname				-0.075*	-0.072*	-0.109**			
				(0.043)	(0.042)	(0.045)			
Door-to-door Treatment * Total number w/ same surname				0.021	0.020	0.035**			
				(0.014)	(0.014)	(0.014)			
Number suppliers w/ same surname				0.035	0.027	0.074**			
				(0.026)	(0.027)	(0.032)			
Total number w/ same surname				-0.008	-0.004	-0.023**			
				(0.008)	(0.009)	(0.009)			
Door-to-door Treatment * Number suppliers same sub-caste							-0.056*	-0.058*	-0.053
							(0.030)	(0.030)	(0.036)
Door-to-door Treatment * Total number same sub-caste							0.011	0.011	0.014
							(0.009)	(0.010)	(0.010)
Number suppliers same sub-caste							0.040*	0.024	0.037*
							(0.021)	(0.023)	(0.021)
Total number same sub-caste							-0.010	-0.007	-0.011
							(0.007)	(0.008)	(0.007)
Strata Fixed Effects	Yes	Yes	No	Yes	Yes	No	Yes	Yes	No
Village Fixed Effects	No	No	Yes	No	No	Yes	No	No	Yes
Household controls	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes
Mean of Dep Variable: Network	0.07	0.07	0.07	0.07	0.07	0.07	0.07	0.07	0.07
Number of Observations	1148	1132	1132	1135	1134	1134	1135	1134	1134
R squared	0.185	0.203	0.413	0.191	0.209	0.419	0.192	0.210	0.413

Dependent variable is 1 if farmer adopted Swarna-Sub1 for 2013 wet season. Standard errors are clustered at the village level. Asterisks indicate statistical significance at the 1% ***, 5% **, and 10% * levels.

a 4 percentage point increase in the probability of adoption, representing a 57% effect. The estimated coefficient on the interaction between the door-to-door indicator and the number of suppliers belonging to the same sub-caste is negative and of similar order of magnitude as the effect in networks. Thus, the effect of belonging to the same subcaste as suppliers becomes effectively zero when door-to-door sales are made. As shown in columns 8 and 9, the results are similar when adding control variables and village fixed effects.

The estimated effects of relationships with suppliers are robust to two natural alternative estimation approaches. First, accounting for the dichotomous nature of the dependent variable by using a probit specification has little impact on the estimates (columns 5 and 6 in Table B.1). Second, an alternative way of measuring relationships with suppliers is to use the share of connected farmers that were selected as suppliers. As shown in Table B.2, using this approach actually improves precision of the estimates.

Compared to the existing literature on peer effects, this result highlights a different mechanism through which peers influence behavior. Namely, when products can be directly traded through networks, one may gain access to a new product via their peers. The literature on peer effects consistently points to peers as a source of learning about new technologies or products (Foster and Rosenzweig, 1995; Munshi, 2004; Bandiera and Rasul, 2006; Kremer and Miguel, 2007; Conley and Udry, 2010; Oster and Thornton, 2012; Cai, de Janvry, and Sadoulet, 2012; Bursztyn et al., 2012). In contrast to this learning channel where peers help to overcome information barriers, the presence of peer effects in trading networks creates *inefficiencies* by limiting trading opportunities.

The results up to this point suggest that adoption in social networks falls short of revealed demand and that this adoption gap is larger for farmers with fewer connections to suppliers. The immediate next question to ask is whether targeting is any more or less effective in networks.

As a first step in answering this, I use data on flooding during the past five years to generate a measure of expected returns for each farmer in the sample,

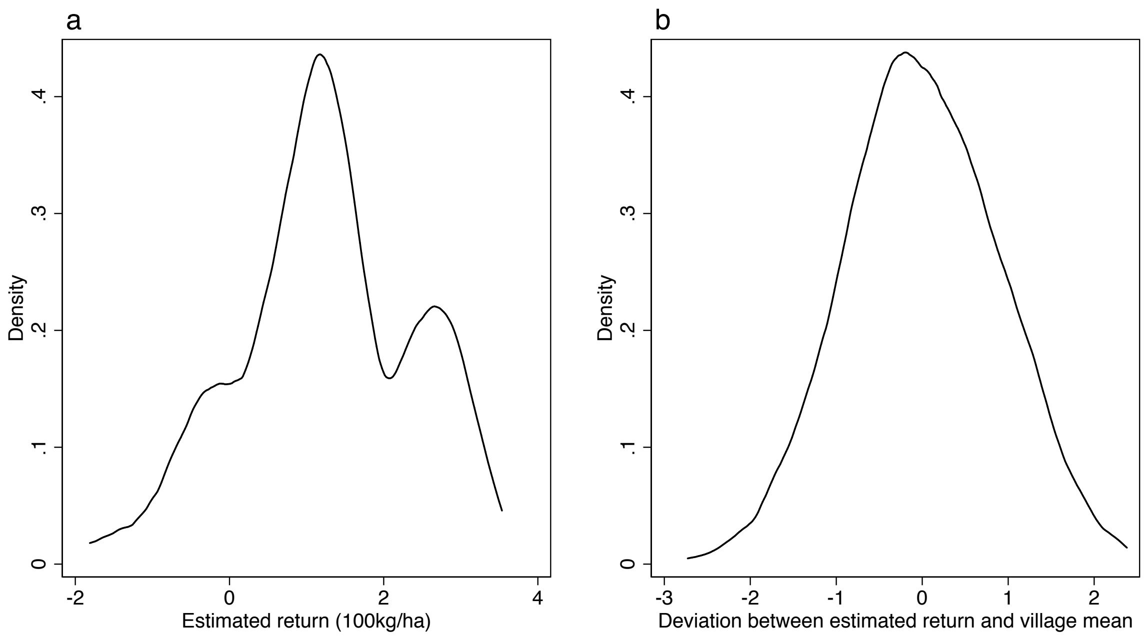
$$return_{ij} = \frac{\frac{1}{5} * \sum_{p=1}^{P_{ij}} \sum_{t=2008}^{2012} R(d_{ijpt}) * area_{ijp}}{\sum_{p=1}^{P_{ij}} area_{ijp}}. \tag{3.5}$$

The term d_{ijpt} represents the duration of flooding for farmer i in village j on plot p during year t, P_{ij} is the total number of plots cultivated, and the function $R(\cdot)$ is the expected agronomic return of Swarna-Sub1, relative to Swarna. The units of measurement of R are kilograms per hectare cultivated. I use estimates of R that were generated using data from a randomized experiment carried out in nearby villages during 2011. Specifically, I use nonparametric estimates of the treatment effect of Swarna-Sub1 as a function of flood duration.[17] The density of estimated returns for the sample of buyers is shown in the left panel of Figure

[17] See the middle panel of Figure 1 in Dar et al. (2013) for the estimates.

3.3. The right panel shows the density of deviations between estimated returns and village means. Variation in topography, and hence flood exposure, generates substantial variation in estimated returns both across and within villages.[18]

Figure 3.3: Distribution of expected returns of Swarna-Sub1



Notes: Figure shows densities of raw estimated returns (Panel A) and deviations between estimated returns and village averages (Panel B). Plot-level recall on flood duration and impact estimates in Dar et al. (2013) were used to calculate expected returns for each farmer in the sample. The only source of variation in expected returns using this methodology is exposure of the farmers' land to flooding.

Following Galasso and Ravallion (2005), I first measure overall targeting performance using a measure that is equivalent to the correlation in a 2 x 2 contingency table. In particular, the measure ϕ is

$$\phi = (a_b - a_{nb}) * \sqrt{\frac{s(1-s)}{a(1-a)}}, \tag{3.6}$$

where $a_b - a_{nb}$ is the difference in adoption rates between farmers with positive net benefits and those with negative or zero benefits, a is the overall adoption rate in the sample, and s is

[18] One caveat is that this approach measures *agronomic* returns rather than *economic* returns. Chapter 2 shows that access to Swarna-Sub1 causes farmers to change several production practices, leading to increases in yield even during years when flooding does not occur. Increases in investment are generally larger for farmers that have more farmers in their peer group also cultivating the variety. Since networks favor adoption by peers, one advantage of farmer-to-farmer exchange is that it could facilitate these behavioral changes.

the share of farmers with positive net benefits. This measure offers two advantages. First, it is neutral to scale, thus allowing an easier comparison between targeting effectiveness in the two treatments where overall adoption varies widely. Second, it is related directly to the objective of a planner that can only observe returns due to flooding. That is, ϕ approaches one if the technology is perfectly allocated to all farmers with positive net benefits. Conversely, ϕ approaches -1 as adoption becomes more concentrated amongst farmers that do not benefit from the technology.

The results in Table 3.4 suggest some slight differences in targeting between the two experimental groups. Neither networks or door-to-door sales are particularly successful at concentrating adoption on farmers with positive expected returns. While the rates of adoption of farmers with positive expected gains are higher by 2.9 and 7.2 percentage points in networks and door-to-door sales, respectively, neither ϕ coefficient is statistically significant. However, social networks are somewhat effective at concentrating adoption amongst farmers with above-median returns. Focusing on the third row of the table, the adoption rate in networks is higher by 5 percentage points – i.e. an increase from 4% to 9% – for farmers with above-median expected returns. This targeting difference is also statistically significant. In contrast, the ϕ coefficient for above-median returns in door-to-door villages is less than half of the size and is not statistically significant.

Table 3.4: Targeting effectiveness of social networks and door-to-door sales

	Networks			Door-to-door sales		
	Targeting Differential	ϕ	p-value	Targeting Differential	ϕ	p-value
Returns greater than:						
0	0.029	0.043	0.302	0.072	0.052	0.212
25^{th} percentile	0.043	0.077	0.067	0.035	0.030	0.476
50^{th} percentile	0.050	0.100	0.017	0.042	0.042	0.312
Flooding 7-14 days during:						
2011	0.011	0.022	0.601	0.114	0.105	0.012
2009	0.015	0.030	0.485	0.009	0.009	0.834
2008	0.050	0.087	0.040	0.044	0.033	0.427

Notes: Targeting differential is the difference in adoption rates between farmers with estimated returns above and below the given threshold. For example, farmers with positive estimated returns are 2.9 percentage points more likely to adopt in social networks when compared to farmers with zero or negative returns. ϕ represents the phi coefficient from the relevant contingency table. Farmers that were flooded for 7-14 days were identified using area-weighted average flood duration collected during survey visits. P-values are calculated using the χ^2 statistic from the relevant 2x2 contingency table.

While door-to-door sales tended to induce adoption by farmers experiencing severe flooding more recently, exchange in social networks induced adoption more by those experiencing flooding in the more distant past. I focus on areas flooded for 7-14 days because this is the range where Swarna-Sub1 has a statistically significant advantage in yield over Swarna (Dar et al., 2013). As shown in the fourth row of Table 3.4, farmers with land flooded for 7-14 days during the 2011 floods were 11.4 percentage points more likely to adopt from the door-to-door salespersons. This targeting difference is statistically significant. In contrast, farmers experiencing flooding from 7-14 days during 2008 were over twice as likely to adopt in networks. Taken together, these results suggest that if anything, targeting is slightly more effective in networks.

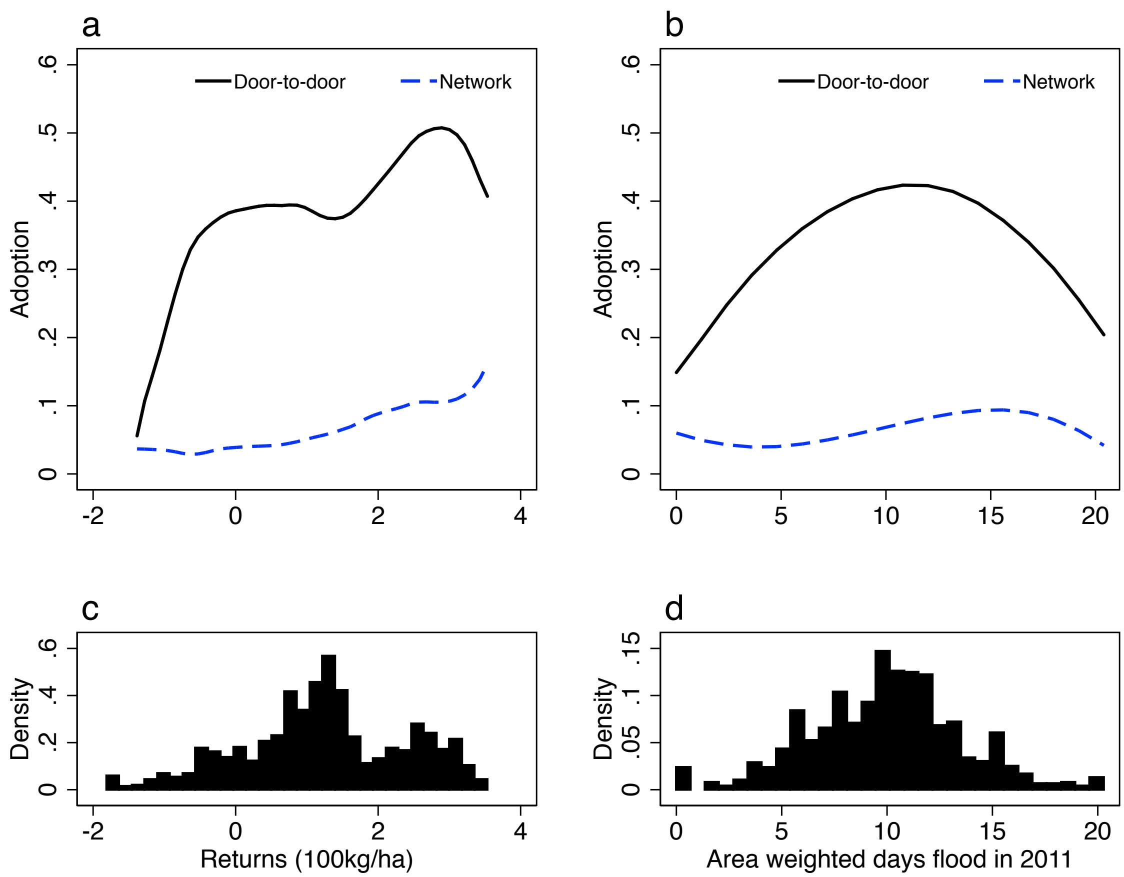
As an additional measure of targeting effectiveness across the entire support of expected returns, Figure 3.4 shows nonparametric fan regressions of adoption on expected returns. Adoption in both treatment arms is positively correlated with expected returns. However, other than for the lowest values of estimated returns, the difference in adoption between networks and door-to-door sales is fairly constant. Following the binary targeting results from Table 3.4, Panel B uses the area weighted average flood duration on the farmer's land during the most recent flood in 2011. Adoption in the door-to-door villages shows a quadratic relationship with flood duration, where the maximum adoption occurs around 12 days. This contrasts with networks where adoption is not strongly correlated with 2011 flood intensity. The pattern is quite remarkable given that impact estimates show that agronomic returns during flooding are maximized at approximately 13 days.

