Social Learning and Parameter Uncertainty in Irreversible Investments: Evidence from Greenhouse Adoption in Northern China

Honglin Wang, Fan Yu, Thomas Reardon, Jikun Huang, Scott Rozelle[1]

Current Version: December 31, 2009

[1] Wang is from the Hong Kong Monetary Authority, Yu is from Claremont McKenna College, Reardon is from Michigan State University, Huang is from the Center for Chinese Agricultural Policy (CCAP) and Rozelle is from Stanford University. We are grateful for comments from Jeffrey Wooldridge, Robert Myers, John Giles, Zhengfei Guan, Songqin Jin, Xiaoxia Dong, CCAP survey team members, and seminar participants in the Red Cedar Conference at Michigan State University. All errors remain our own.

Social Learning and Parameter Uncertainty in Irreversible Investments: Evidence from Greenhouse Adoption in Northern China

Abstract

The adoption of new technology usually involves irreversible investments where the future payoff is uncertain. In addition, investors often have to contend with a limited understanding of the technology itself, which can be modeled as uncertainty regarding the parameters of the stochastic process describing the future payoff. We hypothesize that social learning increases the probability of adoption by reducing parameter uncertainty, and thus the overall level of risk facing the investor. We test this hypothesis using Chinese farm household data on adoption of greenhouses. Our empirical findings support the hypothesis. We also find that market volatility discourages adoption.

Keywords: Social learning, technology adoption, irreversible investment, parameter uncertainty, greenhouse, China.

JEL Classification: D83, O12, Q16.

1 Introduction

Risk and uncertainty have been important themes in the agricultural technology adoption literature since the 1970s. They were included in studies of green revolution technology adoption to explain lagged or partial adoption or even disadoption. Examples include Roumasset (1976) and Feder (1980). This can be seen as part of a wider strand of literature on the economics of risk and uncertainty, and their constraining effects on investment (Newbery and Stiglitz, 1981).

Distinctions in two dimensions in particular that interest us here have been drawn from the initial foundation of inclusion of risk and uncertainty in agricultural technology adoption analysis. The first dimension is the modeling of various forms of "information capital" as part of the vector of capital assets in the adoption function. The earliest forms modeled were public information in the form of farmers' education and access to extension services. Then, and of most interest to us here, came the introduction of personal experience with a technology ("learning by doing") and observation of neighbors' experience with the technology ("learning from neighbors"). These were introduced for example in Besley and Case (1994) and Foster and Rosenzweig (1995).

The modeling of "learning from neighbors" has been further refined in recent papers that model "social learning," such as: (1) Conley and Udry (2001, 2008) in their modeling of Ghana farmers' adoption of fertilizer in pineapple production, conditioned by their incomplete information and communication networks with neighbors; (2) Bandiera and Rasul (2006) in their modeling of Mozambique farmers' adoption of sunflowers, conditioned by their social network (neighbors and friends who have adopted); and (3) Munshi (2004) in his modeling of Indian farmers' adoption of HYV of rice and wheat, conditioned by their neighbors' experiences but differentiated over rice and wheat areas due to the influence of heterogeneous population. This body of work has demonstrated the effects of social learning on technology adoption. In most cases the social learning's effect on adoption is interpreted as increasing the capacity of the farmer to adopt as well as reducing the farmer's uncertainty and perception of risk in adoption.

The second dimension is the modeling of irreversible investments in capital embodying technology, such as tube wells, greenhouses, and so on. This distinction - between reversible investments such as adoption of an annual crop, a hybrid seed, fertilizer, or a new planting technique - and irreversible investments where the salvage value of the asset is negligible or the asset cannot be transferred or sold, is important in the analysis of risk and uncertainty in technology adoption.

Because of incomplete information with respect to the performance, reliability, and appropriateness of agricultural equipment, irreversibility entails substantial risk for the investor (Dixit and Pindyck, 1994, and Sunding and Zilberman, 2000). McDonald and Siegel (1986) and Dixit and Pindyck (1994) show that the ability to delay an irreversible investment can be considered as a real option; a higher level of uncertainty regarding future benefits raises the option value and causes the investment decision to deviate from the classical NPV rule. Specifically, investors may rationally delay investment to gain additional information, reduce the level of uncertainty, and increase discounted expected payoffs. This has been modeled in two strands of literature.

On the one hand, delayed investment to gain additional information in the face of uncertainty has been studied in the economics literature, inspired by McDonald and Seigel and Dixit and Pindyck. Examples include Olmstead and Rhode (1993), Zilberman et al.

(2004), Hassett and Metcalf (1995), and Nelson and Amegbeto (1998), inter alia. These studies have tended to assume that all parameters of the dynamic process are known to agents, and the only uncertainty in the model comes from the future value of the dynamic process.

On the other hand, investment under parameter uncertainty has been examined in the finance literature. Merton (1980) shows that while the variance of the return can be estimated precisely from continuous observations on a finite interval, the estimator of mean return does not converge unless the length of the interval becomes large. Gennotte (1986) studies portfolio choice under incomplete information about the stock return process. He uses tools of nonlinear filtering from Lipster and Shiryaev (1978) to derive the optimal drift estimator as agents continuously observe the returns. Brennan (1998) and Xia (2001) construct similar models to examine how learning about unknown parameters and unknown predictability affects portfolio choice. More recently, Abasov (2005) modeled irreversible investment under parameter uncertainty, and Huang and Liu (2007) modeled learning from discrete noisy signals about the true drift in their study of periodic news on portfolio selection. Note that much of the finance literature is primarily theoretical, with few empirical applications and none in the domain of investment in agricultural capital as an embodiment of agricultural technology adoption.

The present paper aims at a particular, and a particularly important, gap left by the two dimensions discussed. That is, while the literature on social learning and technology adoption has modeled the effect of social learning as a means of reducing uncertainty, that literature has not treated the issue of irreversibility of the investment per se, and thus has not modeled the effect of social learning in a real options context. Moreover, while the literature on irreversible investment and uncertainty has indeed modeled investment in a real options framework, it has not examined uncertainty-reduction measures taken by adopters, in particular, social learning.

There is thus a gap in the literature, both theoretical and empirical, where an analysis of irreversible investment under parameter uncertainty models the effect of social learning. The contribution of the present paper is to address that gap.

We address the gap empirically by modeling greenhouse investments with primary data from Shandong province in China. The data are multi-year, observing the characteristics, including their social network of prior adopters, of the adopters the year before their adoption, and thus, new to this literature, we capture causality of social learning and adoption.

We address it theoretically, by presenting a new model to the literature of these links. Following McDonald and Siegel (1986), we assume that a farmer is considering an investment project, whose value follows a geometric Brownian motion. Departing from the standard framework, we assume that the true drift of the Brownian motion is unobservable to the farmer (we call this parameter uncertainty). In essence, the farmer is imperfectly informed as to the expected rate of return of his investment.[2] He must make an inference about the true expected return based on his information and, at the same time, determine the optimal timing for investing in the project.

The farmer can learn about the unknown parameter in two ways. First, he extracts information on the true drift from a continuous observation of past realized returns on the project value. This captures the process of continuous learning from public information about the project. Second, he obtains discrete noisy signals of the true drift. This represents

[2] We use the male pronoun to refer to farmers because the large majority of farmers in our sample are men.

the process of social learning from early adopters in his social network, who might possess information about the project that the public do not have. In our model, parameter uncertainty adds to the overall risk that the farmer faces; this raises the threshold project value needed to induce the farmer to invest. In contrast, social learning reduces parameter uncertainty, which decreases the overall level of uncertainty and reduces the investment threshold, thereby increasing the likelihood of adoption. In our model, social learning also causes the farmer's belief about the expected return to converge to the average belief of his social network; the higher the average belief, the higher is the investment threshold, and the less likely the farmer will adopt the technology.