The positive correlation between adoption and estimated returns and the quadratic relationship between adoption and flood intensity in 2011 rule out a story where misunderstanding the benefits of the technology drives the results. If farmers did not understand the benefits of the technology, then there would be no reason to expect adoption to be highest in areas exposed to heavy flooding. Farmers appear to have used a combination of available information and their past experiences with flooding, particularly during 2011, to base adoption decisions.

Regression results in Table 3.5 are consistent with the graphical results. The correlation between adoption and expected returns in networks alone is positive, but not quite statistically significant (column 1).[19] An increase from the 25th to 75th percentile in the expected returns distribution leads to an increase in the probability of adoption by 3 percentage points, or 43%. Adding door-to-door sales results in an increase in the correlation between returns and adoption, but the interaction term is not statistically significant. However, the

[19] Two sets of standard errors are used to make statistical inference. First, OLS standard errors are reported in parentheses. Second, bootstrapped standard errors that correct for expected returns being a regressor generated from a separate sample are reported in brackets. The issue is similar to two sample instrumental variables, where authors have calculated standard errors using either the covariance matrix in Murphy and Topel (1985), the delta method, or by bootstrapping (Inoue and Solon, 2010). Following Björklund and Jäntti (1997), I use the bootstrapping method. I draw 200 samples (clustered at the village level) from both the main estimation sample and the sample in Dar et al. (2013). For each sample the nonparametric fan regression relating returns of Swarna-Sub1 to the duration of flooding is re-estimated and expected returns in the sample drawn from the estimation sample are re-calculated using this new mapping between flood duration and estimated returns. I then estimate the regression with these new values of estimated returns. Bootstrapped standard errors for each parameter are calculated as the standard deviations of the 200 estimates.

Figure 3.4: Relationship between estimated returns and adoption, by treatment



Notes: **(a)** Nonparametric fan regression of adoption on estimated returns. **(b)** Nonparametric fan regression of adoption on area weighted duration of flooding during 2011 floods. **(c)** Density of estimated returns. **(d)** Density of area weighted flood duration in 2011.

overall effect in door-to-door sales is statistically significant. Moving from the 25th to 75th percentiles of the expected returns distribution in door-to-door sales leads to a 7.5 percentage point (19%) increase in adoption. The results in column 2 verify that the quadratic relationship between adoption and 2011 flood severity is highly statistically significant in door-to-door sales, but not in networks alone.

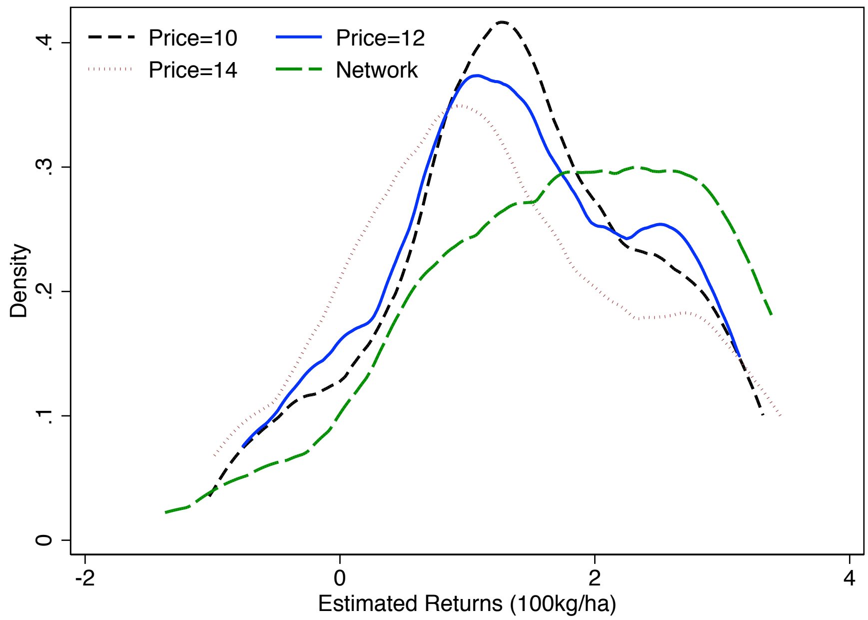
Table 3.5: Estimated correlation between expected returns and adoption

	(1)	(2)
Door-to-door treatment	0.300***	0.130
	(0.049)	(0.126)
	[0.065]	
Door-to-door treatment*Expected Returns	0.031	
	(0.025)	
	[0.024]	
Expected Returns	0.019	
	(0.013)	
	[0.014]	
2011 Area weighted days flood		0.007
		(0.011)
2011 Area weighted days flood^2		-0.000
		(0.000)
Door-to-door treatment*2011 Area weighted days flood		0.044**
		(0.018)
Door-to-door treatment*2011 Area weighted days flood^2		-0.002***
		(0.001)
Strata Fixed Effects	Yes	Yes
Household controls	Yes	Yes
Mean of Dep Variable: Network	0.07	0.07
Number of Observations	1134	1126
R squared	0.212	0.213

Dependent variable is 1 if farmer adopted Swarna-Sub1 for 2013 wet season. *Door-to-door treatment* is 1 for villages where farmers had the opportunity to obtain the technology from a door-to-door salesperson. Conventional standard errors that are clustered at the village level are reported in parentheses. Bootstrapped standard errors that correct for *Expected Returns* being a generated regressor are in brackets. Asterisks (pertaining to conventional standard errors) indicate statistical significance at the 1% ***, 5% **, and 10% * levels.

I consider the average return of adopters as the most direct measure of targeting effectiveness that maps directly to the calculation of efficiency losses. Figure 3.5 displays the densities of estimated returns for adopters across the different treatment groups. Visually, the distribution of estimated returns in network villages shifts to the right when compared to villages where door-to-door sales were made.

Figure 3.5: Densities of estimated returns of adopters, by treatment



Notes: Figure displays kernel densities of estimated returns, by treatment group. Densities are estimated only for adopters.

OLS regression estimates also suggest a moderate improvement in targeting by exchange in networks. The regression results in column 1 of Table 3.6 show that the average return of adopters in door-to-door sales is smaller by 40 kg per hectare, an approximate 23% decrease.[20] The effect is reasonably large, but not quite statistically significant (p=0.11). The average return across the entire sample of 124 kg per hectare can be taken as the return of adopters if the technology had been provided free of cost. Therefore, the average return of adopters in networks (the constant term) represents a 40% improvement over free distribution. This difference is statistically significant (p=0.029).

Focusing on column 2, very similar results are obtained when using a self-reported measure of flood risk for the plot where the new variety would be planted. Farmers were asked to assess on a scale from 1-10 how prone their Swarna-Sub1 plot is to flooding. The average value amongst adopters in networks was 5.25. The predicted decrease with door-to-door sales

[20]Strata fixed effects are dropped in this regression in order to avoid absorbing selection effects.

Table 3.6: Effect of exchange environment on the average return and self-reported flood risk of adopters

	All adopters		Adopters from peers or door-to-door	
	(1) Return	(2) Flood severity (1-10)	(3) Return	(4) Flood severity (1-10)
Door-to-door treatment	-0.402	-0.846	-0.542***	-0.979*
	(0.248)	(0.527)	(0.165)	(0.507)
Door-to-door treatment*(Price-12.4)	-0.037	-0.069	-0.037	-0.069
	(0.061)	(0.109)	(0.061)	(0.109)
Constant	1.742***	5.250***	1.882***	5.382***
	(0.219)	(0.464)	(0.117)	(0.442)
Mean of Dep Variable: Network	1.742	5.250	1.882	5.382
Number of Observations	266	267	264	265
R squared	0.018	0.029	0.031	0.037

Dependent variable in columns 1 and 3 is the expected return of Swarna-Sub1, measured in quintiles (1 quintile = 100kg) per hectare. Dependent variable in columns 2 and 4 is a subjective measure of the flood severity of the plot where Swarna-Sub1 was being planted. This variable ranges from 1-10 and was collected during the final follow-up survey with all potential buyers. *Door-to-door treatment* is 1 for villages where farmers had the opportunity to obtain the technology from a door-to-door salesperson. Standard errors are clustered at the village level. Asterisks indicate statistical significance at the 1% ***, 5% **, and 10% * levels.

is 0.85, or 16.2%. The estimated effect with this separate measure is qualitatively similar, but also not statistically significant (p=0.11).

Columns 3 and 4 show that the results become more precisely estimated when dropping the two farmers that were provided Swarna-Sub1 free of cost from a local disaster management office. These results are largely consistent with the targeting differentials in Table 3.4. Namely, the estimated returns are slightly larger for adopters in networks because there is a larger mass of adopters with expected returns that exceed the median.

Taken together, the results on targeting suggest that improved targeting from exchange in networks will offset a small portion of the inefficiency due to the adoption gap. I next turn to a measure that combines these effects to estimate the overall losses in expected revenue.

Efficiency Loss

As a first step in quantifying the magnitude of the losses to farmers due to trading in networks, I define the gain in expected gross revenue for farmer i as $gain_i = adoption_i * return_i * hectares_i$, where $return_i$ is converted to monetary units by multiplying by the government supported output price of 12.5 Rs per kg. The total gain in expected revenue is then calculated by summing $gain_i$ across farmers. Following the results on peer effects, farmers in network villages are further split into two groups: farmers with one or zero suppliers sharing their surname, and farmers having the same surname as two or more suppliers.[21] I also present potential gains for a scenario where every farmer with positive expected gains adopts.[22] The aggregate gains are then re-weighted to ensure that the total number of observations is held constant across the groups.

Panel A of Figure 3.6 shows that the smallest gains from the technology were amongst the relatively less connected farmers that relied on trading in networks. The total gain in revenue in this group was 16,800 Rs. The gains were 32,800 Rs – or nearly twice as large – amongst the better connected farmers. The aggregate gain in revenue with door-to-door sales is 61,700 Rs. Therefore, approximately 35.6% of the revenue gap between less connected farmers in networks and farmers receiving door-to-door visits can be explained by limited connectivity. Averaging across all network villages, the total gain across all farmers was 23,000 Rs. Thus, the aggregate losses in short-term revenue due to trading in networks represent approximately 63% of the aggregate expected returns generated by door-to-door sales.

While making transactions costless by adding door-to-door sales increases expected revenue, there is still a large gap between door-to-door sales and the allocation where every farmer with positive returns adopts. This is driven by the fact that only 40% of farmers adopt in door-to-door villages even though around 84% of farmers are expected to gain from Swarna-Sub1. Not surprisingly, regardless of the exchange environment, some farmers are likely to wait until additional information about the technology comes available before making adoption decisions.

While there are clear losses from trading in networks, the absolute magnitude of the losses during the first year is small. In particular, the overall loss of 38,700 Rs represents approximately 1.14 USD per farmer. This results for two reasons. First, farmers only cultivate a small amount of the new variety during the first year. Second, the agronomic gains do not account for the changes in farm investment that are induced by the reduction in risk.

Focusing on investments, Chapter 2 shows that farmers with Swarna-Sub1 increase investment in inputs and cultivate more land during the second year of using the technology. I use the estimated parameters in Chapter 2 to estimate gains from the technology over a two year period. The expected gains in revenue during the first year are set equal to the agronomic gains, i.e. those in Panel A of the figure. During the second year, the yield and area cultivated are assumed to increase according to the parameters estimated in Tables 2.8 and 2.3, respectively.[23] The important assumption underlying this calculation is that any additional adoption prior to the second year will be balanced across network and door-to-door villages.