The rest of the paper is organized as follows: In Section 2, we present the theoretical model. In Section 3, we provide background information about the greenhouse technology in northern China. In Section 4, we outline our sample selection and summarize the data. In Section 5, we explain our empirical methodology. In Section 6, we present the empirical findings using linear probability models. We conclude in Section 7.

2 The Model

In this section, we use a real options model to articulate the effect of parameter uncertainty and social learning on technology adoption. We begin with a model of continuous learning, which is essentially that of Abasov (2005). Specifically, a farmer is considering whether to pay a sunk cost of I for an agricultural technology, whose value V evolves according to:

$$dV_t = V_t\left(\mu dt + \sigma dZ_t\right),$$

where Z is a Brownian motion.

Motivated by Merton (1980), we assume that the farmer can observe V continuously and knows its volatility σ; however, he only knows that the drift μ is a normal random variable with mean m_0 and variance γ_0 in the beginning.[3] According to Lipster and Shiryaev (1978), the conditional mean of the drift given the farmer's information set, $m_t = E\left(\mu \mid \mathbf{F}_t^V\right)$, follows:

$$dm_t = \frac{\gamma_t}{\sigma} dZ_t',$$

where $\gamma_t = E\left[\left(\mu - m_t\right)^2 \mid \mathbf{F}_t^V\right]$ is the conditional variance of the drift, satisfying:

$$d\gamma_t = -\frac{\gamma_t^2}{\sigma^2} dt, \tag{2.1}$$

and Z' is a new Brownian motion related to the original Brownian motion through:

$$dZ_t' = dZ_t + \frac{m_t - \mu}{\sigma} dt.$$

We can solve equation (2.1) for γ_t:

$$\gamma_t = \frac{\gamma_0 \sigma^2}{\gamma_0 t + \sigma^2}.$$

This result shows that continuous learning decreases the conditional variance of the unknown parameter. Thus the longer the farmer observes the value process, the less uncertain he is about the drift. This is consistent with Merton (1980)'s results: the uncertainty of the drift is not related to the number of observations, but is rather related to the length of the observation period. However, the conditional mean of the drift can fluctuate up or down, depending on new observations of the Brownian motion Z'.

According to Gennotte (1986), the farmer's decision can be separated into two problems: the inference of the unknown parameter given $\left\{Z_s'\right\}_{0 \le s \le t}$, and the optimal

[3] The drift in this context can be interpreted as the productivity of the greenhouse technology.

stopping decision based on the current state variables (m_t, γ_t, V_t) and the dynamics of (m, γ, V). Putting everything together, we can characterize the farmer's problem using observable processes:

$$\begin{aligned} J(m_0, \gamma_0, V_0) &= \max_{\tau \in \mathbf{F}^V} E\left[e^{-\rho\tau}(V_\tau - I)\right], \\ s.t.\ \ dV_t &= V_t\left(m_t dt + \sigma dZ_t'\right), \\ dm_t &= \frac{\gamma_t}{\sigma} dZ_t', \\ d\gamma_t &= -\frac{\gamma_t^2}{\sigma^2} dt. \end{aligned} \tag{2.2}$$

Here, ρ is the farmer's discount rate, and τ has to be an $\mathbf{F}^V$-stopping time, reflecting that the farmer must make a decision based on his information set. The stopping rule takes the form of:

$$\tau = \inf\left\{t \geq 0 : V_t \geq V^*(m_t, \gamma_t)\right\},$$

where $V^*(m, \gamma)$ is the trigger value of investing, which depends on the state variables.[4]

Abasov (2005) derives the Hamilton-Jacobi-Bellman equation for the optimal stopping problem (2.2) and transforms it into a linear complementarity problem, which he solves with the finite difference method. His numerical results demonstrate that the trigger value of investing, $V^*(m_0, \gamma_0)$, obtained as a part of the solution, increases with γ_0. This result is sensible given that the trigger value in the McDonald and Siegel (1986) and Dixit and Pindyck (1994) model increases with σ, and parameter uncertainty contributes to the total uncertainty in our model. In addition, Abasov shows that V^* increases with m_0; this is also consistent with the traditional real options model without parameter uncertainty.

[4] Since γ is a deterministic function of *t*, we can equivalently formulate the problem in terms of state variables (m, t).

In developed countries, there are public economic forecasts and newsletters informing investors. Therefore, agents can make inferences based on past realized returns. However, in rural China, information is more likely to come from local private sources. Similar to Huang and Liu (2007), we allow farmers to obtain direct signals of the drift from early adopters in their social networks. These signals are noisy, reflecting the fact that even early adopters are unlikely to learn everything about the technology from their own experience. Different from Huang and Liu (2007), we assume that the signals are costless. However, the number of signals to which a farmer has access is limited by the scope of his social network, which we take as exogenous. For simplicity, we also assume that these signals are received at time 0, just as the farmer begins to consider his adoption decision. Since discrete signals are much more effective than continuous learning in changing the farmer's belief, it seems reasonable to assume that he would seek out these signals at the very beginning of his decision-making process. This implies that discrete updating affects the farmer's optimal stopping problem only insofar as it changes his initial belief; discrete updating plays no role in the dynamics of the conditional mean and conditional volatility.

Let signal i be given by:

$$\mu_i = \mu + \varepsilon_i, \tag{2.3}$$

where $\varepsilon_i \sim N\left(0, \sigma_\varepsilon^2\right)$ is independently and identically distributed. After receiving n such signals, it can be shown that the conditional mean and variance of the drift are given by:

$$m_0' = \frac{\sigma_\varepsilon^2}{n\gamma_0 + \sigma_\varepsilon^2} m_0 + \frac{n\gamma_0}{n\gamma_0 + \sigma_\varepsilon^2} \overline{\mu}, \tag{2.4}$$

$$\gamma_0' = \frac{\gamma_0 \sigma_\varepsilon^2}{n\gamma_0 + \sigma_\varepsilon^2}, \tag{2.5}$$

where $\overline{\mu} = \frac{1}{n}\sum_{i=1}^{n}\mu_i$. Equation (2.5) shows that the conditional variance is decreasing in the number of signals, which can be taken as the scope of social learning. Therefore, social learning reduces parameter uncertainty. Using Abasov's numerical results, this implies that social learning decreases the trigger value for adoption, making it more likely that the farmer would adopt the technology.

Considering the conditional mean equation (2.5), we find that as the number of signals increases, $m_0^{'}$ tends to move away from m_0 and approach $\overline{\mu}$. This indicates that social learning causes the farmer's belief about the drift to converge to the average belief in the farmer's social network. The net effect depends on the relation between m_0 and $\overline{\mu}$. If $m_0 > \overline{\mu}$, the farmer is initially too optimistic; social learning causes him to lower his expectation about the project's return. This, in turn, lowers the trigger value and facilitates adoption. If the farmer is, on average, unbiased in his initial belief, then social learning is unlikely to change the probability of adoption through its effect on the conditional mean return.

If we generalize this model to allow the dynamics of social learning to enter the farmer's decision making, then we can write down the following optimal stopping problem, where we combine continuous filtering with discrete updating:

$$
\begin{aligned}
J\left(m_0,\gamma_0,V_0\right) &= \max_{\tau\in\mathbf{F}^V\vee\mathbf{F}^N} E\left[e^{-\rho\tau}\left(V_\tau - I\right)\right], \\
s.t.\ \ dV_t &= V_t\left(m_t dt + \sigma dZ_t^{'}\right), \\
dm_t &= \frac{\gamma_t}{\sigma}dZ_t^{'} + \frac{\gamma_{t-}}{\gamma_{t-}+\sigma_\varepsilon^2}\left(\mu\left(t\right) - m_{t-}\right)dN_t, \\
d\gamma_t &= -\frac{\gamma_t^2}{\sigma^2}dt - \frac{\gamma_{t-}^2}{\gamma_{t-}+\sigma_\varepsilon^2}dN_t.
\end{aligned}
\tag{2.6}
$$

Here, $\mu\left(t\right)$ refers to the independently and identically distributed noisy signals described in equation (2.3), and N_t is a counting process that counts the number of signals that the farmer has received up to time t. It can be periodic and deterministic as in Huang and Liu (2007), or stochastic, as in the case of a Poisson process with arrival rate λ, which describes social interaction as a random phenomenon. In all cases, however, the first part of the dynamic equations for (m,γ) captures the effect of continuous updating as the farmer learns from the past history of V. The second part represents a jump in the conditional mean and variance when the farmer receives a noisy signal of the drift. Because γ and N are deterministically related through the conditional variance relation, we have suppressed the dependence of the value function on N. Similarly, we can write the trigger value as $V^*(m_t,\gamma_t)$, with the understanding that the effect of N_t is already reflected in the conditional variance γ_t.