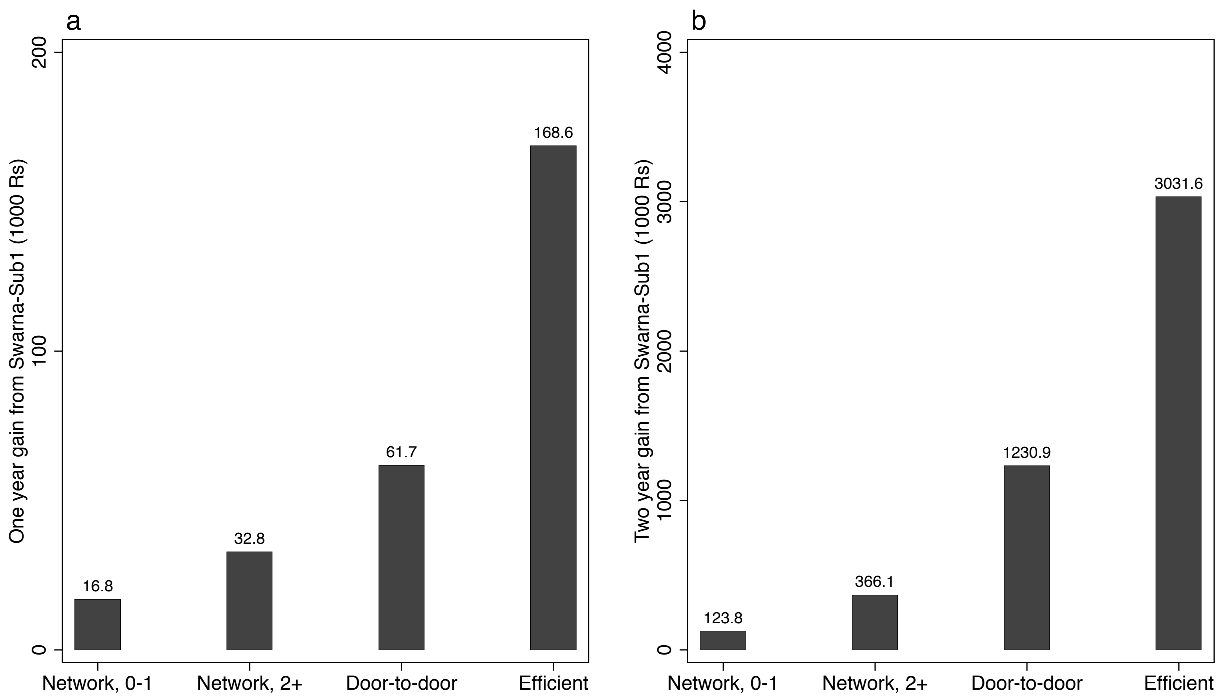
The results in Panel B of Figure 3.6 show much larger gains in expected revenue over a

[21]This threshold is used because the increased adoption in networks is strongest for farmers that have two or more suppliers with the same surname.

[22]Since cultivated area is not observed for non-adopters, it is imputed with average cultivated area of adopters when calculating the aggregate gain in expected revenue for the efficient scenario.

[23]Expected gains during the second year are discounted using a discount factor of 0.9.

Figure 3.6: Losses in expected revenue due to trading in social networks



Notes: Height of bars is the total gain in expected revenue due to adoption of Swarna-Sub1. Bar labels are as follows. Network, 0-1 and Network, 2+ refer to farmers in network villages with 0-1 suppliers having the same surname and 2 or more suppliers having the same surname, respectively. Door-to-door is for door-to-door villages and Efficient refers to a scenario where every farmer with positive expected returns adopts. **(a)** Bars represent total gain in expected revenue due to Swarna-Sub1 *during the first year of cultivation*. Cultivated area for non-adopters is imputed with average cultivated area of adopters for computation of total gains from the efficient network. **(b)** Plot displays total gain in expected revenue *over two years*. The only gains in revenue during the first year are the agronomic gains due to improved flood tolerance, i..e those in Panel A. Following the results in Chapter 2, farmers are assumed to make changes in investment patterns during the second year of cultivation. First, farmers are assumed to cultivate 0.33 hectares with Swarna-Sub1. Second, average yield of Swarna-Sub1 is expected to increase by 283 kg per ha due to investments in fertilizer and modern planting techniques. Third, farmers increase total cutivated area by 0.1 ha. The expected gains in revenue during the second year are discounted using a discount factor of 0.9.

two-year period. Specifically, the net gain with trading in networks is 220,300 Rs. When adding door-to-door sales, the total gains increase by 1.01 million Rs. The per-farmer increase in expected revenue represents approximately 30.73 USD. Thus, when considering all of the measurable benefits of the new technology, there are substantial economic losses to buyers from trading in social networks.

Heterogeneity

Are there some groups that are better off when trading occurs in networks, or is the gap between revealed demand and adoption in networks similar across the population of farmers? As shown in Table B.4, the gain in adoption from adding door-to-door sales is smaller for lower caste (SC) farmers, smaller for the better educated, but larger for those cultivating Swarna – the variety that is otherwise identical to Swarna-Sub1. Put differently, networks are relatively more effective for the lowest caste farmers, the better educated, and farmers not cultivating Swarna.

One implication of this result is that introducing door-to-door sales increases efficiency, but has a smaller effect on equity because lower caste farmers are less likely to be induced to adopt with door-to-door sales. An affirmative action policy that introduces more formal buying opportunities at the same time as favoring lower castes in seed distribution could limit the negative effects on equity because the lower caste farmers would benefit more from peer-to-peer exchange if more of the initial adopters came from their caste group.

Prices as an allocation mechanism

I next use the random variation in prices across door-to-door villages to investigate whether higher prices are more effective at allocating the technology to farmers with the highest returns. I start by estimating the degree to which the demand elasticity is dependent upon estimated returns. I then show how the average return of the pool of adopters varies with prices. Understanding whether prices can be used to more effectively allocate the technology has implications for choosing the most efficient allocation mechanism.

Table 3.7 displays demand estimates. In this analysis, the 4.5% of farmers that adopted from peers in door-to-door villages are considered as non-adopters. The purpose of this is to ensure that the demand analysis reflects only responses to the random price offers. While the linear demand estimates in column 1 imply a demand elasticity of 0.84 when price is 12 Rs per kg, a perfectly inelastic demand curve can't be rejected. This results because power is limited to detect price effects because there is significant clustering in adoption and the number of villages is small.[24] The estimated differences in demand at the lower prices are large, as shown in column 2, but the estimates remain statistically imprecise.

[24] Village-level prices were chosen to avoid perceptions of unfairness and to create a uniform price situation that more closely mimics real-world pricing. The loss in power was acceptable since estimates of demand were of secondary interest.

Table 3.7: Estimated demand functions in door-to-door sales

	(1)	(2)	(3)
Expected Returns	0.048**	0.049**	-0.003
	(0.022)	(0.022)	(0.034)
	[0.024]	[0.023]	[0.027]
Price	-0.025		
	(0.024)		
	[0.031]		
Price = 12		0.101	-0.001
		(0.081)	(0.104)
		[0.096]	[0.154]
Price = 10		0.100	0.001
		(0.095)	(0.121)
		[0.126]	[0.211]
Price=12*Expected Returns			0.086*
			(0.049)
			[0.051]
Price=10*Expected Returns			0.079*
			(0.047)
			[0.062]
Strata Fixed Effects	Yes	Yes	Yes
Household controls	Yes	Yes	Yes
Mean of Dep Variable	0.362	0.362	0.362
Number of Observations	569	569	569
R squared	0.116	0.118	0.125

Data are limited to 41 villages where door-to-door sales were made. Dependent variable is 1 if farmer adopted Swarna-Sub1 for 2013 wet season. Conventional standard errors that are clustered at the village level are reported in parentheses. Bootstrapped standard errors that correct for *Expected Returns* being a generated regressor are in brackets. Asterisks (pertaining to conventional standard errors) indicate statistical significance at the 1% ***, 5% **, and 10% * levels.

Demand is significantly more responsive to price for farmers with larger expected returns. Turning to column 3, the specification includes interaction terms between the two price indicators and estimated returns. Door-to-door sales crowd in farmers with the highest expected returns only when prices are low. The increase in adoption induced by a decrease in price from 14 to 10 is expected to be higher by 16.8 percentage points when estimated returns are at the 75th percentile as compared to when returns are zero. The order of magnitude is similar for a decrease in price from 14 to 12, suggesting that demand at low prices is fairly inelastic across the entire population.[25]

Not surprisingly given these demand estimates, increasing price does not increase the average return of adopters. The regression estimates in columns 1 of Table 3.8 show that if anything, increasing price from 10 to 14 Rs leads to a *decline* in the average return of adopters. The results again become more precise when focusing on the sample of adopters

[25]The quadratic relationship between adoption and 2011 flood intensity is also much more prevalent at low prices (see Figure B.1).

from either peers or door-to-door sales. Focusing on column 3, the average return of adopters at all three price levels is not statistically distinguishable from 1.24 quintiles per ha, which is the overall average across all farmers in the sample. Thus, charging positive prices or increasing those prices does not clearly improve targeting outcomes above the outcome that would be achieved by free distribution. However, decentralized exchange through networks does produce a better targeted pool of adopters.

Table 3.8: Effects of random price variation on the screening of adopters

	All adopters		Adopters from peers or door-to-door	
	(1) Return	(2) Flood severity (1-10)	(3) Return	(4) Flood severity (1-10)
Price=10	-0.323	-0.850	-0.464***	-0.982*
	(0.250)	(0.528)	(0.167)	(0.508)
Price=12	-0.367	-0.485	-0.507**	-0.617
	(0.284)	(0.693)	(0.215)	(0.679)
Price=14	-0.472	-1.150*	-0.612**	-1.282**
	(0.311)	(0.589)	(0.249)	(0.572)
Constant	1.742***	5.250***	1.882***	5.382***
	(0.220)	(0.464)	(0.118)	(0.442)
Mean of Dep Variable: Network	1.742	5.250	1.882	5.382
Number of Observations	266	267	264	265
R squared	0.018	0.047	0.032	0.054

Dependent variable in columns 1 and 3 is the expected return of Swarna-Sub1, measured in quintiles (1 quintile = 100kg) per hectare. Dependent variable in columns 2 and 4 is a subjective measure of the flood severity of the plot where Swarna-Sub1 was being planted. This variable ranges from 1-10 and was collected during the final follow-up survey with all potential buyers. *Door-to-door treatment* is 1 for villages where farmers had the opportunity to obtain the technology from a door-to-door salesperson. Standard errors are clustered at the village level. Asterisks indicate statistical significance at the 1% ***, 5% **, and 10% * levels.

The policy implication of this finding is that in this context, higher prices are not an effective tool for targeting adoption to farmers with the highest expected benefits. While prices are theoretically desirable as a screening tool, this argument relies heavily on the equality of willingness to pay and ability to pay. There is indeed some evidence in the literature that higher prices are an effective tool for targeting the adoption of health products (Ashraf, Berry, and Shapiro, 2010; Cohen, Dupas, and Schaner, 2013). In contrast, Cohen and Dupas (2010) find that varying price subsidies has little effect on targeting of insecticide-treated bed nets in Kenya. The results in this paper are most consistent with this finding and thus add additional evidence suggesting that in some contexts prices may not always be an effective mechanism for screening.

3.5 Why is exchange in networks inefficient?

As a final exercise, I investigate potential explanations for the inefficiency of exchange in social networks. I first present additional evidence suggesting that transactions were limited to family members and close friends and that farmers failed to establish trading links with other farmers. I then consider four alternative explanations: supply effects, quality differences, ineffective choice of suppliers, and increased salience of the technology. I find no evidence consistent with any of these alternative explanations.

Relationships limit trading

Evidence from the final survey with suppliers suggests that only close family and friends approached suppliers to obtain the technology. As displayed in Figure B.2, the most popular reason given by suppliers for not selling or exchanging seeds is that nobody asked. There are two candidate explanations: networks failed to disseminate information on identities of suppliers and farmers knew the identities of suppliers, but failed to establish trading links. The first explanation is unlikely because suppliers were publicly identified at the beginning of the experiment when seeds were disseminated via lottery.

When asked, suppliers openly recognize that existing relationships were important for choosing trading partners. Specifically, 63% and 39% report that trading partners were close friends and close family, respectively. These responses are consistent with the results in Table 3.3 showing that relationships with suppliers are more important for adoption in networks. Interestingly, suppliers clearly expected buyers to initiate trades: only 8% of suppliers reported actively seeking buyers.

In addition to the survey evidence from suppliers, followup social network data indicate that buyers did not make greater contact with suppliers. While suppliers did become more central in the network, this is almost entirely due to additional stated links with other suppliers. To establish increased importance of suppliers, I estimate

$$degree_{ij} = \beta_0 + \beta_1 supplier_{ij} + \beta_2 baselinedegree_{ij} + x_{ij}\delta + \alpha_j + \varepsilon_{ij}, \tag{3.7}$$

where $degree_{ij}$ is the number of links of farmer i in village j during the follow-up survey, $supplier_{ij}$ is an indicator for suppliers, and x_{ij} is a vector of control variables. Regression results are reported in Table 3.9. In columns 1 and 2 degree is measured as the total number of links, regardless of which farmer reported the link. Being randomly selected as a supplier of the technology leads to one additional link, which represents an approximate 14% increase. Columns 3 and 4 show that increases in in-degree – the number of links reported by *other* farmers – account for approximately half of this effect.