Generally, the optimal stopping problem (2.6) must be solved numerically. The adoption decision is related to the amount of social learning that the farmer has experienced. According to the above model, this is measured by N_t. As the conditional variance equation shows, a larger N (more social learning) always reduces γ. We conjecture that the trigger value is increasing in γ, regardless of whether farmers are cognizant or ignorant of future social learning.[5] This implies that social learning can lower the trigger level for adoption.

Summarizing the various models, the classical real options analysis of McDonald and Siegel (1986) predicts that the trigger value for investment increases with the uncertainty of the project value. We show that this result also extends to parameter

[5] One can conceive of cases in which knowledge of the social learning dynamics can actually delay adoption. For example, if the farmer knows that parameter uncertainty will be fully resolved tomorrow, he is unlikely to invest today.

uncertainty. Building from recent work on social learning and technology diffusion (such as Bandiera and Rasul, 2006), we argue that social learning can facilitate adoption by reducing parameter uncertainty. In rural China, where public extension information is not easily accessible to small farmers, information from social learning could play an important role in their adoption decisions. The rest of our paper is dedicated to testing this hypothesis.

3 Greenhouse Intermediate-Technology in Northern China

Before economic reforms, China gave first priority to the development of heavy industry. In agriculture, China emphasized the importance of self-sufficiency for grains - the "iron rice bowl policy." After the "household responsibility system" reform started in 1981, the shortage of grain supply was relieved by a significant increase in grain production. This made it possible for China to diversify into horticulture and livestock husbandry. Meanwhile, rapid income growth in the 1980s and 1990s created an increasing demand for high-value horticultural products.

The huge demand for cheap fresh vegetables led to the development and widespread diffusion of an affordable greenhouse technology for northern Chinese farmers. Rather than the modern, expensive type made of steel frame, plastic or glass walls and ceilings, and requiring energy-using heating and cooling mechanisms (promoted in the 1970s in China but saw very little adoption because of the cost, Wan, 2000), the greenhouse adopted in the 1990s in northern China is of the "intermediate technology" type, which was first created by Shandong farmers in the early 1980s. This type of greenhouse is made of simple clay walls, bamboo frame, a plastic-sheet roof, and a straw mat roll-out awning for cold nights. The sun warms the interior, with the greenhouse built with an orientation to maximize sunlight capture. These greenhouses changed not only the food consumption pattern for hundreds of millions of consumers, but also the face of farming in northern China. They helped to transform China from a modest global player to the volume leader in horticulture - China grew 47 percent of the vegetable volume in the world by 2004 (Weinberger and Lumpkin, 2005). The vegetable greenhouse area in China reached 150,000 hectares in 2004 (Chinese Agriculture Yearbook, 2006), and at least half a million farmers were by that year using the intermediate-technology greenhouse.

The intermediate-technology greenhouse is far cheaper than a modern type, but is still a major investment for the small farmers of Shandong. The construction cost of intermediate-technology greenhouses is roughly four dollars per square meter, much cheaper than modern greenhouses made of glass or plastic, which cost about 70 dollars per square meter to construct. Yet, even four dollars per square meter is a large investment for small farmers. For example, if a greenhouse is 60 meters long and 10 meters wide, the construction cost would be about $2,400, while the average Chinese farmer earned less than $500 in 2005. Moreover, the labor involved in building the greenhouse is substantial; the farmer often spends months creating the main component - the rear-wall of the greenhouse, which is usually made of pounded clay bricks.

In addition, the investment is “irreversible,” in the sense of Bertola and Caballero (1994), as the structure can only be used in immediate production, and has little to no salvage value and cannot be sold or transferred (the structure is not movable). If the farmer decides to demolish the greenhouse, the bricks would most likely be broken into dirt clods, and the old straw awning and old bamboo beams are worth little in salvage.

We must point out that greenhouse vegetable growing is different from open-field vegetable growing along several dimensions. First, greenhouse yields far exceed open-field yields. For example, the tomato yield is about 12 tons/mu/year in a greenhouse, compared to about 2 tons in an open field. Several factors, including a longer growing season, multiple harvests, labor intensive production, and more advanced technology, contribute to the higher yield. Second, the greenhouse growing season lasts for about 9 months between the fall and the spring, while open-field growing lasts for about 3 months during the summer. Third, greenhouse-grown vegetables are often transported to distant markets, while vegetables from the open-field are usually used for self-consumption, with a small surplus sold in local markets. These considerations suggest that there is little overlap or competition between greenhouse and open-field growing. In fact, one could argue that farmers' adoption decision is not about greenhouse vs. open-field growing; rather, it is mainly about making a decent living with their land (greenhouse vegetable growing) or by pursuing off-farm jobs.

4 Data

4.1 Sample Selection

Our survey area is in Shandong province, the leading horticulture province in China. It has seven percent of China's cropland, but 12 percent of China's horticultural land in 2004. The latter share has been steadily rising over time. The number of greenhouses and the level of commercialization as well as yields in Shandong are higher than in the rest of China.

In Shandong, we conducted two coordinated community and household level surveys in 2005 and 2006, respectively. The first one, the Shandong village survey, provided a representative sample of tomato and cucumber growing villages in Shandong.[6] During the first step of the survey, we created sampling frames of county-level tomato and cucumber production in order to select five sample counties per crop. Specifically, with knowledge of county production of each crop, we ranked counties by the output per capita of that crop. For each crop in our sample, one high production county was randomly selected from the counties in the top quintile; the other high production county was randomly selected from the second quintile. The two medium production counties were randomly chosen from the third and fourth quintiles, respectively. After eliminating five percent of the counties with the lowest production, the low production county was randomly chosen from the lowest quintile. In the end, there were two counties in the high production set, two counties in the medium production set, and one county in the low production set.

After the sample counties were chosen, a similar process was used to select sample townships from the counties and sample villages from the townships. For each crop, the survey teams visited a total of ten townships. Moreover, for each crop (among the ten townships), we interviewed respondents in 35 villages (22 in high production counties, 10 in medium production counties, and 3 in low production counties). Since we collected area data on all villages, townships, and counties in the sample, we were able to construct area-based weights in order to create point estimates of our variables that are provincially representative.

[6] The reason why we did not directly stratify on greenhouse use is that our survey is part of a large horticulture production survey, which required stratified sampling of cucumber/tomato and non-cucumber/tomato households. Cucumber and tomato are the two most popular greenhouse crops in the sampling area, and we are able to adjust for the selection bias with knowledge of the distribution of crops and greenhouse use in each village.

Having selected the villages, the enumeration team visited each community and undertook data collection. Specifically, the enumerator conducted a two-hour interview with three village leaders for the village survey. In each village, we divided all households into two groups. For the cucumber sample, they are non-cucumber households and cucumber households. We randomly sample seven cucumber farmers and three non-cucumber farmers. As a result, we obtained 350 households from cucumber growing villages. With knowledge of the distribution of cucumber farmers and non-cucumber farmers, plus the distribution of greenhouse adopters in each village, we calculated the weights to adjust for selection bias. Following this procedure, we also obtained 350 households from tomato growing villages.