I use a dyadic regression model of network formation to investigate whether links at followup were concentrated between buyers and suppliers. The baseline specification is

$$link_{ikj} = \beta_0 + \beta_1 onesupplier_{ikj} + \beta_2 twosupplier_{ikj} + \alpha_j + \varepsilon_{ikj}, \tag{3.8}$$

Table 3.9: Estimated effect of being selected as a supplier on follow-up social network status

	Sharing degree		Sharing in-degree	
	(1)	(2)	(3)	(4)
Supplier	0.998***	1.003***	0.497**	0.473*
	(0.227)	(0.221)	(0.246)	(0.242)
Baseline sharing degree	0.147***	0.148***		
	(0.052)	(0.050)		
Baseline sharing in-degree			0.181***	0.180***
			(0.067)	(0.068)
Farmer is SC		-0.627		-0.872***
		(0.379)		(0.325)
Land cultivated in 2012		0.108**		0.046
		(0.044)		(0.039)
Farmer has BPL card		0.015		-0.004
		(0.156)		(0.162)
Village Fixed Effects	Yes	Yes	Yes	Yes
Mean of Dep Variable	7.28	7.29	3.92	3.92
Number of Observations	1544	1542	1547	1545
R squared	0.341	0.347	0.198	0.204

Asterisks indicate statistical significance at the 1% ***, 5% **, and 10% * levels. Standard errors are clustered at the village level. Degree is the total number of links reported by either the surveyed farmer or other farmers in her village. The in-degree is the total number of *other* farmers in the village that reported a contact with the farmer. Land cultivated in 2012 is measured in acres

where $link_{ikj}$ is an indicator variable equal to one if farmer i stated that he would go to farmer k for sharing seeds, or if farmer k made the equivalent statement for farmer i. The variables $onesupplier_{ikj}$ and $twosupplier_{ikj}$ are indicators for buyer-supplier and supplier-supplier dyads, respectively.[26] Randomization generates exogenous variation in the likelihood that a dyad consists of one or two suppliers. Therefore, both β_1 and β_2 can be interpreted causally. If buyers make new contacts with suppliers, then β_1 should be positive and large.

Results in Table 3.10 show that most of the increase in the degree of suppliers is due to links between suppliers, not links between buyers and suppliers. Specifically, two farmers that were both selected as suppliers are 18.2 percentage points – or 48% – more likely to report being linked. An intuitive explanation for the result is that farmers cultivating the same variety are more likely to go to each other for sharing information, inputs, or even seeds. Conversely, the effect of one farmer in the dyad being a supplier is small.

[26]The symmetry requirement of dyadic regressions with undirected networks is met by definition since $w_{ikj} = w_{kij}$ for all $i \neq k$ (Fafchamps and Gubert, 2007). Also, standard errors in dyadic regressions must be adjusted for correlation of error terms across observations. Observations in the same dyad are obviously correlated, leading to artificially low OLS standard errors. Fafchamps and Gubert (2007) propose a covariance matrix that corrects for correlated observations within dyads. I instead cluster the standard errors at the village level, an approach that is taken in Attanasio et al. (2012). The advantage gained from this approach is that standard errors are robust to arbitrary correlation of error terms between dyads in the same village.

Table 3.10: Dyadic regressions of network formation at follow-up

	(1)	(2)
One farmer is seller	0.013	0.022
	(0.014)	(0.015)
Both farmers are sellers	0.182***	0.207***
	(0.030)	(0.035)
Same sub-caste		0.035*
		(0.018)
Same surname		0.124***
		(0.018)
Houses within 25 m		0.006
		(0.017)
Plots within 100 m		0.009
		(0.015)
Village Fixed Effects	Yes	Yes
Mean of Dep Variable	0.380	0.385
Number of Observations	27633	24837
R squared	0.073	0.088

Data are from follow-up social network survey of all farmers. Dependent variable is 1 if either farmer in the dyad indicated a sharing link (i.e. an undirected network). Standard errors are clustered at the village level. Asterisks indicate statistical significance at the 1% ***, 5% **, and 10% * levels.

Homophily – the tendency of farmers to interact with other farmers having similar characteristics – is present in the data. Turning to the coefficient estimates in column 2, farmers belonging to the same sub-caste are 3.5 percentage points – or 9% – more likely to be linked. Similarly, farmers sharing the same surname are 12.4 percentage points – or 32% – more likely to be linked.[27] As shown in Table B.5, there is significant correlation between common surnames, sub-caste association, and geographic proximity. While networks are formed according to all of these characteristics, sharing a common surname is the most robust predictor of link formation.

Taken together, the results suggest that farmers did not invest effort in establishing trading relationships. Instead, trading was more likely to be limited to existing well-defined social groups. This tendency to transact only with close family and friends therefore explains some of the inability of trading in networks to meet demand.

[27]Similar results were found in network data from southern India (Maertens and Barrett, 2012)

Alternative Explanations

Supply effects and prices

One explanation of the ineffectiveness of trading in networks is that the quantity of seeds available to suppliers was insufficient to meet demand. If scarcity caused low adoption in networks, then having access to door-to-door sales would naturally lead to increased adoption.

The experiment was designed specifically to avoid any effects of scarcity. While only 25 kg of seed were initially provided to suppliers, the average quantity produced with this amount was approximately 1.8 tons – an amount sufficient to meet demand of approximately 180 farmers. As verification, Figure 3.7 shows the distribution of the differences between the Swarna-Sub1 harvest of suppliers during the first year and the total amount of Swarna-Sub1 planted in the village *after* door-to-door sales were made. The total amount planted by all farmers – including suppliers and other farmers outside the sample – was smaller than the total harvest in 40 of the 41 door-to-door villages. In other words, the door-to-door sales did not fill in a gap in supply that could not have been met by suppliers. The average amount harvested exceeded the amount planted by 14 times. Further, the amount harvested by suppliers was more than double the amount planted in all but two villages. Therefore, scarcity can not explain the results.

The limited number of seed transactions can not be explained by output being more valuable under alternative uses. In addition to being used as seeds, the harvest could be consumed or sold as grain for consumption. Since the eating quality of Swarna-Sub1 is identical to Swarna, and the average output price amongst farmers selling for consumption was 10.4 Rs per kg, output could have been sold or exchanged with other farmers without decreasing welfare of suppliers. These transfers were simply not made.

Price differences can not explain the results. In short, the technology was not underpriced in door-to-door sales. The price interval from 10 to 14 Rs covers the range of prices for transactions between farmers. Using the government's minimum support price of 12.5 Rs per kg as a conservative estimate of the price for direct exchanges, the average price of the technology across all farmer-to-farmer transactions was 12.4 Rs. The range of prices in door-to-door sales covers this value and therefore allows for the main effects to be estimated at a price that is equivalent to an average farmer-to-farmer transaction. Further, there is still significant demand at prices *above* the prices in farmer-to-farmer transactions, suggesting that welfare could have been improved if these transactions had been made.

Quality differences

Seed quality is the only potential product attribute that could have varied between networks and door-to-door sales. The seeds that were exchanged between farmers were second generation, i.e. output from the 2012 harvest, while the seeds sold in door-to-door sales were produced by a private seed company in a neighboring state. If farmers fail to produce qual-

Figure 3.7: Distribution of difference between total harvest of Swarna-Sub1 in year 1 and amount planted in year 2 in door-to-door villages

Notes: Data are for door-to-door villages. Figure shows the kernel density of difference between total year 1 harvest of Swarna-Sub1 by suppliers and aggregate amount of Swarna-Sub1 planted in village during year 2 (in kg). The amount planted during year 2 includes amount purchased from door-to-door sales, amount obtained directly from suppliers (by all farmers, not only farmers in the sample), and amount planted by suppliers.

ity seeds, this could potentially explain low adoption in networks.[28] Descriptively, 16% of suppliers reported that seed quality was the reason they chose not to exchange with others (Figure B.2).[29]

I use two proxy measures for quality preferences to investigate whether networks only crowd out farmers with stronger preferences for quality seeds. First, approximately 42% of farmers purchased certified seeds from local government offices for the 2012 season.[30] Given

[28]As an example, if seed is stored without proper drying, then germination ability and vigor of seedlings are negatively affected. Other practices that farmers can do to improve seed quality and purity are hand sorting to remove weeds and seeds of other varieties, winnowing to remove empty grains and chaff, and careful storage to avoid moisture absorption and damage by pests.

[29]Common reasons for poor seed quality were that drought affected production, seeds became wet during harvesting, and that Swarna-Sub1 was mixed with other rice varieties after harvesting.

[30]Seeds that are certified are produced following certain guidelines that ensure purity and higher quality.

the higher quality standards for certified seeds, this serves as a revealed preference measure of demand for seed quality. As a second measure, I use responses to a question asking whether more Swarna-Sub1 seeds would hypothetically be purchased when certified seeds are available at local government offices as compared to when seeds are only available from other villagers. I define those who indicated that a larger quantity of certified seeds would be procured as having a preference for seed quality. This group represents approximately half of the sample. If quality explains the results, then networks should crowd out farmers that either revealed or stated preferences for higher quality seeds.

There is no evidence that exchange in networks differentially crowded out farmers that preferred quality seeds. Table B.6 shows that the correlation between the two measures of quality preference and adoption in networks is small and statistically insignificant. Further, adding door-to-door sales did not lead to significantly larger increases in adoption for these farmers. Overall, the results provide suggestive evidence that differential seed quality does not explain the results.

Selection of suppliers

Another possible explanation is that adoption is low in network villages because suppliers were not selected strategically. A different method commonly used by NGO's would involve a more targeted approach of selecting the most "progressive" or "lead" farmers as initial users of the new technology. In theory, this could result in greater adoption if the more central farmers are either better at demonstrating the technology or if other farmers look to them for the best varieties to cultivate.

I exploit the random selection of suppliers to investigate whether trading in networks is more effective when suppliers are relatively more important, where importance is defined by average degree. I partition villages into two groups according to the ratio of the average degree of suppliers to that of buyers. Villages where suppliers are more central are defined as those where this ratio is greater than the sample median.[31] The regression specification is

$$adoption_{ij} = \beta_0 + \beta_1 door\ to\ door_j + \beta_2 important_j + \beta_3 door\ to\ door_j * important_j + x_{ij}\delta + \varepsilon_{ij}, \tag{3.9}$$

where $important_j$ is an indicator for villages where suppliers were relatively more important than buyers.[32]

The data rule out that networks were more effective at diffusing the technology when suppliers were more central. Focusing on column 1 of Table 3.11, the adoption rate in net-

[31] Randomization of village-level treatment was stratified by the degree ratio for purposes of investigating heterogeneity with respect to importance of suppliers. Using the ratio of average degrees carries one additional advantage since the social network in each village was only partially sampled. Chandrasekhar and Lewis (2011) show that the bias in average degree due to partial sampling of network data is proportional to the sampling rate. Using the ratio of average degrees should therefore minimize concerns regarding biases.

[32] The specification uses block fixed effects rather than strata fixed effects because randomization was stratified by block and the relative importance of suppliers.

works was 4.7 percentage points lower when suppliers were relatively more important. While the estimated coefficient is not statistically significant, large positive effects of importance of suppliers can effectively be ruled out, suggesting that the low adoption in networks is not due to the nonstrategic way in which suppliers were selected.[33] The results in column 2 show that there is no evidence that trading in networks was more effective at increasing adoption when suppliers were relatively larger farmers.

The aggregate demand revealed in door-to-door sales is however larger when suppliers are relatively more important. Returning to column 1, the predicted increase in adoption from adding door-to-door sales is 26 percentage points when suppliers are less important and 41 percentage points when suppliers are relatively more important. This approximate 60% increase in the effect is statistically significant at the 10% level. Two plausible explanations are that farmers learn more effectively from important farmers in the village or that farmers prefer to cultivate the same variety as these farmers. While recent work suggests that both channels are important (Cai et al., 2012; Bursztyn et al., 2012), separation of these channels is outside the scope of this paper.

Salience of the technology

Simply going door-to-door to sell Swarna-Sub1 could have increased awareness about the technology or sent a signal to farmers about its potential value. Increasing salience of the technology is therefore an additional possible explanation for the larger take up in door-to-door sales. Reminding farmers about the technology and its flood tolerance property during the midline survey served the purpose of reducing potential effects of increased salience.

To test salience effects, I take advantage of the fact that while door-to-door visits were only made to a randomly selected group of 15 farmers per village, it was well known that NGO staff were moving between houses to offer seeds. Houses in the sample villages are small and located in close proximity. For instance, there is an average of over two other houses in the sample within a 25 meter radius of each sample household. If door-to-door visits increased salience, then farmers that were outside of the sample would have become aware of the technology and increased purchases from suppliers.

There is no evidence of salience effects in the data. I use data from the final survey with suppliers to test whether suppliers in door-to-door villages transacted with a larger number of farmers from outside the sample. Table B.7 shows that the effect of door-to-door sales on the number of trading partners from outside the sample is negative and statistically insignificant. Moreover, increases in the number of trading partners of over approximately 30% can be rejected. These results provide some evidence that salience effects are not an important driver of the large gap between revealed demand and adoption in networks.

[33]The 95% confidence interval for β_2 is (-0.122,0.028).

Table 3.11: Heterogeneous effects according to baseline importance of suppliers

	(1)	(2)
Door-to-door treatment	0.256***	0.309***
	(0.063)	(0.065)
1 if supplier degree / buyer degree > median	-0.047	
	(0.038)	
Door-to-door treatment*1 if seller degree / buyer degree > median	0.157*	
	(0.088)	
1 if supplier size / buyer size > median		-0.057
		(0.036)
Door-to-door treatment*1 if seller size / buyer size > median		0.063
		(0.089)
Farmer is SC	-0.071*	-0.058
	(0.041)	(0.038)
Farmer has BPL card	-0.061*	-0.067**
	(0.032)	(0.033)
Land cultivated in 2012	0.004	0.005
	(0.007)	(0.007)
Ag. cooperative member	-0.025	-0.019
	(0.024)	(0.024)
Swarna user in 2012	0.074**	0.078**
	(0.032)	(0.033)
Block Fixed Effects	Yes	Yes
Mean of Dep Variable: Network	0.07	0.07
Number of Observations	1134	1119
R squared	0.199	0.195

Dependent variable is 1 if farmer adopted Swarna-Sub1 for 2013 wet season. 1 if supplier / buyer degree > median is a village-level indicator for ratio of average sharing degree of suppliers to average sharing degree of buyers being larger than the median. 1 if supplier size / buyer size > median is a similar indicator, but using average land cultivated during 2012 rather than sharing degree. *Door-to-door treatment* is 1 for villages where farmers had the opportunity to obtain the technology from a door-to-door salesperson. Land cultivated in 2012 is measured in acres. Standard errors are clustered at the village level. Asterisks indicate statistical significance at the 1% ***, 5% **, and 10% * levels.

Summary

Combining the analyses on other possible explanations, the lack of strong evidence for any of these explanations, along with the stronger peer effects in networks, suggest that barriers to exchanging with socially distant farmers represent one important explanation of the inability of decentralized trade in networks to meet demand. The pattern of existing relationships appears to prevent some transactions that otherwise would have been made if buyers and sellers had been anonymous.

3.6 Conclusions

Many products are exchanged directly between individuals that are connected in networks. Put differently, not all goods and services change hands in the textbook marketplace where the identities of buyers and sellers are irrelevant. This paper used a randomized experiment with a new agricultural technology in India to shed light on whether a system of exchanging the technology via networks is able to meet demand. The question is motivated by the idea that network structure may impede the ability of decentralized trade to allocate goods. If transacting with people from other social groups is costly or difficult, then this may present an important friction that limits the ability of buyers and sellers to come together to make transactions.

The results indicate strongly that trading in networks is inefficient. The rate of adoption of the technology was lower by 83% in networks. Trading patterns showed stronger peer effects when exchange occurred in networks. A farmer with a single additional supplier belonging to his sub-caste was approximately 50% more likely to adopt the technology when trading occurred in networks. Similarly, a farmer with one additional supplier having his surname was over twice as likely to adopt from peers. In contrast, being connected to suppliers did not have a positive effect on adoption in villages where farmers had the opportunity to purchase from door-to-door salespersons. However, targeting of farmers with higher expected returns to cultivating the technology was moderately more effective in social networks. In combination, the large decrease in adoption, combined with the only moderate improvement in targeting, cause the aggregate loss due to trading in networks to represent over 60% of the gains from exchange that were achieved with door-to-door sales.

The strong peer effects in networks are consistent with two types of trading frictions. First, there are likely non-trivial costs of interacting with farmers from other social groups. Second, if the flow of information between social groups is limited, then this could limit exchange between farmers from different groups. While information about the technology was provided to all farmers to limit the latter explanation, and five farmers demonstrated the variety in all villages, the experimental design does not fully rule out information as a barrier to exchange. Estimating the extent to which information campaigns can facilitate exchange in networks is an important area for future investigation.

The main contribution of this paper is the quantitative measure of the inefficiency of decentralized trade through networks. Such an exchange environment is common across a variety of situations, including contracting for inputs and trading of informal insurance between family and friends in developing countries. In contrast to a classic marketplace, social relationships are important for decentralized allocation of goods through social networks. Put simply, relationships matter in this exchange environment and this has negative consequences for those with limited connections to suppliers.

An important policy implication of the findings is that dissemination of new technologies using decentralized exchange through networks may be practically desirable, but it is inefficient. Introducing new seed varieties and relying on social networks for diffusion seems desirable in practice because it is an extremely low cost approach to diffusing a product. If the allocation achieved by exchange in networks is efficient, then networks could be relied upon as a sustainable method of ensuring efficient spread of technologies, particularly in the absence of anonymous markets. In terms of agricultural seed varieties, informal exchange between peers is the status quo in many remote areas where formal markets are absent. Introducing more formal channels for adoption can increase access and thus increase efficiency.

One caveat of this result is that the experiment was carried out over a single year, and thus it has little to say about the effectiveness of social networks in allocating the technology over a longer time horizon. Nonetheless, in an environment where farmers commonly learn about the benefits of new technologies from each other, there are clear benefits of having the technology demonstrated in a wide variety of conditions during the initial years. Further, there are short-run benefits to farmers from using a superior technology. My results suggest that taking a hands off approach by relying on trading in networks will leave significant demand unmet and therefore limit these short-run benefits from using the technology.

Bibliography

Abreu, D. and M. Manea. 2011. "Bargaining and efficiency in networks." *Journal of Economic Theory* .

Adamopoulos, Tasso and Diego Restuccia. 2013. "The Size Distribution of Farms and International Productivity Differences." *Working Paper* .

Alchian, A.A. and H. Demsetz. 1973. "The Property Right Paradigm." *Journal of Economic History* 33 (1):16–27.

Alderman, H., J. Hoddinot, and B. Kinsey. 2006. "Long term consequences of early childhood malnutrition." *Oxford Economic Papers* 58:450–474.

Almekinders, CJM, NP Louwaars, and GH De Bruijn. 1994. "Local seed systems and their importance for an improved seed supply in developing countries." *Euphytica* 78 (3):207–216.

Alston, L.J., G.D. Libecap, and R. Schneider. 1996. "The Determinants and Impact of Property Rights: Land Titles on the Brazilian Frontier." *Journal of Law, Economics, and Organization* 12 (1):25–61.

Angelucci, Manuela. 2012a. "Conditional Cash Transfer Programs, Credit Constraints, and Migration." *Labour* 26 (1):124–136.

———. 2012b. "Migration and Financial Constraints: Evidence from Mexico." *Working Paper* .

Angelucci, Manuela and Giacomo De Giorgi. 2009. "Indirect Effects of an Aid Program: How Do Cash Transfers Affect Ineligibles' Consumption?" *American Economic Review* 99 (1):486–508.

Aoki, Masahiko and Yujiro Hayami. 2001. *Communities and markets in economic development*. Oxford university press.

Appendini, Kirsten. 2002. "Land Regularization and Conflict Resolution: The Case of Mexico." *Land Reform, Land Settlement and Cooperatives* 2:37–50.

Ashenfelter, Orley. 1978. “Estimating the Effect of Training Programs on Earnings.” *Review of Economics and Statistics* 60 (1):47–57.

Ashraf, Nava, James Berry, and Jesse M Shapiro. 2010. “Can Higher Prices Stimulate Product Use? Evidence from a Field Experiment in Zambia.” *American Economic Review* 100:2383–2413.

Attanasio, O., A. Barr, J.C. Cardenas, G. Genicot, and C. Meghir. 2012. “Risk pooling, risk preferences, and social networks.” *American Economic Journal: Applied Economics* 4 (2):134–167.

Bailey-Serres, J., T. Fukao, P. Ronald, A. Ismail, S. Heuer, and D. Mackill. 2010. “Submergence tolerant rice: SUB1’s journey from landrace to modern cultivar.” *Rice* 3 (2):138–147.

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

Banerjee, Abhijit, Esther Duflo, and Richard Hornbeck. 2014. “Bundling Health Insurance and Microfinance in India: There Cannot be Adverse Selection if There is No Demand.” *American Economic Review Papers and Proceedings* .

Bardhan, P. and D. Mookherjee. 2011. “Subsidized Farm Input Programs and Agricultural Performance: A Farm-Level Analysis of West Bengal’s Green Revolution, 1982-1995.” *American Economic Journal: Applied Economics* 3 (4):186–214.

Benjamin, D. 1992. “Household Composition, Labor Markets, and Labor Demand: Testing for Separation in Agricultural Household Models.” *Econometrica* 60 (2):287–322.

Besley, T. 1995. “Property Rights and Investment Incentives: Theory and Evidence from Ghana.” *Journal of Political Economy* 103 (5):903–937.

Besley, T. and M. Ghatak. 2010. “Property Rights and Economic Development.” In *Handbook of Development Economics, vol. 5*. 4525–4595.

Binswanger, Hans P and Mark R Rosenzweig. 1993. “Wealth, Weather Risk and the Composition and Profitability of Agricultural Investments.” *The Economic Journal* 103 (416):56–78.

Björklund, Anders and Markus Jäntti. 1997. “Intergenerational income mobility in Sweden compared to the United States.” *The American Economic Review* 87 (5):1009–1018.

Boucher, Stephen R, Michael R Carter, and Catherine Guirkinger. 2008. “Risk rationing and wealth effects in credit markets: Theory and implications for agricultural development.” *American Journal of Agricultural Economics* 90 (2):409–423.

Bursztyn, Leonardo, Florian Ederer, Bruno Ferman, and Noam Yuchtman. 2012. "Understanding peer effects in financial decisions: Evidence from a field experiment." NBER Working Paper.

Cai, Hongbin, Yuyu Chen, Hanming Fang, and Li-An Zhou. 2012. "The Effect of Microinsurance on Economic Activities: Evidence from a Randomized Natural Field Experiment." *Working paper* .

Cai, Jing, A. de Janvry, and E. Sadoulet. 2012. "Social Networks and the Decision to Insure: Evidence from Randomized Experiments in China." Working Paper.

Capell, Teresa, Ludovic Bassie, and Paul Christou. 2004. "Modulation of the polyamine biosynthetic pathway in transgenic rice confers tolerance to drought stress." *Proceedings of the National Academy of Sciences* 101 (26):9909–9914.

Carter, Michael R, Francisco Galarza, and Stephen Boucher. 2007. "Underwriting area-based yield insurance to crowd-in credit supply and demand." *Savings and Development* :335–362.

Cassar, Alessandra, Daniel Friedman, and Patricia Higino Schneider. 2010. "A Laboratory Investigation of Networked Markets." *The Economic Journal* 120 (547):919–943.

Chandrasekhar, Arun G and Randall Lewis. 2011. "Econometrics of sampled networks." Working Paper.

Charness, G., M. Corominas-Bosch, and G.R. Frechette. 2007. "Bargaining and network structure: An experiment." *Journal of Economic Theory* 136 (1):28–65.

Chernina, Eugenia, Paul Castañeda Dower, and Andrei Markevich. 2013. "Property Rights, Land Liquidity and Internal Migration." *forthcoming, Journal of Development Economics* .

Cohen, Jessica and Pascaline Dupas. 2010. "Free Distribution or Cost-Sharing? Evidence from a Randomized Malaria Prevention Experiment." *Quarterly Journal of Economics* 125 (1):1–45.

Cohen, Jessica, Pascaline Dupas, and Simone G Schaner. 2013. "Price subsidies, diagnostic tests, and targeting of malaria treatment: Evidence from a randomized controlled trial." NBER Working Paper.

Cole, Shawn, Xavier Giné, Jeremy Tobacman, Petia Topalova, Robert Townsend, and James Vickery. 2013. "Barriers to Household Risk Management: Evidence from India." *American Economic Journal: Applied Economics* 5 (1):104–135.

Cole, Shawn, Xavier Giné, and James Vickery. 2013. "How Does Risk Management Influence Production Decisions? Evidence from a Field Experiment." .

Coman, K. 1911. "Some Unsettled Problems of Irrigation." *American Economic Review* 1 (1):1–19.

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

Cornelius, W.A. and D. Myhre. 1998. *The Transformation of Rural Mexico: Reforming the Ejido Sector*. Center for U.S.-Mexican Studies, University of California, San Diego.

Corominas-Bosch, Margarida. 2004. "Bargaining in a network of buyers and sellers." *Journal of Economic Theory* 115 (1):35–77.

Dar, Manzoor H, Alain de Janvry, Kyle Emerick, David Raitzer, and Elisabeth Sadoulet. 2013. "Flood-tolerant rice reduces yield variability and raises expected yield, differentially benefitting socially disadvantaged groups." *Scientific Reports* 3.

de Brauw, Alan and Valerie Mueller. 2012. "Do Limitations in Land Rights Transferability Influence Mobility Rates in Ethiopia?" *Journal of African Economies* 21 (4):548–579.

de Ita, A. 2006. *Land Concentration in Mexico after PROCEDE*. Oakland, CA: Institute for Food and Development Policy.

de Janvry, A., M. Gonzalez-Navarro, and E. Sadoulet. 2013. "Are Land Reforms Granting Complete Property Rights Politically Risky? Electoral Outcomes of Mexico's Certification Program." *forthcoming, Journal of Development Economics* .

de Janvry, A., G. Gordillo, and E. Sadoulet. 1997. *Mexico's Second Agrarian Reform: Household and Community Responses, 1990-1994*. Center for US-Mexican Studies, University of California, San Diego.

De Soto, H. 1989. *The Other Path: The Informal Revolution*. New York: Basic Books.

De Soto, Hernando. 2000. *The Mystery of Capital: Why Capitalism Triumphs in the West and Fails Everywhere Else*. New York: Basic books.

De Weerdt, J. and S. Dercon. 2006. "Risk-sharing networks and insurance against illness." *Journal of Development Economics* 81 (2):337–356.

Deere, Carmen Diana and Magdalena Léon. 2001. *Empowering women: Land and property rights in Latin America*. Univ of Pittsburgh Press.

Deininger, Klaus. 2003. *Land Policies for Growth and Poverty Reduction*. Washington D.C.: World Bank Publications.