After data cleaning, we collected 638 valid household observations. Among this sample, 204 (64 percent) out of 317 households from tomato growing villages were found to have adopted greenhouses, while 158 (49 percent) out of 321 households from cucumber growing villages were found to have adopted greenhouses. That a higher share of tomato growers adopted greenhouses is apparently due to the fact that in cucumber production, a shading shed is a substitute for a greenhouse, while in tomato production there is no substitute for a greenhouse, and the options are only growing in the open field or in a greenhouse.

Shandong farmers did not adopt greenhouses all at once, but rather, in a process typical of diffusion of new technology, over years. The greenhouse diffusion process can be roughly divided into three stages: early stage, take-off stage, and slow-down stage. Figure 1 shows that the diffusion process is relatively slow in the early stage before 1990; only a few farmers adopted the technology. Between 1990 and 1995, many more farmers adopted. The diffusion process reached its peak between 1996 and 2000, after which the trend began to slow down. This diffusion curve is similar to the "S-curve" observed by Griliches (1957) for the adoption of hybrid maize in the US, and subsequently documented in many other settings.

4.2 Social Learning

We are interested in the effect of social learning on farmers' adoption of greenhouses. Our theoretical model predicts that social learning helps to reduce uncertainty, thus facilitating adoption. Empirically, however, social learning could be one of many factors affecting adoption. For example, farmers may have other options such as off-farm jobs. Alternatively, farmers may be credit-constrained because greenhouse adoption is a major investment. To disentangle the effect of social learning from other determinants, we need to find appropriate empirical proxies for social learning and control for other factors that might influence farmers' decisions.

Social learning is a key variable in our study. We measure social learning in a way similar to the approach of Bandiera and Rasul (2006). Specifically, we asked the farmers who adopted, "How many people do you know who adopted greenhouses *before you adopted* in your village?" We asked the non-adopters how many adopters they knew at the time of the survey. We control for year with year dummy variables. We then asked, "How many of these people are your relatives and friends?" (We did not include neighbors as a separate category because Chinese farmers usually consider neighbors among friends.) The answer to the second question is taken as our empirical proxy for social learning. Differing from Bandiera and Rasul (2006) (who asked about the social network at the time of the survey, not before adoption), we obtained the size of the farmer's social network of adopters *before* his adoption, so that we can infer causality.

There are several reasons why our measure of social network of adopters is an appropriate measure of social learning before adoption. First, the number of earlier adopters among relatives and friends is likely to be positively correlated with the number of different sources of information on greenhouse adoption that the farmer accessed before adoption, which corresponds to the number of discrete signals in our theoretical model. Second, village membership, kinship, and friends are the defining elements of a farmer's social network, or a group of people with whom the farmer has close contact, and from whom information can be most easily obtained. By concentrating on the number of earlier adopters among relatives and friends, we also mitigate the concern for *ex post* social network formation. While this is obvious for kin adopters, we noticed during our survey that Shandong farmers tended to define friendship based on long-term relation, such as classmates, neighbors, and people who served with them in the army. Typically, they consider a friend someone from whom they can borrow money in case of illness; they would not consider passing acquaintances as friends. Third, we found that farmers were easily able to remember the number of adopters they knew before they adopted; we surmise that this is because a greenhouse is a big investment for local farmers and hence easily observable.

The first two rows of Table 1 provide the means and standard errors of our social learning measures by adoption status. In the last column, tests of equality of the means are provided to examine whether the differences between adopters and non-adopters are significant. The first row indicates that, on average, adopters know about 6.9 earlier adopters among relatives and friends in their own village, while non-adopters only know about 4.7 earlier adopters in their social network. The result of the *t*-test shows that this difference is significant. This implies that there is more social learning for adopters than for non-adopters. When we extend the scope of the social network to include earlier adopters among relatives and friends in nearby villages (the second row), the findings are similar.

4.3 Other Household Characteristics

Table 1 presents other household characteristics by adoption status. There are several salient features.

(1) Demographics differ between adopters and non-adopters. The family size of adopters is significant larger than that of non-adopters, while the amount of farm labor is significant smaller for adopters than for non-adopters. This is because adopters have more dependent family members (either young children or old parents) than non-adopters. For such households, greenhouse adoption could be a good choice because it allows the adults to work close to home, so that they can care for dependent family members. Non-adopters are, on average, substantially older than adopters - a point consistent with younger farmers having more young children and old parents to care for.

(2) Off-farm employment and income are significantly larger for non-adopters than for adopters, which suggests that greenhouse labor and off-farm jobs are substitutes.

(3) There is no significant difference in education between adopters and non-adopters in our sample. This suggests that education is not the main determinant of greenhouse adoption when the main source of information for the technology is social learning.

(4) The farm size of adopters is larger than that of non-adopters, which indicates that farmers with more land are more dependent on agricultural income, and farmers with less land are more likely to favor off-farm jobs.

(5) Irrigation is of course important to greenhouse farming, and 89 percent of the adopters have access to irrigation. However, 80 percent of the non-adopters also have access to irrigation, showing that there is not much variation in irrigation access among farmers in this well-irrigated region.

(6) Adopters have greater land tenure security than non-adopters. This is a sensible result given the long-term nature of greenhouse investment. We proxy land tenure security by the number of land reallocations undertaken by village leaders every few years to ensure relative land distribution equality in the village.

(7) Adopters and non-adopters have no significant difference in grain land share, which suggests that both groups have a similar agricultural production pattern, except that adopters use greenhouses to produce vegetables for income and non-adopters produce vegetables in the open field mainly for self-consumption.

(8) The presence of a credit constraint would in theory undermine an important investment such as greenhouses, all else equal. However, it is difficult to measure the credit constraint of a farmer, as this is equivalent to examining whether he can borrow as much as he would like at the going market interest rate (Banerjee and Duflo, 2002). Since we are focusing on greenhouse adoption rather than testing whether the farmer has invested in a greenhouse of optimal scale, we only need to know whether he is capable of building a greenhouse by borrowing money or using his savings. Therefore, we identify the housing construction cost prior to greenhouse adoption as a proxy for

the financial capacity of a household. We also collect the greenhouse construction cost for each household, and use the ratio of the two construction costs as an indicator of potential credit constraints. This ratio indicates that the housing construction cost is greater than the greenhouse construction cost for both adopters and non-adopters. More importantly, the ratio for non-adopters is 3.24, which is significantly greater than the ratio for adopters, which is 1.95. This suggests that non-adopters are less likely to face credit constraints than adopters.

(9) Although adopters tend to experience more social learning, this is not because they have larger social networks. Since it is not easy to ask farmers to identify their total number of friends, we use the number of people attending the latest wedding of a farmer's family as a proxy of his social network.[7] The data indicates that, on average, non-adopters have more friends (113.07) than adopters (98.45). The difference is not statistically significant, implying that adopters do not have larger social networks than non-adopters.

5 Empirical Methodology

In this section, we illustrate the connection between our theoretical model and the empirical framework. According to our real option model of greenhouse adoption, the farmer decides to adopt or to wait based on a comparison between the current value of the technology and the trigger value. Therefore, we can define the farmer's adoption status at

[7] The number of people attending a wedding in rural China can be easily identified through the number of banquet tables. The number of banquet tables is important to a household, and each table usually seats the same number of guests, so that the total number of guests can be calculated by multiplying the number of tables and the number of guests per table.

time t as:

$$Y_t = 1 \text{ (adopt), if } Y_t^* = V_t - V_t^* > 0,$$
$$Y_t = 0 \text{ (non-adopt), if } Y_t^* = V_t - V_t^* \le 0, \tag{5.1}$$

where V_t is the discounted expected value of all future cash flow from greenhouse vegetable production, and V_t^* is the trigger value.

McDonald and Siegel (1986)'s model, in which the drift μ is known, shows the trigger value V^* as a function of the parameters (ρ, μ, I, σ). However, the drift μ is unknown in our model. Thus, the trigger value also depends on the conditional mean and variance of the drift, (m_t, γ_t). According to the dynamics of (m, γ) in equation (2.6), we can substitute (m_t, γ_t) with functions of $\left(m_0, \gamma_0, Z_t^{'}, N_t, \sigma, \sigma_\varepsilon, \overline{\mu}, t\right)$.[8] Therefore, we can express the trigger value V_t^* as:

$$V_t^* = g\left(\rho, I, \sigma, m_0, \gamma_0, Z_t^{'}, N_t, \sigma_\varepsilon, \overline{\mu}, t\right). \tag{5.2}$$

Following similar reasoning, the current project value V_t can be written as a function of the same group of variables. Therefore, we can express $Y_t^* = V_t - V_t^*$ as:

$$Y_t^* = h\left(\rho, I, \sigma, m_0, \gamma_0, Z_t^{'}, N_t, \sigma_\varepsilon, \overline{\mu}, t\right). \tag{5.3}$$

To motivate the empirical proxies for the variables in equation (5.3), we first note that $Z_t^{'}$ represents the stochastic change in the project value. A good proxy for $Z_t^{'}$ is the observed profitability of greenhouse production in the current period. We proxy that profitability by the ratio of the output price to the input price. Because historical data are

[8] This is only a simplified representation; strictly speaking, the solution of (m_t, γ_t) according to equation (2.6) depends on the paths of $Z^{'}$ and N, as well as the history of the signals up to time t.

not available on vegetable prices in Shandong, we use the ratio of the vegetable price index and the input price index at the national level as a proxy for the profitability of greenhouse production over the years. For the investment cost I, we use the greenhouse construction cost (real value) for each adopter. For non-adopters, we use the average construction cost for adopters in their village or nearby villages as the proxy.

Continuing with the interpretation of equation (5.3), σ is the volatility of the project value, which we measure as the standard deviation of the national vegetable price index over the three years prior to the farmer's adoption. $\bar{\mu}$ represents the average signal received by the farmer from his social network, the proxy for which is the vegetable price index growth rate over the three years preceding the farmer's adoption. This is a reasonable assumption if the expected return of the project is close to the average return in the economy. The time *t* in our model is equated with the amount of time the agent spent in continuous learning. We use the number of years that the farmer had been aware of the technology before adoption to represent the continuous learning effect. As noted above, N_t is the key variable in our study. We measure it by the number of earlier adopters in a farmer's social network, which includes relatives and friends in his own village and nearby villages.

Besides these theoretically motivated variables, there may be other factors that affect greenhouse adoption in practice, such as land tenure security, off-farm employment, and household characteristics. These factors were discussed in the preceding section. In addition, we do not have compelling empirical proxies for farmers' discount factor ρ, their initial values of the conditional mean and variance (m_0, γ_0) before any learning had taken place, and the standard deviation of their signals σ_ε. These parameters, however, are likely correlated with household characteristics such as age, family size, and education, which we include in our empirical analysis to capture potentially omitted factors.

Our theoretical model is based on observables; with knowledge of these observables, the model predicts adoption with certainty. In reality, however, we do not observe all information relevant for determining adoption. Therefore, our empirical model must allow for the presence of unobserved determinants.

In brief, our empirical model can be written as:

$$Y_i^* = f\left(X_i, Z_i, N_i, D_1, D_2\right) + e_i, \tag{5.4}$$

where i denotes a household, Y_i^* is the adoption criterion in year t according to equation (5.1), and X_i are household characteristics before adoption (year $t-1$), which include the age and education of household head, family size, farm size, off-farm income, irrigation conditions, ratio of housing and greenhouse construction costs, grain share, total number of friends, and years of awareness of the technology. Z_i are institutional and market variables at $t-1$, which include the number of land reallocations, the ratio of the output price index to the input price index, the volatility of the vegetable price index, and the average growth rate of the vegetable price index. N_i is the number of earlier adopters in the farmer's social network at $t-1$. D_1 and D_2 are, respectively, year and county dummies that control for heterogeneity in farmers' adoption across different years and counties. Finally, e_i represents the effect of unobservable determinants of adoption. According to equation (5.4), the probability of adoption is:

$$P\left(Y_i^* > 0\right) = P\left(e_i > -f\left(X_i, Z_i, N_i, D_1, D_2\right)\right). \tag{5.5}$$

In our empirical analysis, we estimate a linear probability model (LPM), which specifies the above probability as a linear function of the explanatory variables. The LPM allows us to use year dummies to control for heterogeneities over time, which is important given the structure of our dataset (with different farmers adopting greenhouse growing in different years).

6 Empirical Results

6.1 Identification Strategy

In this section we focus on the potential endogeneity of the social learning effect and our identification strategy. The endogeneity problem is one of the most formidable problems in empirical studies. In order to find an appropriate identification strategy for this study, it is crucial to understand the reasons why we could face the problem.

Manski (1993) uses the reflection problem to describe the tendency for people in the same social network to behave in similar ways. He identifies two possibilities: (1) an endogenous effect, wherein the propensity of an individual to behave in certain ways varies with the prevalence of the behavior in the group; (2) a correlated effect, wherein common environment and personal characteristics produce similar behavior.

In this paper, we attempt to show that farmers' adoption decision is influenced by social learning. Therefore, we need to empirically distinguish the social learning effect from the endogenous effect and the correlated effect.

In our context, the endogenous effect is essentially the social pressure problem. Psychologists often use social pressure as a way of explaining herd behavior. In Shandong, most farmers are free to make production plans for their own farms after the economic reform. Commonly, farmers in a village have multiple ways of making a living (e.g., farming, off-farm jobs in cities, small businesses in the village, etc.). It seems unlikely that farmers would adopt greenhouse vegetable growing because of social pressure.

In our context, the correlated effect poses a more serious challenge. An endogeneity problem could arise from the simultaneous determination of adoption and network formation: for example, a farmer could know more adopters because he adopted the greenhouse. In other words, the adoption could affect social learning instead of social learning affecting adoption (endogeneity from simultaneous determination). To mitigate this problem, we collected household and institutional information for the year before the adoption for adopters. For non-adopters, we collected the information in the year before the survey occurred (2005).

Moreover, farmers who are entrepreneurial in spirit are likely to know more people (hence more adopters). At the same time, they are more likely to try out new things (thus more or less likely to adopt greenhouse, depending on the availability of other options). Therefore, a farmer's adoption could be explained by his personality, rather than by learning from others in his social network. Thus, a key problem is how to identify social learning from unobservable error terms such as similar personalities in the social network. We need to find at least one instrumental variable which is (1) correlated with social learning after we control for other factors, but that is (2) not correlated with the error terms. While we can test the first condition, we cannot test the second condition directly because the error terms are not observable.

Fortunately, we have an appropriate instrument in this study: the walking time from the farm to a farmer's neighborhood. More specifically, we ask farmers the following question in the field survey: "How many minutes does it take to walk by your 20 closest neighbors?" The logic of this question is that social learning could be negatively correlated to the walking time. For example, if a farmer lives in a mountainous area, it could take two hours or even more to walk by his 20 closest neighbors. On the contrary, it only takes 10 minutes for farmers to walk by his 20 closest neighbors if people live closely. We surmise that farmers in the second case are more likely to have access to social learning. We test this hypothesis with data after controlling for other factors: we find that walking time is significantly negatively correlated with social learning (first row of Table 4 for both social learning measures). This result demonstrates that the walking time variable satisfies the first condition for a valid instrument.

For an analysis of whether this instrument meets the second condition (lack of correlation with the error term in the adoption equation), the following discussions provide further justification for the validity of the instrument.