Deininger, Klaus and F. Bresciani. 2001. "Mexico's Second Agrarian Reform: Implementation and Impact." *World Bank, Washington D.C.* .

Demsetz, H. 1967. "Toward a Theory of Property Rights." *American Economic Review* 57 (2):347–359.

Dercon, Stefan and Luc Christiaensen. 2011. "Consumption risk, technology adoption and poverty traps: evidence from Ethiopia." *Journal of Development Economics* 96 (2):159–173.

Dercon, Stefan, Ruth Vargas Hill, Daniel Clarke, Ingo Outes-Leon, and Alemayehu Seyoum Taffesse. 2014. "Offering rainfall insurance to informal insurance groups: evidence from a field experiment in Ethiopia." *Journal of Development Economics* 106:132–143.

Donovan, Kevin. 2014. "Agricultural Risk, Intermediate Inputs, and Cross-Country Productivity Differences." *Working Paper* .

Duflo, Esther, Michael Kremer, and Jonathan Robinson. 2011. "Nudging Farmers to Use Fertilizer: Theory and Experimental Evidence from Kenya." *American Economic Review* 101 (6):2350–2390.

Elliott, Matthew. 2013. "Inefficiencies in networked markets." Working Paper.

Fafchamps, M. and F. Gubert. 2007. "The formation of risk sharing networks." *Journal of Development Economics* 83 (2):326–350.

Fafchamps, M. and S. Lund. 2003. "Risk-sharing networks in rural Philippines." *Journal of Development Economics* 71 (2):261–287.

Feder, G. and D. Feeny. 1991. "Land Tenure and Property Rights: Theory and Implications for Development Policy." *The World Bank Economic Review* 5 (1):135–153.

Feder, G., R.E. Just, and D. Zilberman. 1985. "Adoption of agricultural innovations in developing countries: A survey." *Economic Development and Cultural Change* 33 (2):255–298.

Feder, G., T. Onchan, and Y. Chalamwong. 1988. "Land Policies and Farm Performance in Thailand's Forest Reserve Areas." *Economic Development and Cultural Change* 36 (3):483–501.

Field, E. 2007. "Entitled to Work: Urban Property Rights and Labor Supply in Peru." *Quarterly Journal of Economics* 122 (4):1561–1602.

Field, E. and M. Torero. 2006. "Do Property Titles Increase Credit Access Among the Urban Poor? Evidence from a Nationwide Titling Program." *Harvard University Department of Economics, Working Paper* .

Fogli, Alessandra and Laura Veldkamp. 2012. "Germs, social networks and growth." NBER Working Paper.

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

———. 2010. "Microeconomics of technology adoption." *Annu. Rev. Econ.* 2 (1):395–424.

Fujino, Kenji, Hiroshi Sekiguchi, Yasuyuki Matsuda, Kazuhiko Sugimoto, Kazuko Ono, and Masahiro Yano. 2008. "Molecular identification of a major quantitative trait locus, qLTG3–1, controlling low-temperature germinability in rice." *Proceedings of the National Academy of Sciences* 105 (34):12623–12628.

Fukao, Takeshi and Julia Bailey-Serres. 2008. "Submergence tolerance conferred by Sub1A is mediated by SLR1 and SLRL1 restriction of gibberellin responses in rice." *Proceedings of the National Academy of Sciences* 105 (43):16814–16819.

Galasso, Emanuela and Martin Ravallion. 2005. "Decentralized targeting of an antipoverty program." *Journal of Public economics* 89 (4):705–727.

Gale, D.M. and S. Kariv. 2009. "Trading in networks: A normal form game experiment." *American Economic Journal: Microeconomics* :114–132.

Galiani, S. and E. Schargrodsky. 2010. "Property Rights for the Poor: Effects of Land Titling." *Journal of Public Economics* 94 (9–10):700–729.

Galiani, Sebastian and Ernesto Schargrodsky. 2011. "Land Property Rights and Resource Allocation." *Journal of Law and Economics* 54 (4):S329–S345.

Giles, J. and R. Mu. 2011. "Village Political Economy, Land Tenure Insecurity and the Rural to Urban Migration Decision: Evidence from China." *Working Paper* .

Giné, Xavier, Robert Townsend, and James Vickery. 2008. "Patterns of rainfall insurance participation in rural India." *The World Bank Economic Review* 22 (3):539–566.

Giné, Xavier and Dean Yang. 2009. "Insurance, credit, and technology adoption: Field experimental evidencefrom Malawi." *Journal of Development Economics* 89 (1):1–11.

Goldstein, M. and C. Udry. 2008. "The Profits of Power: Land Rights and Agricultural Investment in Ghana." *Journal of Political Economy* 116 (6):981–1022.

Gollin, Douglas, David Lagakos, and Michael Waugh. 2012. "The Agricultural Productivity Gap in Developing Countries." *Working Paper* .

Hamilton, S. 2002. "Neoliberalism, Gender, and Property Rights in Rural Mexico." *Latin American Research Review* 37 (1):119–143.

Hattori, Yoko, Keisuke Nagai, Shizuka Furukawa, Xian-Jun Song, Ritsuko Kawano, Hitoshi Sakakibara, Jianzhong Wu, Takashi Matsumoto, Atsushi Yoshimura, Hidemi Kitano et al. 2009. "The ethylene response factors SNORKEL1 and SNORKEL2 allow rice to adapt to deep water." *Nature* 460 (7258):1026–1030.

Heath, J.R. 1990. "Enhancing the Contribution of Land Reform to Mexican Agricultural Development." *World Bank Policy Research Working Paper* .

Inoue, Atsushi and Gary Solon. 2010. "Two-sample instrumental variables estimators." *The Review of Economics and Statistics* 92 (3):557–561.

Jackson, Matthew and Yves Zenou. 2013. "Games on networks." *forthcoming, Handbook of Game Theory* .

Jackson, Matthew O. 2009. "Networks and economic behavior." *Annu. Rev. Econ.* 1 (1):489–511.

Karaba, Aarati, Shital Dixit, Raffaella Greco, Asaph Aharoni, Kurniawan R Trijatmiko, Nayelli Marsch-Martinez, Arjun Krishnan, Karaba N Nataraja, Makarla Udayakumar, and Andy Pereira. 2007. "Improvement of water use efficiency in rice by expression of HARDY, an Arabidopsis drought and salt tolerance gene." *Proceedings of the National Academy of Sciences* 104 (39):15270–15275.

Karlan, Dean, Robert Darko Osei, Isaac Osei-Akoto, and Christopher Udry. 2013. "Agricultural decisions after relaxing credit and risk constraints." Tech. rep., National Bureau of Economic Research.

Khush, Gurdev S. 1997. "Origin, dispersal, cultivation and variation of rice." In *Oryza: From Molecule to Plant*. Springer, 25–34.

Kranton, Rachel E. 1996. "Reciprocal exchange: A self-sustaining system." *The American Economic Review* :830–851.

Kranton, R.E. and D.F. Minehart. 2001. "A theory of buyer-seller networks." *The American Economic Review* :485–508.

Kremer, M. and E. Miguel. 2007. "The Illusion of Sustainability." *The Quarterly Journal of Economics* 122 (3):1007–1065.

Kurosaki, Takashi and Marcel Fafchamps. 2002. "Insurance market efficiency and crop choices in Pakistan." *Journal of Development Economics* 67 (2):419–453.

Levy, S. and S. Van Wijnbergen. 1995. "Transition Problems in Economic Reform: Agriculture in the North American Free Trade Agreement." *American Economic Review* 85 (4):738–754.

Lin, J.Y. 1992. “Rural Reforms and Agricultural Growth in China.” *American Economic Review* 82 (1):34–51.

Maccini, Sharon and Dean Yang. 2009. “Under the weather: Health, schooling, and economic consequences of early-life rainfall.” *The American Economic Review* :1006–1026.

Mackill, DJ, AM Ismail, US Singh, RV Labios, and TR Paris. 2012. “Development and Rapid Adoption of Submergence-Tolerant (Sub1) Rice Varieties.” *Advances in Agronomy* 115:299.

Maertens, Annemie and Christopher Barrett. 2012. “Measuring Social Networks Effects on Agricultural Technology Adoption.” *American Journal of Agricultural Economics* 95 (2):353–359.

Magaloni, Beatriz. 2006. *Voting for Autocracy: Hegemonic Party Survival and its Demise in Mexico*. Cambridge University Press.

Manea, M. 2011. “Bargaining in stationary networks.” *The American Economic Review* 101 (5):2042–2080.

Manski, C.F. 1993. “Identification of endogenous social effects: The reflection problem.” *The Review of Economic Studies* 60 (3):531–542.

Mazzocco, Maurizio. 2012. “Testing efficient risk sharing with heterogeneous risk preferences.” *The American Economic Review* 102 (1):428–468.

McKenzie, D. and H. Rapoport. 2007. “Network Effects and the Dynamics of Migration and Inequality: Theory and Evidence from Mexico.” *Journal of Development Economics* 84 (1):1–24.

Milligan, Kevin, Enrico Moretti, and Philip Oreopoulos. 2004. “Does education improve citizenship? Evidence from the United States and the United Kingdom.” *Journal of Public Economics* 88 (9):1667–1695.

Mobarak, A Mushfiq and Mark Rosenzweig. 2012. “Selling formal insurance to the informally insured.” *Yale University Economic Growth Center Discussion Paper* .

Mobarak, Ahmed Mushfiq and Mark Rosenzweig. 2014. “Risk, Insurance and Wages in General Equilibrium.” Tech. rep., National Bureau of Economic Research.

Morduch, Jonathan. 1995. “Income smoothing and consumption smoothing.” *The journal of economic perspectives* 9 (3):103–114.

Munshi, K. 2004. “Social learning in a heterogeneous population: technology diffusion in the Indian Green Revolution.” *Journal of Development Economics* 73 (1):185–213.

Munshi, Kaivan. 2003. "Networks in the Modern Economy: Mexican Migrants in the U.S. Labor Market." *Quarterly Journal of Economics* 118 (2):549–599.

Murphy, Kevin M and Robert H Topel. 1985. "Estimation and inference in two-step econometric models." *Journal of Business & Economic Statistics* 3 (4):88–97.

Navarro, Zander. 2009. "Expropriating Land in Brazil." In *Agricultural Land Redistribution: Toward Greater Consensus*, edited by Hans Binwanger-Mkhize, Camille Bourguignon, and Rogier van den Brink. The World Bank.

Neeraja, CN, R Maghirang-Rodriguez, A Pamplona, S Heuer, BCY Collard, EM Septiningsih, G Vergara, D Sanchez, K Xu, AM Ismail et al. 2007. "A marker-assisted backcross approach for developing submergence-tolerant rice cultivars." *Theoretical and Applied Genetics* 115 (6):767–776.

Nelson, Donald E, Peter P Repetti, Tom R Adams, Robert A Creelman, Jingrui Wu, David C Warner, Don C Anstrom, Robert J Bensen, Paolo P Castiglioni, Meghan G Donnarummo et al. 2007. "Plant nuclear factor Y (NF-Y) B subunits confer drought tolerance and lead to improved corn yields on water-limited acres." *Proceedings of the National Academy of Sciences* 104 (42):16450–16455.

Nishiguchi, Toshihiro. 1994. *Strategic industrial sourcing: The Japanese advantage*. Oxford University Press, USA.

North, Douglass and Robert Thomas. 1973. *The Rise of the Western World: A New Economic History*. Cambridge: Cambridge University Press.

Oster, E. and R. Thornton. 2012. "Determinants of technology adoption: Peer effects in menstrual cup take-up." *Journal of the European Economic Association* 10 (6):1263–1293.

Rao, AN, DE Johnson, B Sivaprasad, JK Ladha, and AM Mortimer. 2007. "Weed management in direct-seeded rice." *Advances in Agronomy* 93:153–255.

Restuccia, Diego, Dennis Tao Yang, and Xiaodong Zhu. 2008. "Agriculture and aggregate productivity: A quantitative cross-country analysis." *Journal of Monetary Economics* 55 (2):234–250.

Rozelle, Scott and Guo Li. 1998. "Village leaders and land-rights formation in China." *American Economic Review* 88 (2):433–438.

Sanderson, Susan W. 1984. *Land Reform in Mexico*. Orlando (USA): Academic Press.

Singh, S., D.J. Mackill, and A.M. Ismail. 2009. "Responses of SUB1 rice introgression lines to submergence in the field: yield and grain quality." *Field Crops Research* 113 (1):12–23.

Singh, Sudhanshu, David J Mackill, and Abdelbagi M Ismail. 2011. "Tolerance of longer-term partial stagnant flooding is independent of the SUB1 locus in rice." *Field Crops Research* 121 (3):311–323.

Sperling, Louise and Michael E Loevinsohn. 1993. "The dynamics of adoption: distribution and mortality of bean varieties among small farmers in Rwanda." *Agricultural Systems* 41 (4):441–453.

Stecklov, Guy, Paul Winters, Marco Stampini, and Benjamin Davis. 2005. "Do Conditional Cash Transfers Influence Migration? A Study Using Experimental Data from the Mexican Progresa Program." *Demography* 42 (4):769–790.

Stephen, L. 1996. "Too Little, Too Late? The Impact of Article 27 on Women in Oaxaca." In *Reforming Mexico's Agrarian Reform*, edited by Laura Randall. 289–303.