First, we use a heuristic explanation to justify the instrument. In rural China, it is not unusual for a family to live in the same place for decades. A well-functioning real estate market does not exist in rural China for several reasons: (1) a farmer could own his house, but not the land on which his house is built because all land is owned by the village collective; (2) it is illegal to buy a house in a village if the buyer is not a member of the village; (3) it is also illegal for a household to buy an additional house from another villager because Chinese law forbids any household to occupy two pieces of land for housing in a village; (4) if a farmer wants to change his house location, either he has to obtain a new piece of land from the village collective under very strict conditions due to land scarcity in Shandong, or he can find another household in the village that is willing to give up its housing land, which is very rare. In addition, in both cases the farmer has to give up his old housing land. Based on these observations, it appears very difficult, if not impossible, for a household to change its location. In other words, the farmer's housing location in rural China can be considered as fixed in most cases.

Second, we construct interaction terms between the IV (distance to neighborhood) and year dummies. We use the Hansen-J over-identification test to examine the validity of the IV given that we believe the other instruments (the interaction terms) to be truly exogenous. The C-statistic from the Hansen-J test (the last row of Table 6) indicates that the distance to neighborhood variable passes the validity test in both social learning measurements. We must be cautious by not over-emphasizing this result, as the power of the Hansen-J test depends on the exogeneity of the other instruments. However, this is the best test we can do to check the validity of an instrumental variable.

Even though the location of a household appears to be fixed, historically, people who live far away from their neighborhood might also live far away from roads or markets. This suggests that they might have fewer opportunities to learn from markets or from the outside world – an important consideration when analyzing the technology adoption decision. To address this concern, we present the correlations between distance to neighborhood and various household characteristics such as distance to roads, distance to markets, education, age, family size, and household wealth proxied by house value. Panel A of Table 2 shows that there is little correlation between distance to neighborhood and the included household characteristics. In Panel B, we group the sample households according to whether their distance to neighborhood is above or below the sample median, and then present the average household characteristics for each group. The results suggest that people who live further away from their neighborhood tend to be richer, and that they are in fact closer to markets. The associated correlations, however, are quite small (see Panel A).

As a result of these discussions, we are fairly confident that the instrumental variable (distance to neighborhood) is exogenous to greenhouse adoption, and therefore it allows us to obtain consistent estimators given that social learning is shown to be endogenous by the Durbin-Wu-Hausman Test (see Table 3).

6.2 Linear Probability Model

Table 3 presents the estimation results for the linear probability model estimated by 2SLS with cluster-robust standard errors using distance to neighborhood as the instrument. The first two columns report the results using a measure of social learning within the farmer's own village; the next two columns report the results using a measure of social learning that also includes the farmer's nearby villages. Generally speaking, the two sets of results are very similar, suggesting that village boundaries are not crucial to how social learning affects greenhouse adoption.

We will focus on the first two columns for a detailed discussion of our results. The first row confirms the key result for our study: social learning has a significantly positive impact on greenhouse adoption. Specifically, one more adopter in a farmer's social network increases the probability of his adoption by about 2 percent after controlling for other factors. In other words, if there are currently 10 earlier adopters in the farmer's social network, his adoption probability in the next year will increase by about 20 percent. Given that the greenhouse adoption rate is still low in rural China, this amount of increasing probability is economically significant.

The third row shows how adoption is affected by the conditional mean return to the greenhouse technology. From our theoretical model, we know that the farmer's belief about the mean return will converge to the average belief of his social network as a result of social learning. Because we cannot observe farmers' expectations, we use the vegetable price index (national level) growth rate before adoption to approximate the average belief of project return in the social network. The coefficient is not significant; however, the sign is consistent with the prediction of our theoretical model, namely, higher expected return results in a higher trigger value for investment and a lower probability for adoption. It is also possible that the price index growth rate is acting as a proxy for farmers' outside opportunities; however, we have already included off-farm income in our regression specification.

We use the market volatility of vegetable prices before adoption to represent the uncertainty in the stochastic project value in our theoretical model. Our result indicates that this source of uncertainty discourages adoption. This finding is consistent with theory, which predicts that the option value of waiting to invest is larger when the future investment value is more uncertain.

We use the number of years that the farmer had been aware of the technology before adoption to represent the continuous learning effect. However, it is not significant according to our estimation. It could be that farmers in rural China simply did not have continuous access to information about the greenhouse technology and its returns. It is also possible that the main source of information about the greenhouse technology is discrete social learning.

Our proxy for the current profitability of the greenhouse technology is the ratio of the output price index to the input price index: the higher is the stochastic project value, the higher is the probability of adoption. Our result confirms this prediction.

We also include the total number of friends in our empirical model to control for the size of a household's social network. This allows us to rule out the possibility that the adoption decision is driven by the size of the social network, and not by our social learning measure. The result confirms that it is social learning that explains greenhouse adoption, not the size of farmers' social networks.

Among the included household characteristics, only the age of the family head is statistically significant. However, the effects of most household characteristics are consistent with our discussion in Section 4.3. The adjusted R^2 of this regression is 0.81, which suggests that we have included most of the factors that could affect the adoption decision. It also reinforces the idea that our irreversible investment model is an appropriate choice for describing the greenhouse adoption behavior.

For comparison purposes, we present OLS estimation results of the empirical model in Table 5. We find that the social learning effect is still statistically significant. However, the effect is much smaller compared to 2SLS estimation, which suggests that social learning could be negatively correlated with omitted factors in the regression error term. To understand this finding, we note that people who are entrepreneurial in spirit are more likely to have larger social networks and know more prior adopters. At the same time, farmers with an entrepreneurial spirit will consider other possibilities such as starting a small business in the village, which might make it less likely for them to adopt the greenhouse. Our finding suggests that we would underestimate the effect of social learning on greenhouse adoption if we fail to control for endogeneity.

A comparison between Table 5 and Table 3 also shows that many household characteristics with insignificant coefficients from the 2SLS regression become significant in the OLS regression. This is likely due to the potential correlation between household characteristics and the social learning variable, which is effectively removed by using the distance to neighborhood instrument in the 2SLS regression (see Table 2 for the low correlation between the IV and household characteristics).

In Table 6, the interaction terms between the distance to neighborhood and the year dummies are included as extra instruments in the regression. The results are very similar to the results in Table 3, which suggests that the results are robust. Moreover, the extra instruments allow us to use the Hansen-J test to test the validity of the IV (distance to neighborhood).

7 Conclusion

In technology adoption with irreversible investment, agents commonly face two sources of uncertainty. First, the future value of the investment is uncertain. Second, agents have incomplete information regarding the parameters of the process describing the future investment value. In this paper, we model social learning as a way of reducing parameter uncertainty, thus facilitating technology adoption with irreversible investment. We use household-level data from intermediate-technology greenhouse adoption in northern China to test the predictions, with the following main results.

(1) Social learning has a significantly positive impact on greenhouse adoption. Ten more adopters in the farmer's social network increase the probability of adoption by 20 percent, which is an economically significant effect.

(2) The empirical data confirms what we know from the conventional theory of

irreversible investment: higher uncertainty about the future investment value results in less adoption.

(3) Social learning could also affect technology adoption through its influence on the farmer's belief about the expected return on the technology. The empirical data offers some support for this hypothesis.

Our paper also provides an answer to the following question: how could small farmers in developing countries deal with the risk from irreversible investment and incomplete information? Our results suggest that social learning can be an effective solution. Therefore, the policy implication from this paper is clear: when small farmers face technology adoptions such as investing in tube wells or machinery, helping several farmers adopt successfully may be the best way to induce more adoption in their village.

Reference

Abasov, T. M. (2005): Dynamic Learning Effect in Corporate Finance and Risk Management, Unpublished Ph.D. Dissertation, University of California, Irvine.

Banerjee, A., and Duflo, E. (2002): Do Firms Want to Borrow More? Testing Credit Constraints Using a Directed Lending Program, MIT Department of Economics Working Paper No. 02-25.

Bandiera, O., and Rasul, I. (2006): Social Network and Technology Adoption in Northern Mozambique, *Economic Journal* 116, 869-902.

Bertola, G., and Caballero, R. (1994): Irreversibility and Aggregate Investment, *Review of Economic Studies* 61, 223-246.

Besley, T., and Case, A. (1994): Diffusion as a Learning Process: Evidence from HYV Cotton, mimeo, Princeton University.

Brennan, M. J. (1998): The Role of Learning in Dynamic Portfolio Decisions, *European Economic Review* 1, 295-306.

Chinese Agricultural Yearbook (2006): Chinese Agricultural Press.

Conley, T., and Udry, C. (2001): Social Learning through Networks: The Adoption of New Technology in Ghanna, *American Journal of Agricultural Economics* 83, 668-673.

Conley, T., and Udry, C. (2008): Learning About a New Technology: Pineapple in Ghanna, forthcoming in the *American Economic Review*.

Dixit, A. K., and Pindyck, R. S. (1994): *Investment under Uncertainty*, Princeton University Press.

Feder, G. (1980): Farm Size, Risk Aversion and the Adoption of New Technology under Uncertainty, *Oxford Economic Papers, New Series*, Vol. 32, No. 2, 263-283.

Foster, A., and Rosenzweig, M. (1995): Learning by Doing and Learning from Others: Human Capital and Technical Change in Agriculture, *Journal of Political Economy* 103, 1176-1209.

Gennotte, G. (1986): Optimal Portfolio Choice under Incomplete Information, *Journal of Finance* 41, 733-746.

Griliches, Z. (1957): Hybrid Corn: An Exploration in the Economics of Technological Change, *Econometrica* 25, 501-522.

Hassett, K. A., and Metcalf, G. E. (1995): Energy Tax Credits and Residential Conservation Investment, NBER Working Paper No. W4020.

Huang, L., and Liu, H. (2007): Rational Inattention and Portfolio Selection, *Journal of Finance* 62, 1999-2040.

Liptser, R., and Shiryaev, A. (2001): *Statistics of Random Processes*, Springer-Verlag, Berlin.

Manski, C. F. (1993): Identification of Social Effects: Reflection Problem, *Review of Economic Studies* 60, 531-542.

McDonald, R., and Siegel, D. (1986): The Value of Waiting to Invest, *Quarterly Journal of Economics* 101, 707-728.

Merton, R. C. (1980): On Estimating the Expected Return on the Market, *Journal of Financial Economics* 8, 323-361.

Munshi, K. (2004): Social Learning in a Heterogeneous Population: Social Learning in the Indian Green Revolution, *Journal of Development Economics* 73, 185-213.

Nelson, A. W., and Amegbeto, K. (1998): Option Values to Conservation and Agricultural Price Policy: Application to Terrace Construction in Kenya, *American Journal of*

Agricultural Economics 80, 409-418.

Newbery, D. and J. Stiglitz (1981), The Theory of Commodity Price Stabilization, Oxford: Clarendon Press.

Olmstead, A. L., and Rhode, P (1993), Induced Innovation in American Agriculture: a Reconsideration, *Journal of Political Economy* 101, 100-118.

Roumasset, J. (1976): *Rice and Risk: Decision Making Among Low Income Farmers*. Amsterdam: North Holland.

Sunding, D., and Zilberman, D. (2000): Research and Technology Adoption in a Changing Agricultural Sector, Draft for the *Handbook of Agricultural Economics*.

Wan, X. (2000): The Chinese Protection Agriculture Outlook and Trend, *Agricultural Machinery (in Chinese)* 2000, 4-6.

Weinberger, K., and Lumpkin, T. (2005): Horticulture for Poverty Alleviation: The Unfunded Revolution, AVRDC Working Paper 15.

Wooldridge, J. (2002): *Econometric Analysis of Cross Section and Panel Data*, MIT Press, Cambridge.

Xia, Y. (2001): Learning about Predictability: The Effects of Parameter Uncertainty on Dynamic Asset Allocation, *Journal of Finance* 56, 205-246.

Zilberman, D., Sunding, D., Howitt, R., Dinar, A., and MacDougall, R. (1994): Water for California Agriculture: Lessons from the Drought and New Water Market Reform, *Choices* (Fourth Quarter), 25-28.

Figure 1. Greenhouse Diffusion Curve at the Household Level

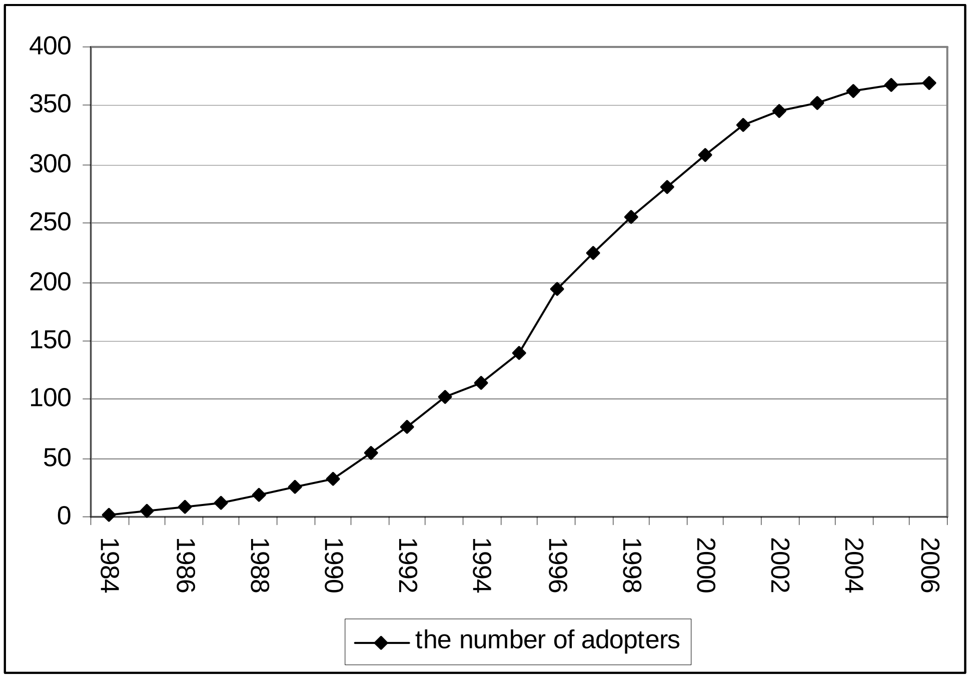

Table 1. Descriptive Statistics: Household Level Data

This table contains the basic household characteristics used in our study. The mean value for each variable is presented with the associated standard error in parentheses. For adopters, all variables are measured in the year before adoption. For non-adopters, all variables are measured in the year before the survey. *** denotes significance at one-percent, ** five-percent, and * ten-percent level.