Suri, T. 2011. "Selection and comparative advantage in technology adoption." *Econometrica* 79 (1):159–209.

Valsecchi, M. 2012. "Land Certification and International Migration: Evidence from Mexico." *Working Paper* .

Vignes, A. and J.M. Etienne. 2011. "Price formation on the Marseille fish market: Evidence from a network analysis." *Journal of Economic Behavior & Organization* .

Voesenek, Laurentius ACJ and Julia Bailey-Serres. 2009. "Plant biology: Genetics of high-rise rice." *Nature* 460 (7258):959–960.

Xu, K., X. Xu, T. Fukao, P. Canlas, R. Maghirang-Rodriguez, S. Heuer, A.M. Ismail, J. Bailey-Serres, P.C. Ronald, and D.J. Mackill. 2006. "Sub1A is an ethylene-response-factor-like gene that confers submergence tolerance to rice." *Nature* 442 (7103):705–708.

Yates, P.L. 1981. "Mexican Land Reform, 1959-1969: A Comment." *Economic Journal* 91 (363):745–752.

Appendix A

Appendix to Chapter 1

This appendix provides additional details on construction of some of the data used in the analysis for Chapter 1.

Progresa Data

Household level migration was taken from the 1998-2000 fall versions of the ENCEL survey. The survey was conducted each fall from 1998-2000 in the 506 localities that were part of the experimental evaluation of Progresa. Since no ejido identifiers were included in these data, we matched the 506 localities to ejidos using a spatial join in ARCGIS. We only observe the coordinates of the centroid of each locality and therefore match localities to ejidos if the center of the locality is located inside the boundaries of the ejido. The digital maps of all ejidos certified from 1993-2006 were obtained from RAN. The spatial merge resulted in 234 of the localities falling into one of 219 different ejidos.[1] The number of households from the 1998 survey that fell inside ejidos as a result of this process is 13,212. Another 4,893 households were removed from the sample as a result of being in ejidos that were certified before 1997. Since permanent migration is being measured, trends in migration are unlikely to be the same in ejidos certified prior to 1997 as those certified later. These ejidos are removed for this reason. It is also important to note that the spatial matching approach does not result in a perfect match between households and ejidos. It is possible that while the centroid of a locality falls into a particular ejido, the outskirts of the locality fall into a different ejido. This is more likely to be an issue in localities that are large. We used census population data to construct the ratio of the population of the locality to the number of ejidatarios in the matched ejido. The matching is more likely to be inaccurate when the locality is large relative to the ejido. We therefore retained only the 200 localities with the lowest values of this metric. This amounted to removing an additional 742 households from the sample. The total number of ejidos in the sample is 127.

[1] This number is roughly consistent with half of Mexico's land being in ejidos. The large number of localities that were not matched to ejidos is therefore not a concern. The matching rate of 46% is actually in line with 50% of land being in ejidos.

1991 and 2007 Ejido Census

The 1991 and 2007 ejido censuses consist of a set of 28,752 ejidos that were surveyed in both 1991 and 2007. We were unable to obtain the name of each ejido due to confidentiality concerns. Further, the 2007 census did not contain information on the time of completion of Procede. A matching process was therefore necessary to make these data usable. The key information used were the state, municipality, and name of the locality where the majority of the ejidatarios live. We used this information along with some common key variables between the census data and the GIS database from RAN to match ejidos based on a 4-step process:

1. There were 22,473 ejidos for which the locality where a majority of the ejidatarios live is located inside the boundaries of the ejido. For these ejidos we were able to use our spatial merge between localities and ejidos to identify the corresponding ejido in the GIS database. There are of course numerous instances where the boundaries of an ejido contain more than one locality centroid. We were unable to include these ejidos in this matching round. This round matched a total of 14,128 ejidos.

2. The second round of matching is meant to partially correct for the fact that matching localities to ejidos in the previous step using only the centroid of the locality is imperfect. The reason for this is that the centroid of the locality could fall outside of the boundaries of the ejido even if there is substantial overlap between the locality and ejido. Further, ejidos with multiple disjoint patches of land pose problems to matching based on locality centroids and ejido boundaries. The distance between the locality centroid for each unmatched census ejido and the center of each unmatched ejido from the GIS database was calculated using a simple distance calculation in ARCGIS. An ejido from the GIS data was matched to an ejido from the census data if the locality where the majority of the ejidatarios live was the closest locaelity to the center of the ejido. Since this match is not perfect, we attempt to minimize errors by only retaining matches where the percentage difference between the number of ejidatarios in the 1991 census and the GIS database was between -46.8% and 29%.[2] This round generated an additional 1,787 matches.

3. In this round we considered the remaining unmatched ejidos for which the locality where the majority of the ejidatarios live is located inside the boundaries of the ejido. We defined a potential candidate match from the GIS database as an unmatched ejido that was located in the same state and municipality. For each of these potential matches we considered 4 metrics of comparison. The first was the similarity between the name of the locality where the ejidatarios live and the name of the ejido in the GIS database.[3] We generated a spelling similarity index using a combination of the COMPARE and

[2] These numbers were chosen as the 10th and 90th percentiles of the percentage difference from the ejidos matched in the previous round.

[3] It is common for ejido names to be the same as locality names in Mexico.

SPEDIS functions in SAS. A match was identified for sufficiently low values of this index. The second metric was the distance between the centroid of the locality and ejido. The ejidos were considered to match if the distance was less than 5.1 kilometers.[4] The third metric was the number of ejidatarios. A match was determined using the same cutoffs as in the previous round. The final metric was the difference between the size of the ejido (in hectares) in the two datasets. The percentage cutoffs were -32.4 and 41.6. We required at least two of these criteria to be satisfied to identify a match between the ejidos. For each census ejido we selected the ejido from the GIS database which matched on the most of these criteria (from 2 to 4). In order to break ties we used the percentage difference in the number of ejidatarios. This round generated a total of 1,878 matches.

4. The fourth round of matches considers the census ejidos where it was stated that the locality where the majority of ejidatarios live is *not* inside the boundaries of the ejido. We used a similar process as in the previous round with only two modifications. First, similarities between the name of the locality and the ejido were not used. Second, the distance requirement was relaxed to 8.6 kilometers (25th percentile). This round generated 1,920 matches.

[4]This value was chosen since it was the 10th percentile in the list of candidate matches.

Appendix B

Appendix to Chapter 3

This appendix provides additional derivations and analysis for Chapter 3.

Derivation of expected returns of adopters

The expected return of adopters in social networks is $E(R|R+u-c > v+\underline{c})$. Using properties of the multivariate normal distribution, this is written as

$$E(R|R+u-c > v+\underline{c}) = \mu_R+\sigma_R\tilde{\rho}E\left(\frac{R+u-c-\mu_r+\mu_c}{\sqrt{\sigma_R^2+\sigma_c^2+\sigma_u^2-2\rho\sigma_c\sigma_R}}\middle|R+u-c > v+\underline{c}\right), \quad \text{(B.1)}$$

where $\tilde{\rho}$ is the correlation between R and $R+u-c$. This expression can be rewritten as

$$\begin{aligned} E(R|R+u-c > v+\underline{c}) =& \mu_R + \frac{\sigma_R}{\sqrt{\sigma_R^2+\sigma_c^2+\sigma_u^2-2\rho\sigma_c\sigma_R}}\tilde{\rho}E(R+u-c|R+u-c > v+\underline{c}) \\ & - \frac{\tilde{\rho}\sigma_R(\mu_R-\mu_c)}{\sqrt{\sigma_R^2+\sigma_c^2+\sigma_u^2-2\rho\sigma_c\sigma_R}}. \end{aligned} \quad \text{(B.2)}$$

Given that $R+u-c$ is distributed normally, $E(R+u-c|R+u-c > v+\underline{c})$ can be written as

$$\mu_R-\mu_c+\sqrt{\sigma_R^2+\sigma_u^2+\sigma_c^2-2\rho\sigma_R\sigma_c} * M\left(\frac{v+\underline{c}-\mu_R+\mu_c}{\sqrt{\sigma_R^2+\sigma_u^2+\sigma_c^2-2\rho\sigma_R\sigma_c}}\right), \quad \text{(B.3)}$$

where $M(z) = \frac{\phi(z)}{1-\Phi(z)}$ is the inverse Mill's ratio. Reinserting this into Equation B.2 gives

$$E(R|R+u-c > v+\underline{c}) = \mu_R+\tilde{\rho}\sigma_R M\left(\frac{v+\underline{c}-\mu_R+\mu_c}{\sqrt{\sigma_R^2+\sigma_u^2+\sigma_c^2-2\rho\sigma_R\sigma_c}}\right). \quad \text{(B.4)}$$

Since $\tilde{\rho}$ is the correlation between R and $R+u-c$, $\tilde{\rho}$ simplifies to

$$\tilde{\rho} = \frac{\sigma_R-\rho\sigma_c}{\sqrt{\sigma_R^2+\sigma_u^2+\sigma_c^2-2\rho\sigma_R\sigma_c}}. \quad \text{(B.5)}$$

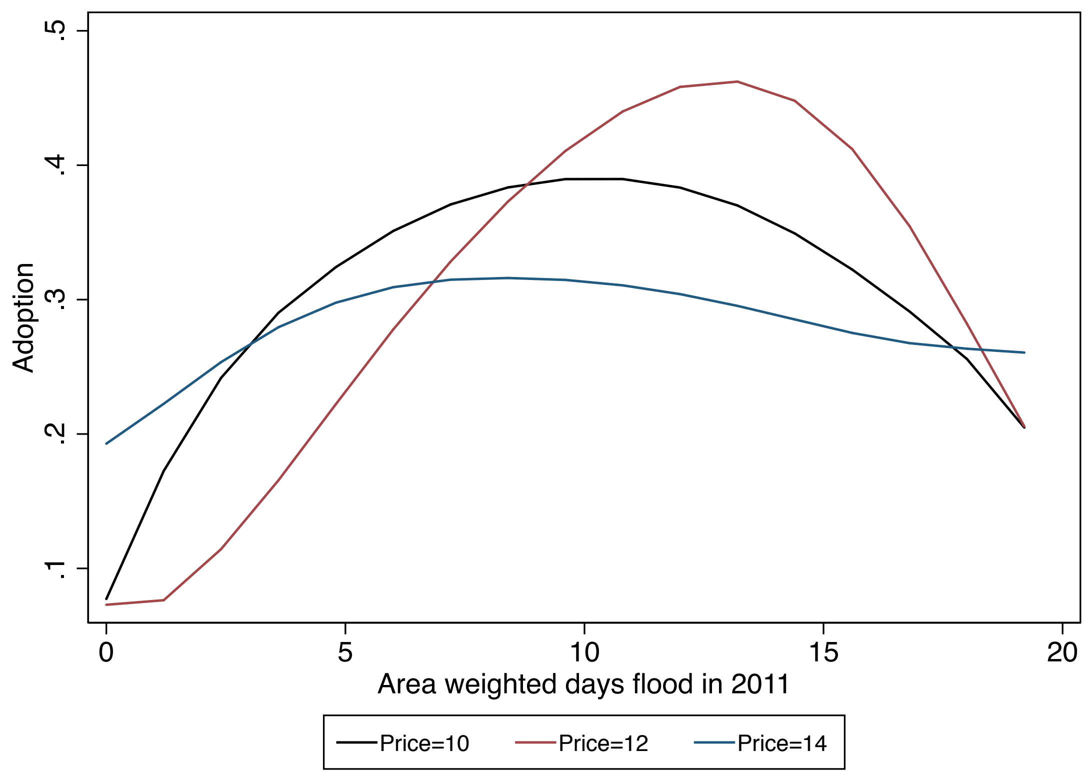
Combining Equations B.4 and B.5,

$$E(R|R+u-c>v) = \mu_R + \frac{\sigma_R(\sigma_R - \rho\sigma_c)}{\sqrt{\sigma_R^2+\sigma_u^2+\sigma_c^2-2\rho\sigma_R\sigma_c}} * M\left(\frac{v+\underline{c}-\mu_R+\mu_c}{\sqrt{\sigma_R^2+\sigma_u^2+\sigma_c^2-2\rho\sigma_R\sigma_c}}\right). \tag{B.6}$$

This establishes the result. A similar derivation is used to verify the formula for $E(R|R+u > v)$ in the main text.

Additional Figures and Tables

Figure B.1: Nonparametric relationship between flooding intensity in 2011 and adoption for 3 different price levels



Notes: Figure shows estimates from nonparametric fan regressions of adoption on area weighted days flood in 2011. Data are limited to door-to-door villages.