Basic characteristics	Non-adopter	Adopter	Test of equality of the means (*p*-value)
Social learning within village	4.7 (0.7)	6.9 (0.67)	0.027**
Social learning within village and nearby villages	5.8 (0.8)	8.45 (0.76)	0.018**
Family size	3.7 (0.07)	3.9 (0.06)	0.016**
Farm labor	2.92 (0.07)	2.46 (0.043)	0.01***
Off-farm employment	0.8 (0.054)	0.24 (0.022)	0.01***
Age of family head	46.4 (0.6)	35 (0.46)	0.01***
Education of family head	7.0 (0.17)	7.24 (0.14)	0.25
Off-farm income (yuan)	8420 (649)	1643 (182)	0.01***
Farm size (mu)	5.6 (0.19)	6.01 (0.16)	0.09*
Irrigation ratio	0.80 (0.019)	0.89 (0.013)	0.01***
Major land reallocations since 1980	1.44 (0.067)	0.79 (0.05)	0.01***
Minor land reallocations since 1980	4.29 (0.26)	3.19 (0.19)	0.01***
Grain Land Share (percent)	0.579 (0.282)	0.577 (0.252)	0.92
Ratio of construction costs	3.24 (0.222)	1.95 (0.151)	0.01***
Total number of friends	113.07 (7.564)	98.45 (5.928)	0.12

Table 2. Relation between Distance to Neighborhood and Household Characteristics

Panel A: Correlations

	Distance to road	Distance to nearest market	Education of family head	Age of head of household	Family size	Wealth of household
Distance to neighborhood	-0.036	-0.034	-0.044	-0.004	0.05	0.057

Panel B: Average household characteristics by distance to neighborhood

Distance to Neighborhood	Distance to road (km)	Distance to the nearest market (km)	Education of family head (year)	Age of head of household (year)	Family size (person)	Wealth of household (1,000 yuan)
Below median	1.179	6.952	7.304	40.02	3.776	14.33
Above median	1.152	5.153	6.932	40.37	3.911	16.57

Table 3. Greenhouse Adoption and Social Learning: LPM Estimated by 2SLS
Dependent variable: 1=adopt, 0=not adopt

Explanatory variables	Coefficient	Robust std error	Coefficient	Robust std error
Social Learning				
Social learning within village	0.020	0.005***		
Social learning within village and nearby villages			0.018	0.008**
Conditional mean of market return	-0.349	0.403	-0.299	0.365
Market volatility	-0.002	0.0006**	-0.002	0.0006***
Years of awareness of the technology	-0.010	0.008	0.011	0.08
Output price/input price	0.758	0.273***	0.796	0.279***
Household Characteristics				
Family size	0.011	0.016	0.013	0.048
Age of family head	-0.002	0.002	-0.002	0.001*
Education of family head	-0.003	0.005	0.001	0.004
Off-farm income	-0.012	0.008	-0.012	0.008
Farm size	0.008	0.007	0.009	0.006
Irrigation ratio	0.049	0.054	0.051	0.045
Ratio of construction costs	0.003	0.007	0.003	0.006
Times of major reallocations	-0.021	0.037	0.028	0.037
Times of minor reallocations	-0.004	0.009	-0.002	0.019
Grain share	0.175	0.106	0.159	0.096
Total number of friends	-0.001	0.002	-0.001	0.002
Dummies and constant terms				
Crop dummy	0.164	0.165	0.174	0.156
County dummies	Yes		Yes	
Year dummies	Yes		Yes	
Interaction terms	Yes		Yes	
Constant terms	-0.757	0.279**	-0.758	0.275**
Observations	616		616	
Adjusted R-squared	0.81		0.83	
Durbin-Wu-Hausman Test for Endogeneity	P-value	0.005	P-value	0.004

*** denotes significance at 1%, ** 5% and * 10%.

Table 4. Greenhouse Adoption and Social Learning: First Stage 2SLS Results
Dependent Variable: Social Learning

	Social learning within village		**Social learning within village and nearby villages**	
Explanatory variables	**Coefficient**	**Robust std error**	**Coefficient**	**Robust std error**
Walking time to 20 closest neighbors	-0.099	0.045**	-0.107	0.048**
Conditional mean of market return	3.259	24.07	0.830	24.95
Market volatility	0.013	0.015	0.016	0.017
Years of awareness of the technology	0.697	0.275**	0.811	0.279**
Output price/input price	12.74	18.18	11.79	18.91
Household Characteristics				
Family size	0.290	0.729	0.227	0.774
Age of family head	-0.164	0.071**	-0.150	0.078
Education of family head	-0.389	0.289	-0.316	0.293
Off-farm income	0.084	0.355	-0.100	0.359
Farm size	-0.101	0.387	-0.175	0.380
Irrigation ratio	2.601	4.776	2.708	4.304
Ratio of construction costs	-0.487	0.452	-0.510	0.462
Times of major reallocations	1.555	1.127	2.076	1.089**
Times of minor reallocations	-0.385	0.563	-0.499	0.568
Grain share	0.162	4.465	1.081	4.452
Total number of friends	0.009	0.010	0.010	0.010
Dummies and constant terms				
Crop dummy	-2.298	3.372	-2.991	3.652
County dummies	Yes		Yes	
Year dummies	Yes		Yes	
Constant terms	-4.977	23.68	-5.358	24.97
Observations	616		616	
Adjusted R-squared	0.284		0.331	

*** denotes significance at 1%, ** 5% and * 10%.

Table 5. Greenhouse Adoption and Social Learning: OLS Results
Dependent variable: 1=adopt, 0=not adopt

Explanatory variables	Coefficient	Robust std error	Coefficient	Robust std error
Social Learning				
Social learning within village	0.006	0.002***		
Social learning within village and nearby villages			0.006	0.002**
Conditional mean of market return	-0.324	0.148**	-0.310	0.146**
Market volatility	-0.002	0.0006**	-0.002	0.0006***
Years of awareness of the technology	-0.001	0.002	-0.001	0.002
Output price/input price	0.949	0.303***	0.954	0.300***
Household Characteristics				
Family size	0.014	0.013	0.014	0.013
Age of family head	-0.004	0.002**	-0.004	0.002**
Education of family head	-0.001	0.003	-0.001	0.003
Off-farm income	-0.009	0.004**	-0.009	0.004**
Farm size	0.006	0.003*	0.006	0.003**
Irrigation ratio	0.082	0.042*	0.082	0.043*
Ratio of construction costs	-0.003	0.003	-0.003	0.003
Times of major reallocations	0.003	0.016	0.001	0.016
Times of minor reallocations	-0.009	0.004**	-0.082	0.040**
Grain share	0.177	0.083**	0.172	0.080**
Total number of friends	0.001	0.001	0.001	0.001
Dummies and constant terms				
Crop dummy	0.018	0.053	0.023	0.050
County dummies	Yes		Yes	
Year dummies	Yes		Yes	
Interaction terms	Yes		Yes	
Constant terms	-0.763	0.311**	-0.759	0.307**
Observations	616		616	
Adjusted R-squared	0.91		0.91	

*** denotes significance at 1%, ** 5% and * 10%.

Table 6. Greenhouse Adoption and Social Learning: LPM with Interaction Terms
Dependent variable: 1=adopt, 0=not adopt

Explanatory variables	Coefficient	Robust std error	Coefficient	Robust std error
Social Learning				
Social learning within village	0.021	0.010**		
Social learning within village and nearby villages			0.019	0.008**
Conditional mean of market return	-0.374	0.439	-0.320	0.404
Market volatility	-0.002	0.0006**	-0.002	0.0006***
Years of awareness of the technology	-0.011	0.008	-0.012	0.008
Output price/input price	0.758	0.273***	0.793	0.283***
Household Characteristics				
Family size	0.010	0.016	0.011	0.015
Age of family head	-0.002	0.002	-0.002	0.001**
Education of family head	0.003	0.005	-0.001	0.004
Off-farm income	-0.012	0.009	-0.012	0.009
Farm size	0.008	0.007	0.009	0.006
Irrigation ratio	0.049	0.058	0.050	0.048
Ratio of construction costs	0.004	0.007	0.003	0.007
Times of major reallocations	-0.022	0.038	-0.031	0.038
Times of minor reallocations	-0.003	0.010	-0.001	0.010
Grain share	0.174	0.109	0.157	0.098
Total number of friends	0.001	0.001	0.001	0.001
Dummies and constant terms				
Crop dummy	0.181	0.168	0.194	0.161
County dummies	Yes		Yes	
Year dummies	Yes		Yes	
Interaction terms	Yes		Yes	
Constant terms	-0.773	0.299**	-0.775	0.295**
Observations	618		618	
Adjusted R-squared	0.80		0.81	
Over-Identification Hansen J Test: C-Statistics	P-value	0.20	P-value	0.19

*** denotes significance at 1%, ** 5% and * 10%.