Table B.1: Robustness of estimated peer effects to different subsamples and nonlinear model

	Variation in adoption		Drop Dhamanagar block		Full sample	
	(1) OLS	(2) OLS	(3) OLS	(4) OLS	(5) Probit	(6) Probit
Door-to-door Treatment	0.268***	0.206**	0.316***	0.328***	0.349***	0.357***
	(0.080)	(0.095)	(0.069)	(0.085)	(0.055)	(0.059)
Door-to-door Treatment * Number suppliers w/ same surname	-0.159**		-0.155***		-0.076**	
	(0.061)		(0.052)		(0.036)	
Door-to-door Treatment * Total number w/ same surname	0.034**		0.048***		0.021	
	(0.016)		(0.016)		(0.014)	
Number suppliers w/ same surname	0.110**		0.061		0.027*	
	(0.053)		(0.038)		(0.016)	
Total number w/ same surname	-0.015		-0.017		-0.010	
	(0.011)		(0.011)		(0.013)	
Door-to-door Treatment * Number suppliers same sub-caste		-0.120**		-0.092**		-0.063**
		(0.045)		(0.035)		(0.027)
Door-to-door Treatment * Total number same sub-caste		0.031*		0.026***		0.017
		(0.017)		(0.010)		(0.012)
Number suppliers same sub-caste		0.075*		0.053**		0.024*
		(0.041)		(0.023)		(0.013)
Total number same sub-caste		-0.024		-0.018**		-0.014
		(0.017)		(0.007)		(0.012)
Strata Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Household controls	Yes	Yes	Yes	Yes	Yes	Yes
Mean of Dep Variable: Network	0.18	0.18	0.09	0.09	0.07	0.07
Number of Observations	744	744	800	800	1134	1134
R squared	0.120	0.118	0.204	0.197		

Data in columns 1 and 2 are limited to villages where at least one farmer adopted Swarna-Sub1 for 2013 wet season. Data in columns 3 and 4 are for villages in Chandabali and Tihidi blocks. Dependent variable in all columns is 1 if farmer adopted Swarna-Sub1 for 2013 wet season. Columns 3 and 4 present marginal effects calculated from probit coefficients, along with standard errors calculated from the delta method. *Door-to-door Treatment* is 1 for villages where farmers could either obtain the technology from door-to-door sales or from peer suppliers. Control variables are indicator for SC, indicator for holding BPL card, land area cultivated in 2012 wet season, indicator for member of agricultural cooperative, and indicator for Swarna cultivator in 2012 wet season. Standard errors are clustered at the village level. Asterisks indicate statistical significance at the 1% ***, 5% **, and 10% * levels.

Table B.2: Robustness of estimated peer effects to measurement of peer influence in shares rather than levels

	(1)	(2)	(3)	(4)	(5)	(6)
Door-to-door Treatment	0.435***	0.439***		0.435***	0.439***	
	(0.049)	(0.047)		(0.057)	(0.054)	
Door-to-door Treatment * Share of same surname that are suppliers	-0.364***	-0.373***	-0.346***			
	(0.115)	(0.112)	(0.130)			
Share of same surname that are suppliers	0.207**	0.206**	0.202**			
	(0.079)	(0.080)	(0.095)			
Door-to-door Treatment * Share of same sub-caste that are suppliers				-0.384**	-0.398**	-0.411**
				(0.172)	(0.167)	(0.184)
Share of same sub-caste that are suppliers				0.175**	0.125	0.174*
				(0.082)	(0.090)	(0.099)
Strata Fixed Effects	Yes	Yes	No	Yes	Yes	No
Village Fixed Effects	No	No	Yes	No	No	Yes
Household controls	No	Yes	Yes	No	Yes	Yes
Mean of Dep Variable: Network	0.07	0.07	0.07	0.07	0.07	0.07
Number of Observations	1009	1008	1008	1056	1055	1055
R squared	0.202	0.220	0.435	0.199	0.218	0.434

Dependent variable in all columns is 1 if farmer adopted Swarna-Sub1 for 2013 wet season. *Door-to-door Treatment* is 1 for villages where farmers could either obtain the technology from door-to-door sales or from peer suppliers. Control variables are indicator for SC, indicator for holding BPL card, land area cultivated in 2012 wet season, indicator for member of agricultural cooperative, and indicator for Swarna cultivator in 2012 wet season. Standard errors are clustered at the village level. Asterisks indicate statistical significance at the 1% ***, 5% **, and 10% * levels.

Table B.3: Estimated peer effects using stated social networks at followup

	(1)	(2)
Door-to-door Treatment	0.316***	
	(0.066)	
Door-to-door Treatment * Followup links with suppliers	0.009	-0.019
	(0.027)	(0.021)
Door-to-door Treatment * Followup degree	0.001	0.004
	(0.010)	(0.008)
Followup links with suppliers	0.002	0.014
	(0.015)	(0.013)
Followup degree	0.006	0.001
	(0.005)	(0.003)
Strata Fixed Effects	Yes	No
Village Fixed Effects	No	Yes
Household controls	Yes	Yes
Mean of Dep Variable: Network	0.07	0.07
Number of Observations	1134	1134
R squared	0.207	0.413

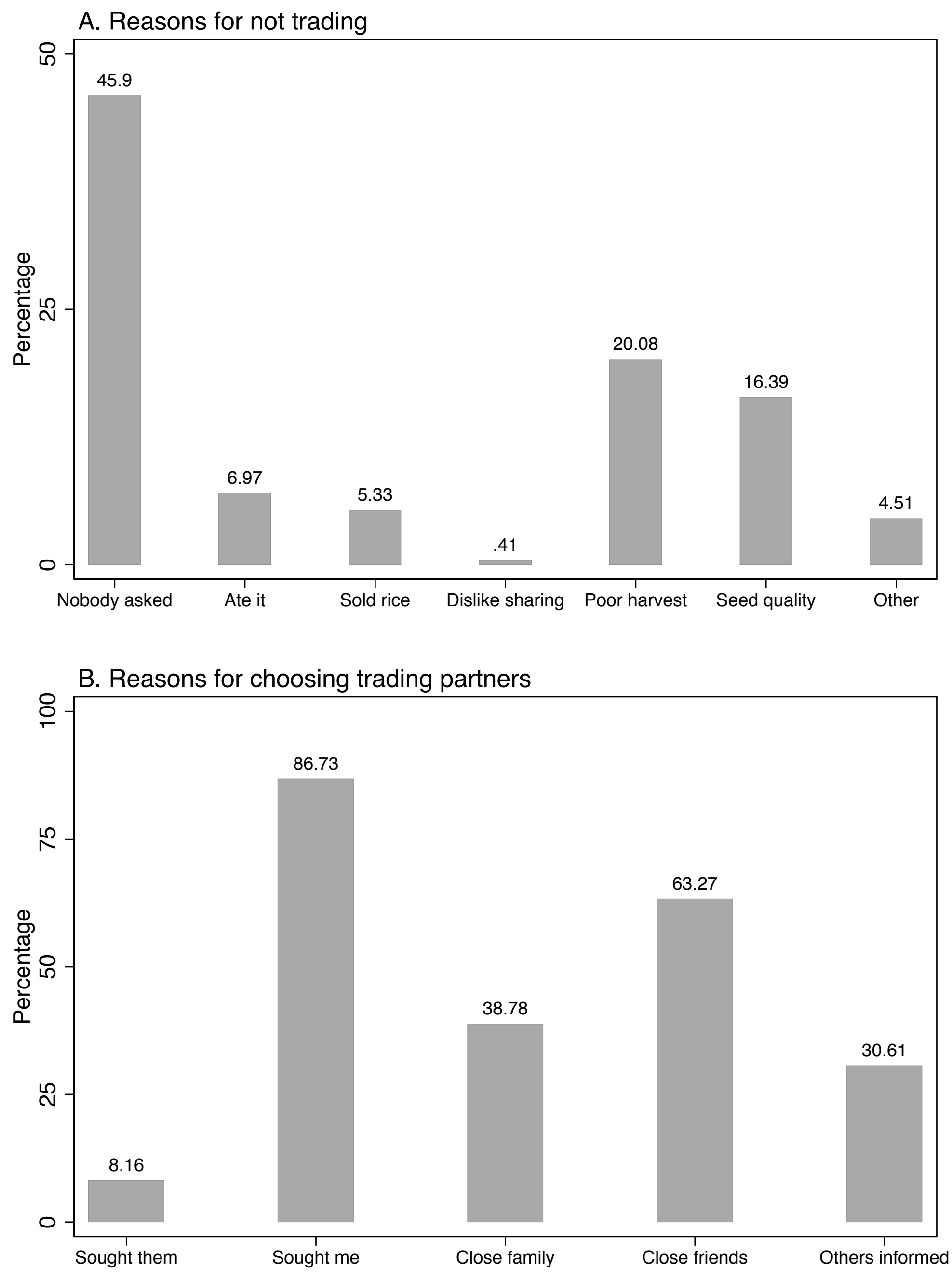
Dependent variable is 1 if farmer adopted Swarna-Sub1 for 2013 wet season. *Door-to-door Treatment* is 1 for villages where farmers could either obtain the technology from door-to-door sales or from peer suppliers. Control variables are indicator for SC, indicator for holding BPL card, land area cultivated in 2012 wet season, indicator for member of agricultural cooperative, and indicator for Swarna cultivator in 2012 wet season. Standard errors are clustered at the village level. Asterisks indicate statistical significance at the 1% ***, 5% **, and 10% * levels.

Table B.4: Heterogeneity in adoption effects by household characteristics

	(1)
Door-to-door treatment	0.409***
	(0.093)
Farmer is SC	0.016
	(0.044)
Farmer has BPL card	-0.014
	(0.033)
Land cultivated in 2012	0.007
	(0.006)
Ag. cooperative member	-0.020
	(0.027)
Swarna user in 2012	0.032
	(0.026)
Education above primary	-0.006
	(0.021)
Door-to-door treatment interacted with:	
Farmer is SC	-0.197**
	(0.076)
Farmer has BPL card	-0.103
	(0.065)
Land cultivated in 2012	-0.001
	(0.014)
Ag. cooperative member	0.009
	(0.046)
Swarna user in 2012	0.115*
	(0.068)
Education above primary	-0.114**
	(0.048)
Strata Fixed Effects	Yes
Mean of Dep Variable: Network	0.07
Number of Observations	1131
R squared	0.224

Dependent variable is 1 if farmer adopted Swarna-Sub1 for 2013 wet season. *Door-to-door Treatment* is 1 for villages where farmers could either obtain the technology from door-to-door sales or from peer suppliers. Standard errors are clustered at the village level. Asterisks indicate statistical significance at the 1% ***, 5% **, and 10% * levels.

Figure B.2: Stated motivation for sharing Swarna-Sub1 by suppliers



Notes: Top panel displays distribution of stated reasons why suppliers chose not to sell, exchange or gift seeds. For instance, 45.9% of farmers that did not transfer seeds indicated it was because nobody came to them asking for seeds. Bottom panel displays distribution of how trading partners were chosen by suppliers that chose to exchange with other farmers. For instance, 86.73% of farmers that exchanged indicated that they were sought out by other farmers.

Table B.5: Dyadic regressions of link formation at follow-up

	(1)	(2)	(3)	(4)	(5)
Same sub-caste	0.079***				0.036**
	(0.016)				(0.018)
Same surname		0.136***			0.127***
		(0.015)			(0.017)
Houses within 25 m			0.043***		-0.002
			(0.015)		(0.017)
Plots within 100 m				0.021	0.006
				(0.014)	(0.014)
Village Fixed Effects	Yes	Yes	Yes	Yes	Yes
Mean of Dep Variable	0.380	0.380	0.380	0.384	0.385
Number of Observations	27633	27633	27427	24979	24837
R squared	0.071	0.080	0.066	0.066	0.080

Data are from follow-up social network survey of all farmers. Dependent variable is 1 if either farmer in the dyad indicated a sharing link (i.e. an undirected network). Standard errors are clustered at the village level. Asterisks indicate statistical significance at the 1% ***, 5% **, and 10% * levels.

Table B.6: Heterogeneity of adoption effect according to preferences for quality seeds

	(1)	(2)
Door-to-door treatment	0.352***	0.376***
	(0.051)	(0.050)
Door-to-door treatment*Seed buyer in 2012	-0.036	
	(0.050)	
Seed buyer in 2012	-0.021	
	(0.024)	
Door-to-door treatment*Quality preference		-0.078
		(0.051)
Quality preference		-0.012
		(0.027)
Farmer is SC	-0.063	-0.054
	(0.041)	(0.039)
Farmer has BPL card	-0.055*	-0.057*
	(0.031)	(0.030)
Land cultivated in 2012	0.004	0.005
	(0.007)	(0.007)
Ag. cooperative member	-0.016	-0.007
	(0.024)	(0.023)
Swarna user in 2012	0.101***	0.091***
	(0.032)	(0.033)
Strata Fixed Effects	Yes	Yes
Mean of Dep Variable: Network	0.07	0.07
Number of Observations	1134	1134
R squared	0.206	0.209

Dependent variable is 1 if farmer adopted Swarna-Sub1 for 2013 wet season. *Door-to-door Treatment* is 1 for villages where farmers could either obtain the technology from door-to-door sales or from peer suppliers. Standard errors are clustered at the village level. Asterisks indicate statistical significance at the 1% ***, 5% **, and 10% * levels.

Table B.7: Effect of door-to-door sales on sales and exchanges to farmers outside the sample

	(1)	(2)
Door-to-door treatment	-0.057	-0.047
	(0.075)	(0.073)
Swarna-Sub1 harvest (100 kg)		0.056***
		(0.018)
Farmer is SC		0.268**
		(0.115)
Age of farmer		-0.002
		(0.002)
Farmer has BPL card		0.034
		(0.067)
Education above primary		-0.046
		(0.075)
Strata Fixed Effects	Yes	Yes
Mean of Dep Variable: Network	0.29	0.29
Number of Observations	394	393
R squared	0.024	0.101

Data are from the final survey with suppliers. Dependent variable is the number of farmers from outside the sample that a given supplier sold or exchanged seeds with. *Door-to-door treatment* is 1 for villages where farmers (in the sample) could either obtain the technology from door-to-door sales or from peer suppliers. Standard errors are clustered at the village level. Asterisks indicate statistical significance at the 1% ***, 5% **, and 10% * levels.